Environmental Compliance and Enforcement Measurement: Why, What, and How?

Shelley H. Metzenbaum
President of The Volcker Alliance

Discussion Paper for the
Penn Program on Regulation's
International Expert Dialogue on
"Defining and Measuring Regulatory Excellence"

March 19-20, 2015

University of Pennsylvania Law School

Environmental Compliance and Enforcement Measurement: Why, What, and How?

Shelley H. Metzenbaum, President of The Volcker Alliance
With assistance from Peter D. Morrissey and Kaeleigh R. Forsyth

Abstract

This chapter discusses the crucial role of measurement in environmental compliance and enforcement efforts. When measurement is wholly integrated into environmental compliance and enforcement management decisions, it can be transformative – allowing regulatory agencies to achieve better environmental outcomes, higher compliance rates, improved fairness, a better understanding of regulated entities and environmental problems, and a higher return on investment. Without a strong system of outcomes-focused measurement, regulators run a high risk of being wasteful and ineffective.

Measurement need not be difficult, but it does require clear thinking about objectives and how to assess progress on those objectives, as well as a commitment to start to measure and continuously use data to find ways to improve. The chapter uses three examples to illustrate the value and feasibility, as well as the variety, of useful compliance and enforcement measurement approaches: water quality data to clean up of a river, maps to reduce wetlands loss, and self-reported compliance information complemented by government-conducted audits of randomly sampled facilities to increase compliance. It also explores why, what and how to measure as well as measurement challenges.

Keywords: environment, compliance, enforcement, measurement, regulation, performance measurement, return on investment, fairness, goal-setting, transparency, outcomes, deterrence, mapping, technology, evidence, improvement science, EPA, Massachusetts, United States

I. Introduction
II. Objectives
III. Case 1: Cleaning Up the Charles River
IV. Case 2: Reducing Wetlands Loss in Massachusetts
V. Case 3: Counting the Frequency and Noting the Nature of Non-Compliance
VI. What, Why and How to Measure
VII. Measurement Challenges
VIII. Conclusion
Works Cited

Introduction

This chapter discusses the role of measurement in environmental compliance and enforcement efforts. It begins with a consideration of the objectives of environmental compliance and enforcement because discussion of any and all government programs and associated activities should start there -- with a clear understanding of the objectives of government action. The chapter next provides three examples illustrating the power and feasibility of useful compliance and enforcement measurement, describing a variety of measurement approaches that can be highly effective.[1]

- **Cleaning Up the Charles River.** First, the Charles River clean-up starting in the mid-1990s illustrates the value of using geographically, temporally frequent, outcome-focused measurements to assess progress toward a goal and determine where to focus. Fresh, frequent outcomes data not only reveal previously unknown environmental problems and suggest how to target resources to more consequential problems. The data also give government fast feedback to determine which actions to continue and possibly promote for broader adoption and which actions to fix or stop because they are not working well. In addition, the Charles River example illustrates how clearly stated achievable-but-stretch goals (combined with frequent measurement) encourage innovative thinking and lead to higher returns on government spending. It also shows how goals, with frequent measurement, can enlist ideas, expertise, and resources from outside the goal-setting organization, accelerating progress on a goal.
- **Reducing Wetlands Loss in Massachusetts.** The Massachusetts wetlands enforcement example shows why it is so important to be clear about program objectives and measure progress toward those objectives. Measurement not only helps government determine if it is making progress toward its objectives, it also informs resource deployment decisions. Beyond that, the wetlands example vividly illustrates how failure to measure progress toward objectives can result in government spending significant time and money doing activities that fail to achieve their intended impact or achieve them far less well than an alternate set of actions would do.
- **Counting the Frequency and Noting the Nature of Non-Compliance.** Lastly, the Environmental Results Program (ERP) demonstrates the value of having individual regulated parties self-report performance and use of key practices, combined with government-prepared compliance-assistance manuals and post-reporting auditing and analysis. Beyond that, the ERP analytic approach – involving more consistent, systematic coding and entry of the status of compliance with individual requirements – can be applied to traditional government inspection findings with great benefit. Coding, collecting, analyzing, and displaying inspection findings from all regulatory programs can spotlight the

[1] This chapter draws from my experience as Undersecretary of the Massachusetts Executive Office of Environmental Affairs, Associate Administrator of the U.S. Environmental Protection Agency (EPA) for Regional Operations and State/Local Relations, and Executive Director of the Environmental Compliance Consortium (state environmental compliance and enforcement officials who came together to find better ways to measure and manage their programs), plus my experience as Associate Director of the U.S. Office of Management and Budget for Performance and Personnel Management.

most prevalent non-compliance problems and inform the design of more tailored, and therefore cost-effective, compliance-increasing actions.

The chapter then explores why, what and how to measure, explaining why effective measurement is essential to good government, useful both to regulators and regulated parties, and need not be complicated or burdensome. A brief discussion of compliance and enforcement measurement challenges follows.

I hope that this chapter does more than introduce readers to this important subject. I hope it also convinces you to contribute to more robust, useful measurement in whatever role you play in the world. The potential payoff is huge.

Objectives

Compliance and enforcement programs have multiple objectives – conditions they seek to improve and harmful incidents they seek to reduce.

The ultimate objective of virtually all environmental compliance and enforcement efforts, I would argue, is to improve human, animal, plant, and ecosystem health and reduce the frequency, severity, and consequence of environmental stressors such as chemical or oil spills and permit exceedances. Compliance is another ultimate objective, in its own right, because we value lawfulness as a society.[2]

In addition, compliance with environmental laws, regulations, and individual and general permit conditions is an intermediate objective that contributes to better environmental conditions and fewer, less consequential incidents. Other intermediate objectives are awareness of regulatory requirements, understanding of those requirements, motivation to comply, and motivation to adopt "beyond compliance" practices such as pollution prevention that are not required but lead to better outcomes.[3]

Environmental outcomes and lawfulness are not the only objectives of interest, though. Higher public-value return on government investment—more bang for the taxpayer's buck—is another important objective. Whether programs are funded with taxes or fees, we want government to push continually to find ways to achieve more with less in areas where government tries to have an impact.

Fairness, and perceptions of fairness, are also important objectives of compliance and enforcement programs. [4] Compliant parties are placed in a competitive disadvantage if non-

[2] Tom Tyler, 'Procedural Fairness and Compliance with the Law' (1997) v.133 2/2 Swiss Journal of Economics and Statistics 219.
[3] Michael Wogalter, et al, 'Research-Based Guidelines for Warning Design and Evaluation' (2002) 33 Applied Ergonomics 219-230 <http://www.who.int/fctc/guidelines/ArtElevenWogalterNine.pdf> accessed 29 December 2014. See also Shelley Metzenbaum, 'Compliance and Deterrence Research Project: Measuring Compliance Assistance Outcomes' State of Science White Paper (EPA 2007).
[4] Tyler (n 4).

compliant companies and facilities are able to avoid compliance costs others bear. Unfairness can become a serious problem if not monitored and addressed.

In addition, the quality of government interaction with people, including but not limited to owners and employees of regulated entities, is important. If people think a government official is rude or incompetent, it hurts government's ability to get needed changes to law and the funding needed for effective programs. Government employees must be courteous and helpful, as appropriate to the situation. Because compliance and enforcement officials must, on occasion, interact with regulated parties who aggressively resist compliance and enforcement orders, compliance and enforcement officials would be well served to think carefully about and then specify interaction objectives for different compliance situations, including contested enforcement cases. Interaction objectives to consider include: awareness of compliance obligations by all regulated parties; full understanding of these obligations; and positive and productive interactions between government officials and those who contact them seeking assistance and clarifying information.

Finally, I would suggest that accountability to the public is an important outcome objective. Accountability is a popular term with multiple definitions so let me suggest one. Accountability means that the public understands what government organizations are trying to accomplish, why, how, progress on objectives, adjustments made to announced strategies and tactics, and why those adjustments were made. It also means that the public can constructively provide feedback and engage on all of the above. This approach to accountability (informed in part by the public accountability and transparency lessons of EPA's Charles River experience described below[5]) was adopted by the Obama Administration in its performance measurement and management efforts, explained and reflected on the website Performance.gov, and subsequently incorporated into U.S. law.[6]

You may agree with my assumptions about the objectives of environmental compliance and enforcement programs or you may not. If you do not agree, I urge you to develop your own list of compliance and enforcement objectives, discuss them, and then figure out how to measure progress and find ways to improve. The alternative is taking action without being clear about what you are trying to accomplish and hoping, but not confirming, that those actions have the intended effect.

Case 1: Cleaning Up the Charles River

---

[5] Charles River Conservancy, 'A Swimmable Charles? Water Quality and Public Access with Examples from Swiss Urban Rivers' (Charles River Conservancy, 2013) <http://www.thecharles.org/media/uploads/2013/04/Charles03_web.pdf> accessed 29 December 2014; Shelley Metzenbaum, 'Performance Accountability: The Five Building Blocks and Six Essential Practices' (IBM Center for the Business of Government 2006); and Harvard Kennedy School Executive Session, 'Get Results Through Performance Management: Memorandum to New Government Executives' (2001).

[6] United States Government, <www.performance.gov> accessed 29 December 2014 and United States Government, 'The Government Performance and Results Modernization Act' (2010) <http://www.gpo.gov/fdsys/pkg/BILLS-111hr2142enr/pdf/BILLS-111hr2142enr.pdf> accessed 29 December 2014.

Let me begin discussion of compliance and enforcement measurement with the story of the ongoing clean-up of the lower Charles River, which runs between Cambridge and Boston in eastern Massachusetts. In 1995 on the eve of the world-famous Head of the Charles regatta, EPA New England's Regional Administrator John DeVillars publicly announced a goal: the lower Charles River would be swimmable in ten years. At the time, the river was so badly polluted that boaters who fell in were advised to get a tetanus shot.[7]

DeVillars put EPA staff attorney Ken Moraff in charge of this goal. Moraff recruited two engineering colleagues to help him, and working part-time the team quickly realized that EPA would need better data to figure out how to meet this ambitious target. Fortunately, the non-profit Charles River Watershed Association (CRWA) had recently begun a volunteer monitoring effort. CRWA had trained volunteers to collect fecal coliform and dissolved oxygen samples, and they collected them every month at 37 points along the 80-mile stretch of the river. CRWA posted these data almost immediately in an online spreadsheet.[8] Moraff recognized that these data were more "actionable" than anything else EPA or the state had. Although volunteer-gathered, they were valid enough to trigger focused follow-up questions.

Each month, the EPA team analyzed the CRWA data to look for anomalies and patterns. When found, they tried to understand the likely cause. In one case, when EPA saw a downstream reading worse than an upstream one that could not be explained by a permitted discharger between the two monitoring points, EPA called the Boston Water and Sewer Commission (BWSC) to see if it could explain the anomaly. BWSC could not, so the two teams joined together to "walk the pipes." Doing so, they discovered an illicit hookup to the storm drain sending untreated waste that should have been directed to the sanitary sewers for treatment flowing instead through the storm sewers and directly into the Charles.

The EPA team wondered if more illicit hook-ups existed. Brainstorming, someone suggested lifting manhole-covers over storm sewers on dry days to look for running water. EPA asked Massachusetts cities and towns to do this and to eliminate illicit hook-ups they found.[9] This worked remarkably well. Within five years, the Charles was safe for boating 92% of the time (up from 39%) and safe for swimming 59% of the time (up from 19%).[10]

EPA did not meet its ten-year target. What is interesting is that this has never become a political issue. Perhaps that is because EPA shares its progress, problems, and strategies with the

---

[7] Shelley Metzenbaum, 'Measurement that Matters: Cleaning Up the Charles River' in D Kettl, ed., Environmental Governance: A Report on the Next Generation of Environmental Policy (Brookings 2002) 58.

[8] Charles River Watershed Association <http://www.crwa.org/field-science/monthly-monitoring/water-quality-data> accessed 29 December 2014.

[9] Most municipalities did this cooperatively, but EPA took legal steps to compel local government action where needed. See, for example, EPA, 'Good News for Charles River Rowers: the River is Getting Cleaner' (EPA, 1998) <http://yosemite.epa.gov/opa/admpress.nsf/6427a6b7538955c585257359003f0230/a2064493471efa37852574d300024786!OpenDocument> accessed 29 December 2014.

[10] Charles River Conservancy (n 5).

public every year, usually around Earth Day. One sign of the political viability and sustainability of this transparent goal-setting, data-rich approach is that multiple regional administrators working for several U.S. Presidents from different political parties have continued the goal of a swimmable Charles. In September 2014, nineteen years after the goal was first set, EPA released its annual grade for the river, an A-. The river was safe for boating 96% of the time, and swimmable 70% of the time, the best result to date.[11]

It is hardly intuitive to think of water quality data as an effective way to measure compliance and enforcement activities, but the Charles River example shows just how powerful it can be – stimulating innovative thinking and revealing ways to improve. Beyond that, the case shows that useful measurement is not always costly to gather and can come from multiple sources, not just from government.

Case 2: Reducing Wetlands Loss in Massachusetts

In the late 1980s, I ran the capital budget for the Commonwealth of Massachusetts. My deputy, Beth Storey, argued that we should approve the state environmental agency's capital funding request to map the wetlands of Massachusetts, using capital (as opposed to operating) dollars to pay both for high-resolution flyover photographs of the state and for experts to review those photographs closely to look for wetlands markers. I was dubious, I must admit. Was this a legitimate capital project that would produce long-term returns? Still, I greatly respected my deputy's intelligence and expertise in this area so I agreed -- and I'm glad I did. This capital investment created a baseline for measuring wetlands loss, identifying violations, and winning environmental enforcement cases, including some requiring wetland restoration. This investment in measurement has delivered enormous returns and will continue to deliver more for years to come.

Throughout the 1990s and into the early millennial years, the Massachusetts Department of Environmental Protection (DEP) gathered and analyzed aerial photographs to map over 70% of the state's wetlands.[12] Massachusetts has since been able to update these photographs, thanks to cost-reductions in high-resolution digital photography and greater opportunities for camera-carrying flyovers.

As technology evolved, state officials realized they could digitize the original photographs and program computers to detect significant, measurable differences in photos from different time periods.[13] The images in Figure 1 below show how readily apparent illegal

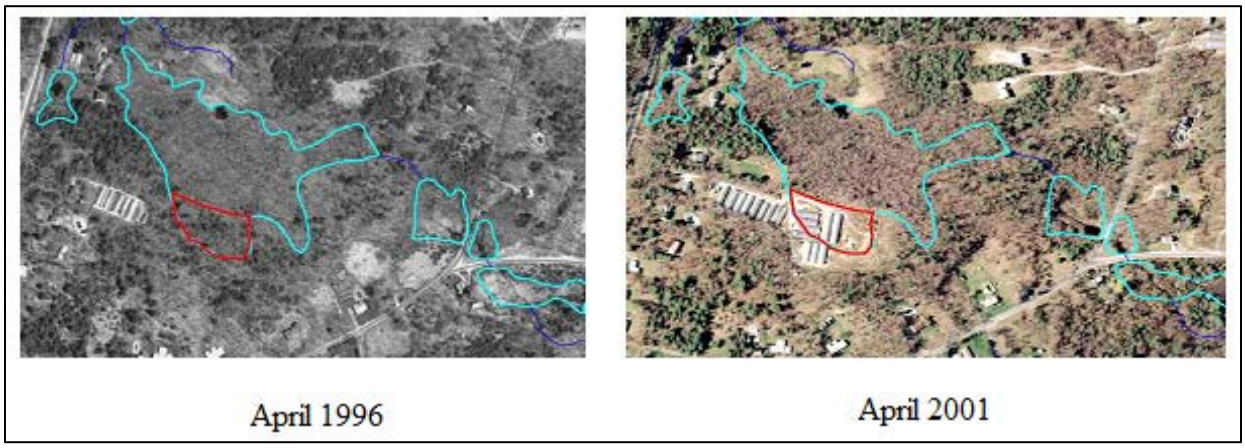
[11] EPA, 'Clean Charles River Initiative' (EPA) <http://www.epa.gov/region1/charles/reportcards.html> accessed 29 December 2014; and EPA, 'Charles River 2013 Report Card Reflects Cleanest Water in Decades' (EPA, 2014) <http://yosemite.epa.gov/opa/admpress.nsf/6d651d23f5a91b768525735900400c28/473e88ce3d5353a485257d48005a4ea0!OpenDocument> accessed 29 December 2014.

[12] Lealdon Langley, 'Wetlands Enforcement Using Aerial Surveillance' (Spring 2007) *ECOStates* 15. See also 'MassGIS Data – MassDEP Wetlands Change' (2013) <http://www.mass.gov/anf/research-and-tech/it-serv-and-support/application-serv/office-of-geographic-information-massgis/datalayers/wetchange.html> accessed 29 December 2014.

[13] Cynthia Giles, 'Next Generation Compliance' (2013) *The Environmental Forum* 22.

construction in wetlands is in aerial photographs. The turquoise lines indicate the originally mapped wetlands, while the red lines show the location of illegal activity.[14]

Figure 1:



Comparing more recent photographs with the baseline photographs, the state identified 850 acres of wetland loss spread over 3,248 different sites in a ten-year period. In 2004, then-DEP Commissioner Robert Golledge reported that DEP had successfully concluded ten enforcement cases, ordered the restoration of 25 acres of illegally filled wetlands, assessed some $640,000 in penalties, and referred two violators to the Attorney General for civil prosecution.[15]

Maps of Massachusetts' wetlands also allowed regulators to allocate staff resources more wisely. Cynthia Giles, who headed the Massachusetts wetlands programs at the time and pioneered use of comparative maps for enforcement cases, realized that while 80 percent of wetland loss occurred on unpermitted lands, 90 percent of her staff's time was spent reviewing permit applications and conducting inspections—clearly a misallocation of scarce resources.[16] Further, she and her staff realized that their inspections were primarily road-based, entirely missing the large volume of off-street violations the photographs revealed. Today, Massachusetts makes highly detailed wetland maps available online for citizens and communities to track the health of their wetlands.[17]

A critical lesson from Massachusetts's experience mapping wetlands is that investment in measuring outcomes can have a huge compliance and enforcement payoff, although one that may require some patience, persistence, and innovation to realize. Not only was DEP's understanding of the contours of its challenge enhanced by geo-spatial measurement, it was also better able to identify non-compliance problems, win enforcement cases, and maximize return on government spending.

---

[14] Cynthia Giles, 'Protecting Massachusetts Wetlands Through Innovative Technology' presentation to Governing/Environmental Compliance Consortium Forum on Managing Environmental Information (2004) Slide 22.

[15] Robert Golledge, 'Massachusetts Makes Enforcement Strides Despite Funding Woes' (Fall 2004) *ECOStates* 4-5.

[16] Giles (n 14) 19-20.

[17] "MassGIS Data – MassDEP Wetlands Change" (n 13).

Case 3: Counting the Frequency and Noting the Nature of Non-Compliance

In 1997, Massachusetts launched the Environmental Results Program (ERP) to promote environmental compliance in business sectors with too many small companies to be adequately inspected but with significant enough environmental risk to warrant attention. The ERP program pioneered a new approach to recording and analyzing non-compliance findings to assess program impact and find ways to improve in three sectors that use small volumes of hazardous chemicals that need proper handling and disposal – dry cleaners, commercial printers, and photo processors.

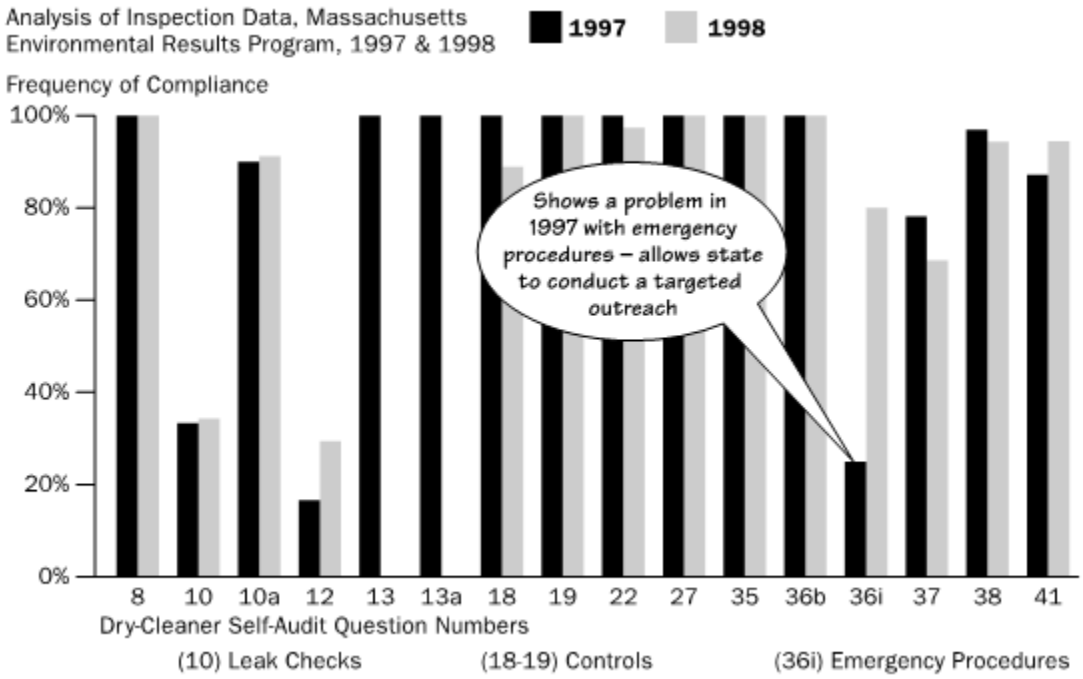
To initiate the ERP program, the state first worked to identify every facility in each selected sector. It then randomly sampled those facilities to create a representative group and inspected them to create a baseline of compliance findings to compare to future results and to detect common types of non-compliance problems for targeted compliance assistance outreach.

Simultaneously, working with trade associations, the state created a plain language, user-friendly workbook for each sector explaining compliance obligations and recommending beyond-compliance practices, such as pollution prevention. Each workbook includes a checklist of required and recommended practices, which the Massachusetts Department of Environmental Protection (DEP) calls Environmental Business Practices Indicators. Every company in the selected sector is required to complete the checklist, certify its accuracy, and submit the information to Massachusetts DEP. In addition, regulated entities that find they are non-compliant with a requirement must submit and execute a corrective action plan. Post-certification, Massachusetts randomly samples and inspects a representative group of facilities and compares inspection data to baseline and prior-year findings to determine how well the sector is performing and assess the accuracy of the self-reporting forms.[18]

Figure 2 presents the incidence of dry cleaner non-compliance with respect to specific performance standards for 1997, just before adoption of ERP, compared to 1998, a year after companies filled out their first self-certification forms. The figure makes clear that compliance with leak check and emergency procedure requirements was a problem prior to program implementation and that emergency procedures practices improved after implementation. A similar chart for 2000 showed compliance levels for leak checks and perchloroethylene record-keeping, low in 1997, had improved to over 60 percent.[19]

[18] National Center for Environmental Innovation, 'ERP States Produce Results' (December 2007) <http://www.epa.gov/erp/files/2007reportfull.pdf> accessed 29 December 2014.
[19] Shelley Metzenbaum, 'More Nutritious Beans' (2003) *The Environmental Forum* 28.

Figure 2:



Source: Massachusetts Department of Environmental Protection[20]

The strength of this chart is that it spotlights problems and progress with regard to specific regulatory requirements. Traditional environmental compliance and enforcement programs can similarly code their inspection findings to gain valuable insights about the nature of non-compliance in facilities. Doing so requires agencies to standardize what they inspect and centrally record their inspection findings, something that unfortunately is too often not done.

The New Hampshire Department of Environmental Services (DES) decided to test this approach in a traditional program, for facilities regulated under the Resource Conservation and Recovery Act (RCRA). It asked its top inspectors to identify ten priority practices for RCRA-regulated parties: seven compliance obligations and three "beyond compliance" practices. It then trained summer interns to assess business performance in these ten areas at a random sample of facilities. Somewhat to the surprise of the state, this project did more than just help it estimate a compliance rate, better understand the nature of non-compliance, and identify specific non-compliance problems needing enforcement follow-up by state inspectors. It also greatly increased awareness within the regulated community of compliance obligations and beyond-compliance opportunities because the interns reached so many more facilities than the small number of state inspectors had previously been able to and did so in an initially collegial rather than adversarial manner. In addition, following the initial intern inspections, a few trade associations and suppliers strengthened the compliance assistance they offered their clients.[21]

[20] Metzenbaum (n 19) 18.

[21] New Hampshire Department of Environmental Services, 'Measuring the Effectiveness of Partial RCRA Inspection, Risk-based Facility Targeting and Compliance Activities' at U.S. Environmental Protection

Colorado similarly found coding its RCRA inspection findings useful, enabling it to detect previously unknown patterns of problems and get fast feedback on the effectiveness of targeted compliance assistance efforts.[22]

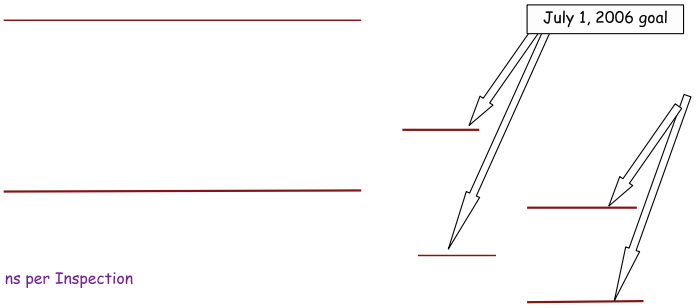
Building on New Hampshire's and Colorado's RCRA measurement efforts, the San Diego County Department of Environmental Health tested a new way to gauge the impact of compliance assurance work with its biotechnology and research and development firms. It measured the average number of violations and the average number of "top 10" violations per inspection, as well as the total number of violations and top 10 violations. As Figure 3 shows, tracking this information enabled the county to assess the overall effectiveness of its compliance assistance efforts and quickly detect when a favorable trend line started to turn in a less favorable direction.[23]

Agency Office of Enforcement and Compliance Assurance Abstracts and Results for State and Tribal Assistance Grants (STAG) <http://cfpub.epa.gov/compliance/state/grants/stag/abre.cfm?id=41802&templatePage=2> accessed 29 December 2014. This information also comes from discussions at meeting convened by the Environmental Compliance Consortium.

[22] Tiffin Shewmake, 'Using Compliance Rates to Manage' (Fall 2004) *ECOStates*.

[23] Shelley Metzenbaum, 'Can Environmental Agencies Make Better use of Accident, Incident, and Inspection Information?' (Spring 2007) *ECOStates* 44.

Figure 3:



Why, What and How to Measure

The preceding three cases demonstrate ways to measure effectively, sources of useful measurement, and a variety of different measurement users and uses. The cases demonstrate that useful measurement can be applied to a whole program or to a specific intervention, and can be produced by government, volunteers, and even regulated parties. Beyond that, the cases illustrate the value and feasibility of enlisting indicators as varied as the quality of ambient conditions, before-and-after maps, and non-compliance findings reported by facilities, inspectors, and trained assessors.

Why Measure. Measurement serves many purposes. It can illuminate what works and likely paths for improvement. It can motivate regulators and regulated parties, communicate priorities, and enlist expertise and assistance. Measurement also reinforces the message that a previously announced objective remains a priority, especially when leaders (whether the head of a government organization or a local team leader) hold data-rich discussions to explore what the data say.

Measurement helps determine if government action had the intended impact. It can sometimes make a compelling case for an action's or program's impact on its own, as with the monthly Charles River water quality and annual ERP compliance data. Often, however, measurement needs to be paired with evaluation methods to distinguish the impact of agency action from that of other variables possibly influencing the same outcomes.

Regression analyses and measured trials are two such evaluation methods. In recent years, EPA has supported regression analyses to understand factors affecting compliance assurance and ways to measure the impact of the full spectrum of compliance assurance activities.[24] It has also supported randomized trials comparing a treatment group to a control group.[25] Noteworthy progress has also been made in the United States, the United Kingdom, and other countries integrating small, iterative randomized control trials into regulatory agency operations to improve compliance.[26]

---

[24] Jay Shimshack, 'Monitoring, Enforcement, & Environmental Compliance: Understanding Specific & General Deterrence' (EPA State-of-Science White Paper 2007); Wayne Gray and Jay Shimshack, 'The Effectiveness of Environmental Monitoring and Enforcement: A review of the Empirical Evidence' (2011) *Review of Environmental Economics and Policy* 1.

[25] Dyke-Redmond T, C Leggett, and M Crow, 'Measuring the Effects of EPA Compliance Assistance in the Auto Body Sector: A Statistically Valid Pilot Project' (EPA, November 2012) <http://www.epa.gov/evaluate/pdf/enforcement/measuring-compliance-autobody-sector.pdf> accessed 29 December 2014.

[26] Re: United Kingdom, see Michael Hallsworth et al, 'The Behavioralist As Tax Collector: Using Natural Experiments to Enhance Tax Compliance' National Bureau of Economic Research Working Paper No. 20007 (2014); Richard Thaler, 'Watching Behavior Before Writing the Rules' New York Times (July 7, 2012) <http://www.nytimes.com/2012/07/08/business/behavioral-science-can-help-guide-policy-economic-view.html?pagewanted=all&_r=1&> accessed 29 December 2014;
Tamsin Rutter, 'What Next for the Nudge Unit?' (June 1, 2014) *The Guardian* <http://www.theguardian.com/public-leaders-network/2014/jun/01/nudge-unit-behavioural-insights-team-conference> accessed 29 December 2014. Re: United States, the U.S. Internal Revenue Service and the U.S. Department of Labor have both made significant progress

Often under-appreciated, measurement and evaluation should answer more than the question, "What works?" Arguably even more important, they can and should answer the question, "What works better?" [27] The Charles River and Environmental Results Program examples demonstrate the power of finding ways to improve. "Positive outliers," regulatory units, regulated parties, or locations trending better or doing better on an absolute basis than peers should trigger questions such as: are there practices that explain better performance worth trying to replicate in other situations? If so, should they be promoted for broader adoption? Measurement should also be used to look for negative outliers and patterns of problems pointing to issues needing attention, their relative import, their key characteristics, and their root causes. This kind of analysis is essential for designing and implementing more effective, efficient prevention and response strategies.

Measurement can also motivate. Lab and field research finds that people like to do well, and well-designed measurement systems can motivate them to do better. In addition, measurement, especially when paired with a compelling goal, communicates and energizes people inside a goal-setting organization.[28] It also enlists expertise and assistance from beyond an organization's boundaries, as was evident with the Charles River goal and New Hampshire RCRA effort. Finally, measuring and publicly reporting progress on goals strengthens accountability to the public.

What and How to Measure. Many environmental agencies try to measure their compliance and enforcement work by counting activities that are easy to measure, such as the number of inspections completed, enforcement actions initiated, and penalties levied. So, too, does the media.[29] Counting these activities provides helpful information for running programs, budgeting for those activities, and understanding the percentage of initiated enforcement actions leading to enforcement initiation and to penalties. Using those counts as performance indicators creates serious problems, however. For example, local offices may be tempted to initiate enforcement actions for trivial violations, fearing they will be chastised if their enforcement numbers are too low.

EPA and several states have tried a variety of ways to make tallies of enforcement activity more valid as outcome indicators, including measuring penalties levied and pounds of

---

integrating evaluation into program management. The IRS has adopted "test and learn approaches" in areas as diverse as taxpayer authentication, small business underreporting, and reducing improper payments for the Earned Income Tax credit, from correspondence between author and Dean Silverman, Senior Advisor and Director of the Office of Compliance Analytics (December 2014). The Occupational Safety and Health Administration, U.S. Department of Labor, has created a website to share its evaluation work -United States Department of Labor, 'Clearinghouse for Labor Evaluation and Research' (Department of Labor) <http://clear.dol.gov/synthesis-report/evidence-effects-osha-activities> accessed 29 December 2014.

[27] Gawande A, *Better: A Surgeon's Note on Performance* (Picador 2007); Anthony Bryk, *Learning to Improve: How America's Schools Can Get Better at Getting Better* (Harvard Education Press: forthcoming).

[28] For a review of the relevant literature, see Metzenbaum, 'Performance Accountability' (n 5).

[29] See, for example, Emily Yehle, 'Slide in EPA Clean Water Criminal Enforcement Continues Under Obama' (2010) Greenwire <http://www.eenews.net/stories/1059941312> accessed 29 December 2014.

pollution reduced as a result of enforcement cases. Calculating pollution reduced can have value if it encourages agency staff to think more about the relative environmental import of the cases they pursue *before* initiating the case, and measuring penalties levied can provide useful insights about the seriousness of the cases being pursued, especially if penalties actually collected are also measured and compared to those levied. Neither, however, works well as a performance indicator. For starters, major enforcement cases often take more than a year to develop and resolve, so a year with high levels of pollution reduced and fines levied can easily be followed by year with low levels, implying worse performance that is not, in fact, worse. Beyond that, counting pollution reduced through enforcement would double count pollution reduced through permits if that too is being measured. Finally, counting penalties levied without also counting penalties collected does not provide an accurate picture of enforcement fines. In short, dressing up the number of enforcement actions with penalty dollars and pollution reduction estimates does not accurately reflect compliance and enforcement performance.

Also, as compliance rises, the number and gravity of enforcement should fall. True success therefore would appear as declining enforcement cases and penalties. Yet downward trends in enforcement activity might worry observers without the context of outcome trends, such as the percent of inspected facilities with high priority violations, and other information, such as compliance monitoring levels, as EPA shows on its Environmental Compliance History Online (ECHO) dashboards.[30]

Outcomes information is needed to put activity trends in perspective to avoid the risk of managing activities that are easy to measure rather than managing the outcomes and intermediate outcomes government wants to improve. Activity data needs to be paired with outcomes information such as compliance levels, non-compliance patterns, harmful environmental incidents, discharge trends, air quality, or animal or human health. Many have long considered outcome measurement too hard or costly. As the examples in this chapter suggest, this need not be the case. This is especially true in the information age when the costs of data collection and analysis are so low compared to the past and when data-rich presentations and discussions are everywhere. Interesting outcome indicators tried in recent years include mercury levels in fish tissue, real-time transmitted signals of where asthma inhalers are being used, discharge trends, and remote monitoring of air and water releases crossing geographic boundaries. Interesting intermediate outcome indicators include awareness of regulatory obligations and perceptions of compliance monitoring likelihood.[31]

Many compliance and enforcement lawyers use data to look for the most egregious violators and precedent-setting cases, but do not follow up after the case to assess whether the precedent-setting case had the expected effect. Was, for example, a highly publicized

[30] See EPA, 'Analyze Trends: State Air Dashboard,' (EPA Enforcement and Compliance History Online) <https://echo.epa.gov/trends/comparative-maps-dashboards/state-air-dashboard?redirect=page> accessed 29 December 2014.

[31] Les Carlough, 'General Deterrence of Environmental Violation: A Peek into the Mind of the Regulated Public' (Oregon Department of Environmental Quality, 2010) <http://www.deq.state.or.us/programs/enforcement/DeterrenceReport.pdf> accessed 29 December 2014; Wogalter (n 3).

enforcement action for failure to get a required permit followed by a sudden rise in permit applications in that program? Looking at permit application trends before and after the case can provide a sense of whether the deterrence message was successfully sent. Gathering data about compliance characteristics of the entire caseload, as the ERP program did, can put a precedent-setting case in context and suggest ways to influence other non-compliant parties.

Failure to measure progress on the outcome objectives discussed earlier, both ultimate and intermediate, results in the possibility that regulators and regulated parties will spend time and money doing activities that do not improve outcomes or do so sub-optimally, as the wetlands example illustrated. Related to that, the wetlands example illustrated how reviewing costs relative to outcomes helps allocate scarce resources to higher-yield activities.

It can also be useful to measure process efficiency and adequacy attributes of compliance and enforcement activities. Useful process (and intermediate outcome) measures include risks, timeliness, timing, coverage, and completeness. Measuring risks (such as high staff turnover rates) and unwanted side effects (such as complaints about overly aggressive inspectors) provides warning signs to government agencies and regulated parties, allowing them to prevent problems before they become serious.

One particularly important process measure is the compliance monitoring or coverage rate when inspection or other compliance monitoring findings are used as the outcome indicator. When fewer than 100 percent of facilities in a program or sector are inspected as is often the case, it is important to note the reason for monitoring (e.g., inspections conducted to follow up on complaint or suspected problem, random sampling, follow-up on past problem) as well as the type of compliance monitoring action (e.g., government-conducted inspection, self-reporting with or without validation, remote sensing) along with the compliance monitoring rate. This allows more accurate interpretation of the data.

Noting the time and location of incidents and inspections can inform inspection scheduling and improve results. That is what happened when a local office of the Coast Guard realized that most of its oil spills occurred at night while all of its inspections took place during the day. Rescheduling inspections to align with spill hours reduced the number of oil spills.[32] Tracking the percentage of mandated corrective actions completed, penalties collected relative to penalties levied, and penalties paid on time promotes compliance and advances fairness. Tracking time needed for internal clearances reveals bottlenecks, which can reduce processing time, frustration, and costs.

Measurement Challenges

The value of effective measurement and management is so large that it is a baffling that it is so rare (and where it exists, so rarely used), especially in the information age. Several possible reasons may explain the paucity and inadequacy of compliance and enforcement measurement.

One is frustration with past measurement efforts. Poorly designed measurement systems add to people's workload without giving them offsetting returns. One feature of a well-designed

[32] Metzenbaum, 'Compliance and Deterrence Research Project' (n 3) 22.

measurement system is not just the choice of data to collect and the means for collection, but also adopting the discipline of returning data to data suppliers with value added through analysis so that the data suppliers learn from and otherwise benefit from the data they supply. This motivates data suppliers to pay more attention to the quality of the data they submit.

Measurement can also be scary if people fear they will be punished for not meeting or making adequate progress toward an announced goal. Some argue that measurement should be used punitively to cut funding for ineffective programs[33] or to compare, name, and shame the worst performers. This can be problematic, encouraging measurement manipulation. It can also discourage government officials from setting the kind of hard-to-achieve outcomes-focused stretch targets that lab-based and field evidence find focus and energize staff and encourage innovation, ultimately leading to greater impact per dollar spent. The Charles River example shows how stretch targets need not be scary nor missed targets embarrassing if goals, progress, problems, and strategies are clearly communicated and the community engaged.

Certainly, the intergovernmental nature of the environmental regulatory system complicates measurement. Common measures that facilitate comparison to find positive outliers can be achieved by a top-down mandate from the federal government or by agreement across EPA, 50 states, and local jurisdictions with environmental regulatory responsibilities. Either option, unfortunately, can be politically difficult to achieve. Some resist collecting the data that will make it easier to compare, fearing comparison will make them look environmentally weak; others fear comparisons will make their environmental protection look too strong and hurt their ability to attract and retain business. Others, often those who already invested in sophisticated measurement systems, resist because if they think national data standards have been reduced to the lowest common dominator and they will need to invest in new systems and use different metrics after having just made major changes.

Another possible problem is fear of data or the lack of analytic skills. Some environmental compliance and enforcement lawyers have strong quantitative backgrounds, but many do not. Also, I suspect few law schools offer courses on how to analyze data or run measured trials to find improvement opportunities, let alone require courses teaching their students these skills.

The traditional organizational structure of most environmental agencies poses another problem. Few environmental agencies have offices responsible for data analysis, for running data-rich reviews of progress, or for sharing data analyses with peers to learn.

These impediments to better measurement, analysis, and data-rich decision-making are unfortunate because reams of relevant data (e.g., inspection findings, self-reported findings, discharge trends, environmental incidents, and root cause analyses) go under-examined. Information is collected, but seldom wisely organized, analyzed, and distributed.

New technologies, new attitudes towards big data, and more strategic thinking afford unprecedented opportunities to make data more useful. The question is: who will push to make

[33] Virginia Thomas, 'Restoring, Results, and Accountability' in D Forsyth ed., *Quicker, Better, Cheaper?* (Rockefeller Institute Press 2001) 11.

that happen? Cynthia Giles, the program director who started using digitized maps for wetlands enforcement in Massachusetts and Assistant Administrator of the EPA Office of Enforcement and Compliance Assurance in the Obama Administration as this chapter is being written, is one such leader. She has called increasing use of strategic, transparent measurement and monitoring through law, regulation, enforcement agreements, and program design.[34] Let us hope she succeeds and that other leaders add their voices and action to hers.

Conclusion

Effective measurement is essential to good government and need not be complicated, burdensome, or scary. It is tempting to think of measurement as extra work, sometimes distracting, or work that is nice but not necessary. That is not the case. Measurement need not be difficult, but it does require clear thinking about objectives and how to assess progress on those objectives, as well as a commitment to start to measure and continuously use data to find ways to improve. Without a strong system of outcomes-focused measurement, regulators run a high risk of being wasteful and ineffective.

This chapter has offered examples and reasons for integrating measurement wholly into environmental compliance and enforcement decisions. When done well, it can be transformative, allowing regulatory agencies to achieve better environmental outcomes, higher compliance rates, improved fairness, a better understanding, and higher return on investment. I hope the reader has been convinced to play a part building the measurement capacity of regulators and regulated parties to achieve better environmental and compliance results.

[34] Cynthia Giles, 'Annotations and References to "Next Generation Compliance' published in *The Environmental Forum*' (Environmental Law Institute 2013) <http://www2.epa.gov/sites/production/files/2014-10/documents/giles-next-gen-article-forum-eli-sept-oct-2013-appendix.pdf>.

Works Cited

Bryk, A, *Learning to Improve: How America's Schools Can Get Better at Getting Better* (Harvard Education Press: forthcoming)

Carlough L, 'General Deterrence of Environmental Violation: A Peek into the Mind of the Regulated Public' (Oregon Department of Environmental Quality, 2010) <http://www.deq.state.or.us/programs/enforcement/DeterrenceReport.pdf> accessed 29 December 2014

Charles River Conservancy, 'A Swimmable Charles? Water Quality and Public Access with Examples from Swiss Urban Rivers' (Charles River Conservancy, 2013) <http://www.thecharles.org/media/uploads/2013/04/Charles03_web.pdf> accessed 29 December 2014

Charles River Watershed Association <http://www.crwa.org/field-science/monthly-monitoring/water-quality-data> accessed 29 December 2014

Correspondence between author and Dean Silverman, Senior Advisor and Director of the Office of Compliance Analytics (December 2014)

Dyke-Redmond T, C Leggett, and M Crow, 'Measuring the Effects of EPA Compliance Assistance in the Auto Body Sector: A Statistically Valid Pilot Project' (EPA, November 2012) <http://www.epa.gov/evaluate/pdf/enforcement/measuring-compliance-autobody-sector.pdf> accessed 29 December 2014

EPA, 'Charles River 2013 Report Card Reflects Cleanest Water in Decades' (EPA, 2014) <http://yosemite.epa.gov/opa/admpress.nsf/6d651d23f5a91b768525735900400c28/473e88ce3d5353a485257d48005a4ea0!OpenDocument> accessed 29 December 2014

—— 'Analyze Trends: State Air Dashboard,' (EPA Enforcement and Compliance History Online) <https://echo.epa.gov/trends/comparative-maps-dashboards/state-air-dashboard?redirect=page> accessed 29 December 2014

—— 'Clean Charles River Initiative' (EPA) <http://www.epa.gov/region1/charles/reportcards.html> accessed 29 December 2014

—— 'Good News for Charles River Rowers: the River is Getting Cleaner' (EPA, 1998) <http://yosemite.epa.gov/opa/admpress.nsf/6427a6b7538955c585257359003f0230/a2064493471efa37852574d300024786!OpenDocument> accessed 29 December 2014

Gawande A, *Better: A Surgeon's Note on Performance* (Picador 2007)

Giles C, 'Next Generation Compliance' (2013) *The Environmental Forum* 22

—— 'Annotations and References to "Next Generation Compliance' published in *The Environmental Forum*' (Environmental Law Institute 2013) <http://www2.epa.gov/sites/production/files/2014-10/documents/giles-next-gen-article-forum-eli-sept-oct-2013-appendix.pdf>

—— 'Protecting Massachusetts Wetlands Through Innovative Technology' presentation to Governing/Environmental Compliance Consortium Forum on Managing Environmental Information (2004) Slide 22

Golledge Jr. R, 'Massachusetts Makes Enforcement Strides Despite Funding Woes' (Fall 2004) *ECOStates* 4-5

Gray W and J Shimshack, 'The Effectiveness of Environmental Monitoring and Enforcement: A review of the Empirical Evidence' (2011) *Review of Environmental Economics and Policy* 1

Hallsworth M, et al, 'The Behavioralist As Tax Collector: Using Natural Experiments to Enhance Tax Compliance' National Bureau of Economic Research Working Paper No. 20007 (2014)

Harvard Kennedy School Executive Session, 'Get Results Through Performance Management: Memorandum to New Government Executives' (2001)

Langley L, 'Wetlands Enforcement Using Aerial Surveillance' (Spring 2007) *ECOStates* 15
'MassGIS Data – MassDEP Wetlands Change' (2013) <http://www.mass.gov/anf/research-and-tech/it-serv-and-support/application-serv/office-of-geographic-information-massgis/datalayers/wetchange.html> accessed 29 December 2014

—— 'Can Environmental Agencies Make Better use of Accident, Incident, and Inspection Information?' (Spring 2007) *ECOStates*

—— 'Compliance and Deterrence Research Project: Measuring Compliance Assistance Outcomes' State of Science White Paper (EPA 2007)

—— 'Performance Accountability: The Five Building Blocks and Six Essential Practices' (IBM Center for the Business of Government 2006)

—— 'More Nutritious Beans' (2003) *The Environmental Forum*

—— 'Measurement that Matters: Cleaning Up the Charles River' in D Kettl, ed., Environmental Governance: A Report on the Next Generation of Environmental Policy (Brookings 2002) 58

National Center for Environmental Innovation, 'ERP States Produce Results' (December 2007) <http://www.epa.gov/erp/files/2007reportfull.pdf> accessed 29 December 2014

New Hampshire Department of Environmental Services, 'Measuring the Effectiveness of Partial RCRA Inspection, Risk-based Facility Targeting and Compliance Activities' at U.S.

Environmental Protection Agency Office of Enforcement and Compliance Assurance Abstracts and Results for State and Tribal Assistance Grants (STAG) <http://cfpub.epa.gov/compliance/state/grants/stag/abre.cfm?id=41802&templatePage=2> accessed 29 December 2014

Rutter T, 'What Next for the Nudge Unit?' (June 1, 2014) *The Guardian* <http://www.theguardian.com/public-leaders-network/2014/jun/01/nudge-unit-behavioural-insights-team-conference> accessed 29 December 2014

Shewmake T, 'Using Compliance Rates to Manage' (Fall 2004) *ECOStates*

Shimshack J, 'Monitoring, Enforcement, & Environmental Compliance: Understanding Specific & General Deterrence' (EPA State-of-Science White Paper 2007)

Thaler R, 'Watching Behavior Before Writing the Rules' New York Times (July 7, 2012) <http://www.nytimes.com/2012/07/08/business/behavioral-science-can-help-guide-policy-economic-view.html?pagewanted=all&_r=1&> accessed 29 December 2014

Thomas V, 'Restoring, Results, and Accountability' in D Forsyth ed., *Quicker, Better, Cheaper?* (Rockefeller Institute Press 2001)

Tyler T, 'Procedural Fairness and Compliance with the Law' (1997) v.133 2/2 Swiss Journal of Economics and Statistics 219

United States Department of Labor, 'Clearinghouse for Labor Evaluation and Research' (Department of Labor) <http://clear.dol.gov/synthesis-report/evidence-effects-osha-activities> accessed 29 December 2014

United States Government, <www.performance.gov> accessed 29 December 2014

United States Government, 'The Government Performance and Results Modernization Act' (2010) <http://www.gpo.gov/fdsys/pkg/BILLS-111hr2142enr/pdf/BILLS-111hr2142enr.pdf> accessed 29 December 2014

Wogalter M, et al, 'Research-Based Guidelines for Warning Design and Evaluation' (2002) 33 Applied Ergonomics 219-230 <http://www.who.int/fctc/guidelines/ArtElevenWogalterNine.pdf> accessed 29 December 2014

Yehle E, 'Slide in EPA Clean Water Criminal Enforcement Continues Under Obama' (2010) Greenwire <http://www.eenews.net/stories/1059941312> accessed 29 December 2014