

## 1. Worker Protection Overview

### **U.S. Department of Labor (DOL) Worker Protection Responsibilities**

Through its worker protection agencies DOL enforces laws and regulations governing wages, workplace safety, work flexibility, discrimination, pension benefits and other rights and benefits guaranteed to the nation's workers. At the direction of Secretary Solis, worker protection agencies are undertaking strategies focused on evaluation, innovation, and improved implementation in an effort to reform how they operate. These reform principles are guiding their efforts to more effectively carry out the Secretary's Vision of *Good Jobs for Everyone*.

Central to these efforts is the ongoing evaluation of strategies and measuring outcomes. Two reform principles speak to this expectation directly:

- Establish regular processes for evaluating the success of enforcement and regulatory strategies in helping to achieve the desired outcomes.
- Foster a culture that emphasizes continuous improvement in our regulatory and enforcement programs.

Also central to the responsibility of securing worker protections is accountability to the public. DOL will continue to use performance measures to account to the public how well it is performing the crucial worker protection responsibilities of the DOL mission. To these ends, DOL has developed a *New Approach* to measuring outcomes for the worker protection agencies, which is discussed in this paper.

**Figure 1**

### **Strategies for Reforming Worker Protection at the Department of Labor**

#### **A New Strategic Outlook**

- Target the most egregious and persistent violators.
- Protect the most vulnerable populations while assuring broad-based compliance.
- Regulate to eliminate or reduce the hazards with the broadest and most serious consequences based on sound science.

#### **Evaluation**

- Establish regular processes for evaluating the success of enforcement and regulatory strategies in helping to achieve the desired outcomes.

#### **Innovation**

- Foster a culture that emphasizes continuous improvement in our regulatory and enforcement programs.
- Shift the burden of compliance to the employer or other regulated entity rather than relying exclusively on enforcement interventions. "No more 'catch me if you can' regulation and enforcement,"
- Emphasize corporate-wide or enterprise-wide enforcement and other strategies for leveraging limited resources to accomplish the broadest possible compliance.
- Use openness, transparency, and effective press and communications strategies as means of ensuring broad-based and continuing compliance.

#### **Improved Implementation**

- Implement collaborative enforcement strategies with other DOL, federal, state, and local agencies to leverage limited resources and ensure broad compliance in the whole workplace.
- Establish and maintain an effective partnership between the worker protection agencies, and the Solicitor's Office and the Policy Office in all aspects of regulation and enforcement.
- Impose penalties and other remedies that are consistent with the seriousness of the violation and act as effective deterrents, including identifying and aggressively pursuing appropriate cases for criminal prosecution.
- Partner with worker and community-based organization to identify likely violations and educate workers, small businesses, and others about the law's requirements.
- Involve workers and worker organizations in the workplace to prevent, identify, and remedy violations.
- Provide meaningful compliance assistance and engage the regulated community in designing and implementing compliance assistance.

Figure 2

<b>Department of Labor Worker Protection Agencies and their Missions</b>
<b>Employee Benefits Security Administration (EBSA)</b> The mission of the Employee Benefits Security Administration is to assure the security of the retirement, health and other workplace related benefits of American workers and their families. We will accomplish this mission by developing effective regulations; assisting and educating workers, plan sponsors, fiduciaries and service providers; and vigorously enforcing the law.
<b>Mine Safety and Health Administration (MSHA)</b> The purpose of the Mine Safety and Health Administration is to prevent death, disease, and injury from mining and to promote safe and healthful workplaces for the nation's miners.
<b>Occupational Safety and Health Administration (OSHA)</b> The purpose of the Occupational Safety and Health Administration is to ensure safe and healthful working conditions for working men and women by setting and enforcing standards, and by providing training, outreach, education and assistance.
<b>Office of Federal Contract Compliance Programs (OFCCP)</b> The purpose of the Office of Federal Contract Compliance Programs is to enforce, for the benefit of job seekers and wage earners, the contractual promise of affirmative action and equal employment opportunity required of those who do business with the Federal government.
<b>Office of Labor-Management Standards (OLMS)</b> The Office of Labor-Management Standards (OLMS) enforces standards for union democracy, including provisions for union officer elections; protects union funds and assets; and ensures financial transparency by enforcing public reporting requirements. OLMS also administers employee protections under Federal transit law, ensuring that fair and equitable arrangements are in place to protect the interests of employees affected by such Federal assistance.
<b>Wage and Hour Division (WHD)</b> The Wage and Hour mission is to promote and achieve compliance with labor standards to protect and enhance the welfare of the Nation's workforce.

## **A New Approach to Worker Protection Outcomes**

Newly revised mission statements for DOL's worker protection agencies express their intended outcomes (see adjacent text box). But how well are the agencies achieving those outcomes?

Three key considerations in developing outcome measures for worker protection agencies help to explain the intent behind the *New Approach*. First, the regulatory and enforcement challenge faced by these agencies' regulated populations far outstrips the size of the enforcement staffs. Second, each agency operates within a unique statutory framework and with a distinct set of compliance and enforcement tools. Third, worker protection agencies previously have not applied a uniform approach to evaluating their strategies or measuring outcomes. Each of the above discussed considerations will be explored here.

## **Responsibilities and Resources**

The combined goals of the worker protection agencies cover the areas of health and retirement benefits, wages and hours, safe and healthy workplaces, discrimination and equal employment opportunity, and union elections and financial reporting. Each agency carries out its enforcement responsibilities through a combination of complaint and directed investigations<sup>1</sup> and outreach. With

the exception of MSHA, these agencies reach a fraction of the entities they regulate. Table 1

<sup>1</sup> Each worker protection agencies characterizes its enforcement interventions differently. For example, OSHA conducts *inspections*, OFCCP conducts *compliance evaluations*, and OLMS conducts *audits*. Additionally, there are varying degrees of enforcement options that the agencies may employ. WHD, for example, may conduct a *full investigation* or pursue a *conciliation* depending on the nature of a complaint. For the purposes of this paper, interventions will be referred as either interventions or investigations, when not referencing a specific agency

A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

provides a snapshot of the coverage of DOL’s worker protection agencies alongside their FY 2010 resource levels. Note that full-time-equivalent (FTE) employment levels include non-enforcement employees as well as enforcement workers.

**Table 1. Worker Protection Agencies’ Jurisdiction**

<b>Agency and Resources</b>	<b>Who is Covered</b>
<p><b>EBSA</b></p> <p>FY 2010: \$155 million</p> <p>FTE: 910</p>	<p>EBSA’s jurisdiction includes 150 million Americans – including workers, the unemployed, and retirees – covered by more than 708,000 private retirement plans, 2.8 million health plans, and similar numbers of other welfare benefit plans holding over \$5 trillion in assets; as well as plan sponsors and members of the employee benefits community.</p>
<p><b>OSHA</b></p> <p>FY 2010: \$559 million</p> <p>FTE: 2,335</p>	<p>OSHA covers most of the nation’s private sector employers, either directly through Federal OSHA or through an OSHA-approved state program.</p> <p>Employees who work for state and local governments are not covered by federal OSHA, but have OSH Act protections if they work in those states that have an OSHA-approved state program. OSHA also covers Federal employees, and Federal agencies must have a safety and health program that meet the same standards as private employers. Approximately 72 million workers fall under OSHA’s direct jurisdiction, and an additional estimated 61 million private sector workers are covered through OSHA-approved state programs.</p> <p><b>Who is not covered by the OSH Act:</b></p> <ul style="list-style-type: none"> <li>• Self employed;</li> <li>• Immediate family members of farm employers with no outside employees; and</li> <li>• Workers who are protected by another Federal agency (i.e., MSHA).</li> </ul>
<p><b>MSHA</b></p> <p>FY 2010: \$159 million</p> <p>FTE: 2,425</p>	<p>MSHA inspects all of the nation’s roughly 15,000 surface and underground mines multiple times annually. Each surface mine is inspected at least 2 times a year and each underground mine at least 4 times a year.</p>
<p><b>OFCCP</b></p> <p>FY 2010: \$105 million</p> <p>FTE: 788</p>	<p>OFCCP regulations cover approximately 90,000 to 100,000 Federal contractor establishments with various contract award thresholds tied to specific statutes. Further, contractors and subcontractors with a non-exempt federal contract or subcontract of \$50,000 or more, and 50 or more employees must develop a written affirmative action program that is designed to ensure equal employment opportunity for minorities and women, and sets forth specific and action-oriented programs to which a contractor commits itself to apply every good faith effort.</p> <p>Additionally, all firms that have a nonexempt Federal contract or subcontract in excess of \$50,000 and 50 or more employees must have an affirmative action program for individuals with disabilities. Firms with contracts of \$100,000 or more (or \$25,000 or</p>

A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

Agency and Resources	Who is Covered
	more if entered into before 12/31/2003) and more than 50 employees must have an affirmative action program for covered veterans.
<b>OLMS</b>  FY 2010: \$41 million  FTE: 269	OLMS covers private sector labor unions with different filing requirements based on the amount of annual receipts. Roughly 23,000 unions file disclosure reports annually, including approximately 19,000 locals, 2,200 intermediates, and 174 national unions.
<b>WHD</b>  FY 2010: \$243 million  FTE: 1,582	WHD is responsible for administering and enforcing of a wide range of laws which collectively cover virtually all private and State and local government employment. This represents an estimated 8 million workplaces and 120 million workers.

**Statutory Requirements and Strategies**

Each of the worker protection agencies strives to maximize resources through the most effective and efficient use of their enforcement activities. Agencies prioritize and target areas where violations are most likely to be found. These efforts complement their complaint-driven investigative work. The list below highlights the agencies' current enforcement priorities.

- EBSA: General targeted enforcement
- MSHA: Targeting specific hazards, but at all mines
- OSHA: High hazard industries; vulnerable workers
- OFCCP: Targeting (in an administratively neutral way) based on likelihood of finding violations
- OLMS: General targeted enforcement
- WHD: Vulnerable workers; low-wage industries

DOL worker protection agencies have different methods for identifying likely violators, including complaints, intelligence gathering, data from required filings, recidivism data, or any combination thereof. Additionally, worker protection agencies have differing parameters within which they can pursue violators and potential violators, return to prior violators, and different legal remedies for entities found in violation of associated statutes.

In their efforts to maximize resources by targeting those regulated entities in which violations are most likely to be found, it is important that a worker protection agency's success not be defined solely or primarily by their rate of success in identifying violators. An agency's success in identifying those entities at a high rate is not the same thing as achieving that agency's

A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

intended objectives or outcomes. As such, worker protection agencies are responsible for the entire universe of entities they regulate, and not just those that they interact with in a given year. This theme of a broad conception of accountability – being responsible for more than just those entities a worker protection agency touches – is recurring in this paper and is the imperative behind the *New Approach* to measuring worker protection performance.

## **2. An Overview of the *New Approach***

### **Performance Measurement and Public Accountability**

Through the preparation of the 2011 – 2016 Strategic Plan, DOL is developing performance measures for its agencies that meet the standard of measuring *outcomes*. While output measures gauge activities, or the goods and services an agency produces or conducts, outcome measures gauge success in achieving the intended result of carrying out a program or activity. Outcome measures are often separated into measures of *intermediate outcomes* and measures of *end outcomes*. Examples for measures of end outcomes include fatality, injury, and illness rates for the Mine Safety and Health Administration (MSHA) and the Occupational Safety and Health Administration (OSHA). Measures of intermediate outcomes are leading or indirect indicators of outcomes, often chosen to provide early signs of changes in behavior. While fatality, injury and illness rates represent the ultimate outcomes of concern for OSHA and MSHA, compliance with specific health and safety standards associated with them represent important intermediate outcomes because the failure to adhere them can create the preconditions for accidents or adverse health exposures. Similarly, the discrimination rate among all Federal contractors and subcontractors would be an end outcome for OFCCP, and contractors' compliance rate with equal opportunity laws and regulations is an associated intermediate outcome. Failure of contractors to comply with equal opportunity laws and regulations can lead to discrimination. Importantly, outcome measures should cover the universe of regulated entities, not just those directly reached through an agency's enforcement activities.

Through the preparation of the 2011 – 2016 DOL Strategic Plan, DOL is moving toward developing performance measure schemes for its worker protection agencies that meet this standard of measuring outcomes. While there is some expense and great effort required to implement these measures, DOL's accountability to the public – to report the extent to which employers deliver on the rights and benefits guaranteed to the nation's workers and to use its resources for strategies that achieve the greatest effect – is vitally important.

### **Committing to a New Way of Measuring Success**

The *New Approach* represents several important commitments on the part of worker protection agencies.

1. To develop strategies that aim to affect change beyond immediate interventions.

## A New Approach to Measuring the Performance of U.S. Department of Labor Worker Protection Agencies 6/28/2010

2. To report on compliance<sup>2</sup> levels that represent the universe of regulated entities—not just those where an intervention occurred.
3. To continually evaluate whether strategies are leading to improved outcomes.
4. To use both routine quantitative analysis and program evaluations to assess both the extent to which the agency is achieving its outcome goals and the effectiveness of the particular strategies pursued.

These commitments are tied to the performance measures and methods that are driving the *New Approach* and that have shaped the strategic planning process for worker protection agencies. Under the *New Approach*, worker protection agencies will be aligned by a similar set of performance measurement methods that reinforce these commitments while accounting for the considerations discussed previously. Additionally, DOL has developed new methods for target-setting and establishing trends which should contribute to the aforementioned culture of continuous improvement. The following sections will explain the worker protection outcome measures and the methods by which DOL will hold the agencies accountable for improvement.

### Using Sampling to Estimate Performance Results

DOL's *New Approach* to measuring the performance of its worker protection agencies is centered on the idea that outcome measures should characterize the entire population of regulated entities. To accomplish this, DOL worker protection agencies will rely on the statistical techniques of random sampling and stratified random sampling and the application of social science research methods to collect and analyze performance data. Through this approach, DOL can make credible estimates about the behaviors and conditions within groups of regulated entities that may be in the thousands or hundreds of thousands by systematically collecting data from small samples – typically numbering only in the hundreds – of those entities. In statistics, such generalizations about a population from the collection of sample data are called inferences.<sup>3</sup>

---

<sup>2</sup> What constitutes compliance differs by agency and sometimes by measures within agencies. EBSA's compliance measure, for example, tracks compliance with the civil provisions of its governing legislation, the Employee Retirement Income Security Act of 1974 (ERISA). OSHA's measures track the absence of compliance, violations, or more specifically, "serious, willful, repeat violations." Less grave violations do not meet the specific criteria for serious, willful, repeat violations, are therefore not counted as such.

<sup>3</sup> Reliable inferences can be drawn only when the samples are drawn so that every entity in the population has a positive, known probability of being chosen for the sample. Most commonly, samples are drawn so that each regulated entity has an equal chance of being selected into the sample. Such a sample is representative of the population from which it is drawn, and ensures that average outcomes (e.g., compliance rates) in the sample are indicative of average outcomes in the population. A more complex approach is stratified random sampling, where the probability of selection into a sample varies across classes of entities. This approach can be useful when subgroups within a population have distinct and diverse characteristics. Statistical analyses of stratified random samples are more complex, but also permit reliable inference. Moreover, stratified random samples can also permit inferences about average outcomes for subgroups that might represent a small fraction of the overall population but that are of distinct interest.

## A New Approach to Measuring the Performance of U.S. Department of Labor Worker Protection Agencies 6/28/2010

The outcome measures from the *New Approach* accomplish another important strategic planning objective – holding agencies accountable for the universe of entities they regulate. Historically, many DOL performance measures have focused on evaluating the efficiency or results of program activities. For a program in which success is defined by the outcome achieved for program participants, such measures are appropriate. For worker protection agencies, however, measuring the results of only those investigations, inspections, or audits they conduct provides a very limited view of whether an agency is fulfilling its mission. With the exception of MSHA, worker protection agencies touch only a fraction of the universe they regulate. Measuring the success of those few interventions alone provides little insight on the actual levels of compliance or the agencies' affect on compliance throughout the universe of regulated entities.

### **The Promise of a Broader Understanding of Accountability**

Many worker protection agencies have taken targeted approaches to enforcement and measured success based on the findings of those enforcement actions alone. EBSA, for example, previously defined successful performance based on finding high violation rates in the subset of benefit plans for which cases were opened, as this indicated that enforcement efforts were well targeted. Even if EBSA were successful in identifying and correcting fiduciary violations in 100% of the cases it opens, this measure would only inform the public about EBSA's success with the benefit plans where it conducted its work. This is valuable information, and corrected fiduciary violations represent the protection or restoration of medical, retirement, or other benefits. However, these measures fail to identify the broader conditions within EBSA's sphere of responsibility. In particular, they are uninformative about whether EBSA has achieved high compliance rates for the overall population of regulated plans. Additionally, such measures, especially when tied to targeted annual performance improvements, create perverse incentives for seeking improvements within a narrow sphere potentially at the expense of broader worker protection. Even before any analysis is done, merely reporting estimated outcomes, such as compliance rates, for a universe or population represents greater public accountability than establishing and reporting against targets for measures that are limited to the entities where a worker protection agency investigates, inspects, or audits.

The Wage and Hour Division (WHD) serves as a model for the *New Approach*. WHD has adopted methods for not only measuring the compliance levels of industries, but has developed strategies to expand its impact on compliance beyond those workplaces it investigates.

- **Industry Structure:** WHD has conducted numerous studies examining the structures of various industries to identify more strategic approaches to enforcement. Understanding how the structure of industries (e.g., franchises or company-owned, supply chain) relates to the compliance behavior of employers within industries helps WHD make informed decisions about how those industries are affected by various compliance and enforcement tools. This, in turn, allows WHD to make more effective use of its resources.

A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

- Deterrence Effects: Exploiting deterrence effects is a fundamental feature of how WHD approaches its enforcement efforts. Through deterrence effects, WHD strives to influence the behavior of employers in the same industry and/or the same geographic area. Deterrence can be achieved through several means, including expanding the scope of investigations, increasing the use of penalties, litigation, and using media and other disclosure and transparency tools to broaden the impact of its efforts.

WHD will be applying the *New Approach* to report on compliance rates for targeted low-wage industries.

In Table 2 below, the universe for each agency is defined alongside the measures that will allow them to draw inferences about covered populations. The table is meant to highlight the different applications of the *New Approach*. As discussed earlier, agencies prioritize their enforcement work based upon those areas where violations are most likely to be found. Those priorities are reflected in the populations covered by the measure. Using these measures, agencies can assess their effect on compliance, both broadly and with respect to investigated entities. Of course, with this broader conception of compliance, it will be important to keep in mind that outcomes are affected by a variety of influences beyond the specific actions of the regulatory agency. Thus, an important component of the *New Approach* to accountability must be a recognition that, while these measures are informative about the success or failure of an agency's efforts, the agencies cannot be held solely accountable for changes in the outcomes.

**Table 2. Outcome Measures' Coverage**

<b>Agency</b>	<b>Universe</b>	<b>Measures</b>	<b>Methods Used and Measures' Coverage</b>
<b>EBSA</b>	All employer provided retirement plans, health plans, and other welfare benefit plans	Compliance rate and Investigation effect	A random sample of entities (mostly benefit plans), grouped into classes (health, defined contribution pension, etc.). Sampling method will match, proportionately, classes of current EBSA workload. Sampling method will be used to estimate performance of approximately 90% of benefit plan universe.
<b>MSHA</b>	All surface and underground, metal and non-metal mines	Fatal injury rate. Injury and illness rate beginning in 2013	Entire universe of mines covered. No sampling used.
<b>OSHA</b>	The majority of U.S. employees, except miners, transportation workers, domestic workers, some public employees, and the	Violation rate, recidivism rate, and investigation effect	Two groups of emphasis covered through sampling-based performance measures. Sampling will be used to estimate performance,



A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

Agency	Universe	Measures	Methods Used and Measures' Coverage
	self-employed		separately, in high hazard manufacturing and construction projects over \$5 million.
<b>OFCCP</b>	Federal contractors and sub-contractors	Compliance rate, discrimination rate, and compliance rate of previously audited	Sampling used to estimate performance of entire universe of Federal contractors.
<b>OLMS</b>	Private sector labor unions	Proxy measures of chronic delinquent filing rate and recidivism for chronic delinquent filing	Entire universe of private sector labor unions covered. No sampling used.
<b>WHD</b>	Fair Labor Standards Act (FLSA) covers all employees engaged in or producing goods for interstate commerce and all employers with annual gross sales over \$500,000. Also covers public sector employers, employers operating hospitals, nursing homes, and educational institutions.	Compliance rate and compliance of previous violators	Four groups of FLSA emphasis covered through sampling-based performance measures. Sampling will be used to estimate performance, separately, of H-2A certified employers, and the janitorial, construction, and hotel/motel industries. Family and Medical Leave Act is not covered by sampling-based measures.

### **3. Performance Accountability**

#### **Determining Improvement in DOL Worker Protection Agencies**

Historically, for its publicly reported measures, the Department of Labor (DOL) has assessed its performance against targets that were both established and reported against annually. For most measures, that approach will still be in effect for agencies reporting on goals included in the 2011 – 2016 Strategic Plan. For some performance measures, however, establishing and reporting against annual performance targets is not appropriate to the nature of the outcomes being measured.

Reporting target attainment for individual measures or portraying a worker protection program's performance as a function of the number or percentage of targets attained can interfere with other, more substantive analyses and may, indeed, mask improvements or declines in performance. The development of the measures identified in this paper grew, in part, out of a concern that gauging success in worker protection agencies by using measures that are amenable to annual target setting tends to over-emphasize measures that are tied to specific strategies and activities at the expense of measuring outcomes. Performance measures

## A New Approach to Measuring the Performance of U.S. Department of Labor Worker Protection Agencies 6/28/2010

that define success by activities, even successfully implemented activities, serve to reinforce those activities. Such measures will never provide specific evidence to either support or refute that an agency is achieving the objectives (outcomes) it was established to achieve.

DOL is striving to sever the tie between existing methods of measuring performance and assumptions that current activities will lead desired outcomes. Those existing measurement methods also confuse improved outcome performance with meeting targets that are tied to activities. The *New Approach* to measuring worker protection agency performance is designed to provide the data that will allow the Department to do two things: 1) provide a more accurate and definitive assessment of agency outcome performance, and 2) test the connections and assumption between existing activities and the achievement of desired outcomes<sup>4</sup>

For many, establishing and reporting against targets are cornerstones of government accountability. Agencies use annual reporting to demonstrate progress on major goals and to draw conclusions about goal achievement. For most programs, their measures are suitable for this type of reporting. For example, knowing that a targeted percentage of WIA beneficiaries entered employment represents the key outcome of WIA-provided training. Reducing the number of lost production days for the Federal Employees' Compensation Act demonstrates that federal workers are timely returning to employment. A targeted rate of universal compliance, for both practical and theoretical reasons, however, is ill-suited for this type of reporting.

Thus, DOL will move away from numeric performance targets for the outcome rate measures associated with the *New Approach* (i.e., compliance or violation rate for most worker protection agencies and discrimination rate for OFCCP) for external reporting purposes. Given the nature of these worker protection outcome measures, the traditional GPRA-based approach to target-setting would undermine their very purpose. Outcome rates require time, analysis, and contextual data and performance information to determine whether and the extent to which success was truly achieved. Alternatively, attempting to pinpoint a single numeric target or even a range fails to recognize the sustained effort required to truly change behaviors.

### **Increased Results Analysis and Performance Trends**

Worker protection agency activities are one of many factors influencing compliance rates. Their work is performed in a vast, complicated, and continuously evolving regulatory environment. An agency's ability to change overall compliance rates in the near- or long-term is contingent upon many factors outside its control. To determine a worker protection agency's influence, rigorous program evaluation methodologies are needed. However, conducting annual program evaluations to assess a program's impact on goal achievement is not a realistic undertaking for these agencies. Decision-makers require performance goals that aim for the broader effect, but

---

<sup>4</sup> The importance of rigorously testing the connection between activities and outcomes is explored further in the Appendix, in the *Using Performance Measures to Drive Strategy* section.

A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

allow them to gauge progress differently from previous performance measures. To that end, DOL will evaluate performance trends.

Reporting results for each of the new performance measures will be accompanied by increased analysis of the trends, including the direction and extent of movement, and as more data are collected, the reasons underlying the reported level of performance. The next section will explore DOL's approach to performance targets and how DOL will hold its agencies accountable to report not just levels of performance, but also whether or not strategies are leading to goal achievement and a commitment to change strategies as evidence warrants.

For measures within the DOL FY 2011 – 2016 Strategic Plan that indicate a compliance, violation, or discrimination rate, agencies will target for either an increasing or decreasing trend. The magnitude of the desired increase or decrease is not predetermined (not targeted at a specific percent increase or decrease). Any degree of improvement will be assessed within the context of prior years' results. Performance results for these measures will be reported in the DOL Annual Performance Report (APR). Those results will be evaluated through rigorous analysis of the annual results and trends, including:

- Discussion of the direction and extent of the change in results.
- Detailed results of the major outputs and strategies that are most closely associated with the outcome rate results, as well as external factors that are known to influence the outcomes.
- Describing what is known regarding the connection between agency outputs and the outcome rates, and whether the connections are grounded in empirical evidence or based on experience and hypotheses.
- Where evidence is lacking, a plan for testing those hypotheses will be provided and then reported on in subsequent APRs. Evaluations, discussed later in this plan will be encouraged as a tool for extracting empirical evidence on the connection between strategies, outputs, and outcomes.

The DOL strategic plan will commit agencies to improving performance, as measured by outcome data trends, analysis of annual performance, and the corresponding outputs tracked internally and through the budget process. Agencies will be accountable for achieving performance goals by adjusting strategies based on that evidence.

Trend analysis strikes a balance between holding agencies accountable to goals while promoting transparent assessments of strategies and outcomes. As noted earlier, developing measures tied to strategies assumes reaching those performance targets will lead to the desired outcome. By contrast, aiming for improvements in the outcome rate promotes adopting the strategies that are empirically tied to improving that rate.

A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

**Population Performance Measures Can be Used to Strengthen the Targeting of Likely Violators**

DOL's worker protection agencies have often selected sites for inspection, investigation, or audit on a targeted basis. The targeting of specific sites or entities is usually based on information available to the agencies through required reporting or available information and is assumed to tie to increased likelihood of findings violations. Employing targeted investigations is sound management where the ratio of investigative resources to the sphere of regulatory responsibility is small. Such targeted investigations have proven useful in recovering or restoring benefits (EBSA), identifying unsafe or unhealthful working conditions (OSHA), discriminatory hiring or promotion practices (OFCCP), and labor standards violations. It may seem counter-intuitive, therefore, to displace resources now dedicated to targeted investigations to randomized investigations.

Diverting any resources from directed enforcement for the purposes of a measure, particularly for agencies already operating with limited resources, was a concern for many DOL agencies. Without this data, however, agencies have no evidence-driven methods for evaluating the success of their strategies or their targeting methods. Absent overall compliance or other outcome rates, how can a worker protection agency determine whether their targeting efforts or strategies affect those outcomes? Gathering information on compliance will become a routine part of how these agencies conduct business, and will eventually create an empirical foundation that will lead to the retention of some targeting approaches and the refinement or termination of others. In this light, it is not a diversion of resources but rather an essential dedication of resources that the Department has long neglected.

#### **4. Program Evaluation**

##### **Rigorous Program Evaluation Complements Performance Measurement by Identifying Agency Impacts and by Validating Agency Strategies and Activities**

In addition to failing to provide the public suitable information on the outcomes of worker protection agencies, using outputs and/or processes to measure and report performance assumes that the outputs, and their associated strategies, activities, and processes are positively tied to outcomes. These strategies and activities include inspections, public outreach and branding campaigns, back wages, civil monetary penalties, and litigation. These are all strategies that agencies believe to be effective in achieving compliance, but agencies cannot continue these practices; much less use them to report performance, without supplying empirical evidence to demonstrate their positive relationship to desired worker protection outcomes. Rigorous program evaluation will play a crucial role in providing the empirical evidence that will validate the use of specific agency strategies and activities to improve outcomes.

### **Outcome Performance Measures Establish Data Needed for Program Evaluation**

Population outcome measures based on random samples allow for estimates of compliance rate and other related outcome levels in a given year, comparison of compliance rates or related outcome levels from year to year and allow for determining trends over lengthier periods of time.

Performance measures also serve to establish data collection ground work that will allow for program evaluations that will serve to better understand the relationship between agencies' strategies and their performance outcomes and to more accurately assess agencies' impacts. Without data on compliance or other outcome rates, the ability to draw empirically-based conclusions about the effectiveness of existing strategies is severely limited. Measuring the performance results of strategies can serve to support evaluation efforts if executed within the context of collecting data on strategies. To serve this end, performance reporting should be carried out in the spirit of evaluation and not under the pass/fail conditions encouraged by reporting performance against targets.

### **Program Evaluation Identifies Program Impacts**

Even with these more sophisticated approaches to measuring worker protection outcomes, and given the measures of uncertainty and precision that will accompany the sample values used to estimate what's happening within and entire population, the ability to isolate the effects of an agency's activities or to measure the impact of an agency's activities (what would have happened, all else being equal, in the absence of the agency) requires a program evaluation. For any given Federal program's reported performance, there are several factors (external independent variables) over which the agency has neither jurisdiction nor control, that will affect the level of performance. Separating out the influence of these external independent variables on an agency's performance so that you can have a clearer idea of the specific influence of the agency is done through program evaluation.

An example of such an external independent variable affecting agency performance is the influence of commodity demand on mining safety. Higher demand for a given mineral, for example, will result in market responses to meet that demand. Market responses might include increased production at existing mines or the creation of new mines; either of which could have adverse effects on mine safety. Increased production at existing mines may result in greater fatigue and accompanying inattention to sound safety practices. Getting a mine up and running quickly may take precedence over and compromise the adoption of sound safety practices in an effort to quickly to mine the commodity prior to a return to normal demand. Even with the methods and measures described above DOL will not be able to distinguish the influence of such a phenomenon (and any number of other variables) from the effects of agency efforts on performance without rigorous program evaluations of the latter.

## **Accountability for DOL Worker Protection Agencies**

DOL is establishing a foundation of outcome-oriented performance measures for its worker protection agencies that go beyond measuring just those entities DOL targets to measuring the sphere of its responsibility. It will build on that foundation by developing rigorous program evaluation plan to test worker protection strategies and assess the impacts of its programs. Taken together, outcome measures and program evaluation create a new framework for DOL worker protection accountability that will report to the public our stewardship over wages, workplace safety, work flexibility, discrimination, pension benefits and other rights and benefits guaranteed to the nation's workers.

### **5. Appendix**

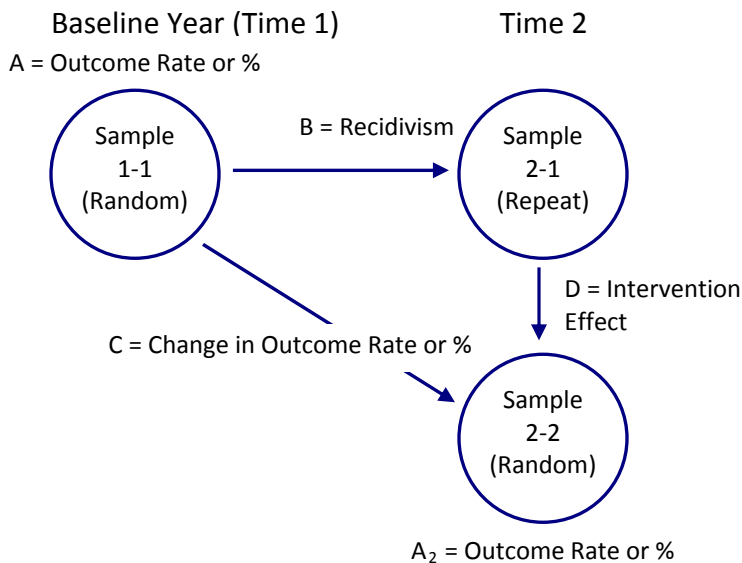
#### **Analytic Techniques to be Used in Worker Protection Agencies to Determine Outcomes through Performance Measurement and to Identify Impacts and Validate Strategies**

##### **Measurements of Outcomes and Changes in Outcomes Over Time**

Through collection of a representative sample, a given DOL worker protection agency will estimate the performance as expressed as either an intermediate outcome or an end outcome. In this paragraph, generating measures of outcome rates  $A$  and  $A_2$ , through taking random samples 1-1 and 2-2, is explained. For some agencies, a sample will be used to determine the level of compliance with that agency's applicable standards, regulations, and laws. The level of compliance, or a compliance rate, will become a measure which that agency will report to the public on an annual or biennial basis. As mentioned above, compliance is an intermediate outcome that provides information about behaviors of regulated entities that are tied to the desired end outcomes, or the ultimate intended result for which the worker protection agency was created.

An example of the use of a compliance rate for reporting performance is the Employee Benefits Security Administration (EBSA). EBSA will report compliance with the provisions of the Employee Retirement Income Security Act (ERISA), which is an intermediate outcome measure that is a reflection of EBSA's performance in meeting its end outcome of securing health and retirement benefits, as expressed in DOL's strategic and outcome goals. In terms of the above model, a performance measure of the compliance rate (% in compliance with the provisions of ERISA) in the baseline year (Time 1) is  $A$ , using random sample 1-1. The same performance measure for a subsequent year (using random sample 2-2 at Time 2), is denoted as  $A_2$ . An additional indicator of performance, denoted as  $C$ , is the change in compliance rate from Time 1 to a subsequent year (Time 2). This is a comparison from the outcomes measured in random sample 1-1 to the outcomes measured in random sample 2-2. These measures, of course, are not limited to intermediate outcomes, such as compliance. In addition to compliance, an end outcome, such as the Office of Federal Contractor Compliance Programs (OFCCP) measuring the discrimination rate of Federal government contractor and subcontractor employees at Time 1 and Time 2, and the decline between those years ( $A$ ,  $A_2$ , and  $C$ , respectively), can also be used.

## The Model



### Recidivism and Intervention Effect

Using the approach identified above also allows worker protection agencies to gauge the effectiveness of their corrective actions and other interventions in two ways: recidivism and intervention effect<sup>5</sup>. Recidivism<sup>6</sup> can be defined as the recurrence of an undesirable behavior. Within the worker protection context, recidivism could be continued non-compliance in the above EBSA example or continued discrimination in the above OFCCP example. Measures of recidivism require the passage of time after Time 1 (Baseline year). In addition to the random sample taken at Time 2 (2-2) that is used to determine changes in the outcome level or rate within the universe from Time 1 to Time 2, the worker protection agency can return to the exact same entities previously sampled. This return is depicted as sample 2-1 in the above model. For compliance, the recidivism rate is the percent of those previously non-compliant (when sampled at Time 1) that are also non-compliant at Time 2. Recidivism is a measure of agency performance because between Times 1 and 2, there are follow-up activities, such as corrective actions and penalties, at those regulated entities that were non-compliant at Time 1. Recidivism is, therefore, a measure of the effectiveness of those follow-up activities and also of the way those activities are carried out. The more effective the follow-up, in both substance and implementation, the lower the recidivism rate will be. In order for the follow-up activities

<sup>5</sup> While there is a comparison between two samples for the intervention effect measure – one consisting of entities recently investigated and one not – this is not to suggest that this measure represents the isolated effects, or impact of an agency's intervention. Determining impact requires the methods discussed previously in the *Program Evaluation* section.

<sup>6</sup> Some DOL agencies will report a recidivism rate, or the recurrence of violations of those entities that have previously been investigated and found in violation. Other agencies will report on the behaviors of previously investigated entities through the inverse of recidivism, which can be expressed as the compliance rate of previously investigated entities. This choice is usually driven by methodological concerns, and will be clearly identified by the wording of the measures.

A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

to have had an effect requires the passage of time, such as two years, between Time 1 (sample 1-1) and Time 2 (sample 2-1).

Time 2 includes two samples: the follow-up visit to all those entities previously sampled at Time 1 (sample 2-1) and the new random sample drawn from the universe (sample 2-2) for the purposes of determining both the outcome level at Time 2 and the change in outcome level between Times 1 and 2. A comparison between these samples is another measure of performance: intervention effect, which is designated as D in the above model. There was some intervention at all the regulated entities in sample 2-1. Some of the names worker protection agencies give for these interventions are audit, investigation, inspection, etc. In addition to the basic level of intervention, there are the follow-up interventions at those entities found to be non-compliant, possibly including fines, penalties, required training, restitution, etc. Because those entities receiving follow-up intervention at Time 2 were originally selected on a random or stratified random sample at Time 1, the comparison between the outcome levels of sample 2-1 and sample 2-2 (also selected at random) is the effect of the worker protection agency' interventions.

### **Using Performance Measures to Drive Strategy**

From a management perspective, one of the weaknesses of using output measures is that it creates incentives to continue the existing and in many cases untested practices, activities, and strategies. The relationship between practices, activities and strategies and the production and level of outputs is not is not a complicated one. If a worker protection agency measures and reports its effectiveness in terms of outputs, such as numbers of inspections, investigations, or audits (an output not recently been used for DOL performance reporting), it creates incentives for the agency to continue engaging in the behaviors tied to increased investigations. What remains unknown under this scenario is the relationship between the existing practices, activities, and strategies with the end outcome, or the ultimate intended result for which the worker protection agency was created. In such instances, the positive relationship between what the worker protection agency is now doing and improvements in end outcomes is assumed.

It is important to test such assumptions, but to test them empirically. Establishing and reporting against outcome measures is the first step to questioning and testing the assumptions connecting current practices, activities, and strategies and end outcomes. Establishing outcome measures and targeting for improvement allows analysis exploring the connection between what an agency does and its outcome-related performance. Such analysis can provide a first level of empirical evidence supporting existing practices or identifying new strategies. This analysis should be accompanied by rigorous testing of practices, activities, and strategies against outcomes that can be achieved through program evaluations (discussed below).



A New Approach to Measuring the Performance of  
U.S. Department of Labor Worker Protection Agencies 6/28/2010

**Worker Protection Performance Results Will be Estimates and will Only Account for External Factors in Limited Ways**

While more sophisticated than previous approaches to measuring the performance DOL worker protection efforts, a certain level of humility is warranted. Each measure result will actually be an estimate which will be accompanied by measures of precision and uncertainty. Additionally, in order to isolate the impact of a worker protection agency's efforts requires program evaluation.

Confidence Intervals (Statistical Precision)

The measure of precision is called a confidence interval, often referenced in terms of a margin of error. Statistical estimates of performance within a population will be accompanied by a margin of error, which is both added and subtracted from the estimated level of performance to create a range (confidence interval) within which the true population value (level of performance) is found. For demonstration purposes, a 3 percentage point margin of error accompanying an estimated discrimination rate of 18% would mean that the population discrimination rate is 18% + or – 3 percentage points; or that the true population discrimination rate falls between 15% and 21%.

Confidence Levels

The measure of uncertainty is called the confidence level. A confidence level is a pre-selected value that, simply expressed, indicates how willing you are to be wrong. Commonly accepted confidence levels are 90%, 95%, and 99%, which correspond to probabilities of being wrong of .10, .05, and .01, respectively. "Wrong," is defined as the true population discrimination rate actually being outside of the margin of error. Using the example above, if the sample discrimination rate is 18%, the margin of error is 3 percentage points, and the confidence level is 95%; indicates that there is a .05 probability (5% chance) that the true population discrimination rate falls outside 15% - 21%.