Measuring Environmental Enforcement Success: The Elusive Search for Objectivity

by Joel A. Mintz

Joel A. Mintz is Professor of Law at Nova Southeastern University’s Shepard Broad Law Center, Member-Scholar at the Center for Progressive Reform, and Elected Member of the American Law Institute.

I. Introduction

How best to measure the effectiveness of environmental enforcement is a long-standing question. Over the 42 years during which the U.S. Environmental Protection Agency (EPA) has engaged in enforcement work, the agency’s system for evaluating its own enforcement efforts (as well as those of the states) has evolved considerably. Beginning with a simple counting of numbers of formal enforcement actions taken, the Agency has added measures such as the volume of pollutants cleaned up as a result of enforcement activities, the beneficial public health impacts of enforcement actions, and the number of years of incarceration of criminal defendants convicted of environmental crimes.

Recently, EPA stirred renewed interest in how to measure enforcement success. Its Draft FY 2014-2018 Strategic Plan suggested a new set of “Next Generation Compliance” metrics. Those “Next Gen” measurements include the number of regulated sources using advanced monitoring technologies to measure their own emissions, the number of enforcement settlements that resulted from or that incorporate advanced monitoring, the number of sectors for which measurable compliance rate strategies have been adopted, and the percentage of facilities that electronically report Clean Water Act (CWA) compliance data to EPA and to authorized states and tribal governments. Although the Agency deleted those measurements from its final FY 2014-2018 Strategic Plan, the questions they raised remain.

In this Article, I will attempt to take a fresh look at the enforcement measurement issue. In doing so, I will make two assumptions. First, I will presume that the public is entitled to know as much as possible about the nature and efficacy of government environmental enforcement programs; and, second, I will assume that government environmental agencies, at all levels, are likely to operate in a climate of resource-scarcity for some time to come. After describing the contentious political context in which the success of environmental enforcement programs has historically been asserted and debated, I will examine the relative merits of various alternative approaches, including the measurement of environmental enforcement outputs, outcomes, and organizational factors. I will then critique the enforcement measurements suggested in EPA’s draft five-year plan. Finally, I will suggest a combination of metrics that is likely to provide a reasonable balance of openness and frugality, in an era of challenge and change for environmental regulation.

II. The Politics of Measuring Enforcement Effectiveness: A Look Back

Measurement of the success of environmental enforcement over given periods of time has long had important political implications. Incumbent elected officials have frequently pointed to various statistical yardsticks to support claims that their administrations—or the administrations run by their political allies—protected the public through vigorous enforcement of environmental requirements. Conversely, political opponents of incumbent administrations have invoked numerical enforcement indicators to bolster a case that their rival ignored environmental concerns by being unduly permissive toward polluting industries. The history of EPA’s enforcement efforts provides several examples of these patterns—as well as one illustration of an administration using enforcement statistics to defend itself against opponents’
claims that it unfairly singled out a particular industry for overly aggressive enforcement.

In the mid-1970s, soon after President Jimmy Carter assumed office, the administration radically changed EPA’s enforcement program to emphasize civil judicial enforcement against violators of environmental laws. In explaining the Agency’s actions, Marvin Durning, EPA Assistant Administrator for Enforcement, cited Agency records reflecting a paucity of civil enforcement actions taken by EPA during the Richard Nixon and Gerald Ford Administrations. During the George W. Bush Administration, in contrast, Grant Nakayama, the head of the Agency’s enforcement office, made use of EPA enforcement statistics to rebut criticism that EPA had been too lax in enforcing federal environmental laws.

Enforcement metrics also played a major role in the efforts of a Democratic-controlled U.S. Congress to investigate—and castigate—the lax enforcement efforts of EPA during the Republican Administrations of Ronald Reagan and George H.W. Bush. Conversely, Democratic allies of the Barack Obama Administration have contrasted the statistical enforcement efforts taken by EPA during the Obama years with those taken under the previous (Republican) administration, to defend against Republican charges that EPA’s Dallas, Texas, regional office had conducted an enforcement vendetta against the oil and gas industry in New Mexico, Oklahoma, and Texas.

Each of these instances of reliance upon enforcement statistics to bolster claims about an administration’s enforcement performance at the EPA was successful, to one extent or another, in political terms. Nonetheless, in each case, the enforcement metrics in question were cited by political actors in the hurry and heat of political controversies. Little or no thought was given as to whether the measurements cited in political battle reflect the most accurate and objective way of assessing the efficacy of an agency’s environmental enforcement performance. Given the sharp, argumentative tone of political discourse, and the hurried nature of political controversies, this situation is, perhaps, understandable.

Nonetheless, a dispassionate analysis of the extent to which particular environmental enforcement measurements are accurate and reliable gauges of governmental enforcement performance remains much needed. As discussed further below, however, such an analysis must take into account both the inherent complexity of the environmental enforcement process and the limited extent to which the environmental impact of particular enforcement cases may be assessed.

### III. Examining the Alternatives

#### A. Outcome Measures

One potential set of measurements of an enforcement program’s success may be termed *outcome measures*. By this phrase, I mean the actual results of governmental enforcement work with respect to public health and the environment. Since at least the 1990s, EPA has made regular use of two such metrics: the number of pounds of pollutants that will no longer be discharged or emitted to the environment as a direct result of enforcement orders or agreements; and the amount of money that would have been spent on health care by people who would have been exposed to pollutants in the absence of such enforcement agreements or orders.

While both of these measurements have value, each has some drawbacks. Although measuring the volume of pollution whose release will be prevented because of enforcement actions provides a rather crude indication of how well the environment has been protected, not all pollutants have the same effect. The release of even a small amount of a highly toxic compound may do more harm to health and the environment than the introduction into air, water, or land of a much larger quantity of a more benign substance. Moreover, this mode of measurement does not take into account the extent to which particular cleanup agreements have deterred other similarly situated firms or municipalities from polluting the environment unlawfully.

 Likewise, although estimations of health care expenditures foregone as a result of compelled environmental cleanups may also shed some light on an environmental enforcement program’s effectiveness, those estimations, too, fail to take account of the extent to which that sort of program is having a deterrent impact on would-be polluters. Moreover, as a practical matter, such measurements seem limited to violations of Clean Air Act (CAA) requirements. At present, they are difficult to apply to hazardous waste cleanups compelled under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) and the Resource Conservation and Recovery Act (RCRA), and are virtually impossible to apply to the abatement of CWA violations.

One evaluation tool that, at first blush, has considerable appeal is measurement of the extent to which environmental enforcement has resulted in environmental improvement. Unfortunately, however, in many cases, this approach is more theoretical than real. Variations in weather conditions, which are entirely unrelated to environmental enforcement, can have enormous impacts on ambient measurements of air and water quality. Thus, year-to-year changes in measurable levels of environmental pollutants may be as much or more a result of meteorological factors as they are a reflection of the success or failure of environmental enforcement efforts. In addition, some

---

4. Id. at 165-66.
5. Id. at 52-61, 68-69, 88-91, 100-02.
environmental pollution may simply be beyond the jurisdictional authority of environmental regulators. Under the CWA, for example, a considerable amount of water pollution, such as rainfall runoff from croplands, is considered to be nonpoint source pollution and subject to only nominal federal and state regulation. To the extent that such pollution has a significant effect on water quality data, that data is obviously a poor gauge of the effectiveness of environmental enforcement programs that, by law, must focus solely on pollutant discharges from point sources.

Another potentially useful way in which enforcement effectiveness may be assayed is to focus on the overall levels of compliance in polluting industries. Unfortunately, however, overall compliance rates in many industries are frequently not measured by government regulators. Doing so is a resource-intensive process that state and federal environmental regulators often lack the funds to undertake. Another drawback is that overall industrial compliance rates do not reveal the relative size and environmental impact of differently sized facilities within a given industry. Thus, for instance, an industry may have a 90% overall compliance rate and still do considerable aggregate harm to health and the environment because the 10% of facilities that are in noncompliance are larger and dirtier than their law-abiding competitors.

Since the deterrence of violations by similarly situated companies and municipalities is an important goal of an enforcement program, it would, of course, be useful to measure the number of regulated companies that have decided to comply with applicable environmental laws out of concern that those firms will be subject to enforcement action. In reality, however, any data that purports to provide such information is not likely to be reliable. For understandable reasons, companies that decide to comply with environmental laws will be extremely reluctant to tell survey researchers that, but for the threat of enforcement action against them, they would have violated the law. Moreover, in some settings, corporate decisionmaking about expenditures is a complex process. Various factors may be taken into account, and a number of individuals may play a role. Even if a corporate representative is prepared to be fully candid with a researcher, he or she may simply not be able to pinpoint the extent to which the likelihood of enforcement played a role in an institutional decision regarding environmental compliance.

B. Output Measures

A second set of possible metrics for evaluating the success of environmental enforcement programs may be referred to as output measures. By this term, I mean measurements of levels of activity of enforcement program personnel, in particular categories, over given periods of time. EPA and the states have made use of these evaluative tools since the 1970s. Among other things, they have kept (and assessed) records of such enforcement-related activities as numbers of inspections and evaluations of pollution sources, the volume of initiated and concluded administrative and civil judicial enforcement cases, compliance with open consent decrees, amounts of civil penalties by violators, numbers of enforcement-compelled Superfund settlements with potentially responsible parties (PRPs), amounts of money recovered in CERCLA cost recovery cases, numbers of filed criminal cases, and years of incarceration of defendants convicted in criminal prosecutions. Although some state environmental officials object to the compiling of these sorts of records as consuming too many of their limited resources, output measures do have genuine value. When accurately kept, they provide at least a rough indication of how much work an enforcement program has done. They also yield insights into the nature of that work.

At the same time, however, output measurements have shortcomings. First, mere counts of orders issued, cases filed, and the like tend to overlook the ways in which environmental enforcement cases differ. In reality, these cases vary immensely in complexity. Enforcement responses to failures to prepare spill prevention control and countermeasure plans or to particular misapplications of pesticide products, for example, require a great deal less time and effort on the part of Agency staff professionals than do multimedia enforcement actions against multiple oil refineries that are owned by the same corporation. It is thus essential—if case counting is to have meaning—that enforcement agencies weigh the enforcement actions they have taken for purposes of evaluation to differentiate among those actions based upon their relative complexity and resource-intensiveness.

Additionally, mere output measurements provide no information as to the timeliness and promptness of enforcement work. One crucial aspect of effective environmental enforcement is early discovery of violations and prompt follow-up action after violations have been identified. Data regarding that component of enforcement work can yield useful measures of an enforcement program’s vigor. Nonetheless, traditional output measurements generally ignore the time it takes to pursue and close out cases.

Beyond this, output measurements shed little or no light on the deterrent impact of environmental enforcement programs. The fact that over a given time period an agency has issued a good many orders, or initiated numerous civil and criminal enforcement actions, does not necessarily indicate that industrial compliance levels have increased or that public health and the environment have been well-safeguarded. Output measures also reveal nothing of the effectiveness of an enforcement agency’s organizational arrangements, which often provide important (albeit indirect) indicators of how successfully the agency is meeting its enforcement responsibilities.

C. Organizational Measures

The final type of environmental enforcement measurements are what I will refer to as organizational measures. By this term, I mean indicators of the extent to which an
A. EPA Proposals

Next Generation (or Next Gen) compliance is the name EPA has given its newest initiative to encourage compliance with environmental standards. This approach was described in a 2013 article by Cynthia Giles, the current Assistant Administrator for Enforcement and Compliance Assurance.10 Ms. Giles’ prescription for improving environmental enforcement has several distinct components. These include: (1) crafting agency rules that are simpler and clearer; (2) allowing third-party validation of compliance in some cases; (3) requiring additional self-certifications of compliance; (4) publishing information regarding company environmental compliance; (5) making wider use of advanced technology for pollution monitoring; and (6) requiring electronic self-reporting.11 Ms. Giles takes the position that “[v]igorous enforcement of the law will always be the backbone of environmental protection.”12 Nonetheless, in her view, “[a]s we continue to learn about ways to strengthen compliance, and take advantage of advances in technology, Next Gen can transform our protection work even in a time of declining budgets.”13

Ms. Giles’ article did not indicate how EPA plans to evaluate the success of the next-generation compliance measures that she favors. However, a table included in the Agency’s final FY 2014-2018 Strategic Plan indicated that “EPA is initiating a discussion with states and the public about ways to incorporate Next Gen approaches into our measures.” It then identified several possible metrics that it described as “a few examples of the types of measures that may be discussed.” These new metrics included quantitative measurements of: (1) enforcement settlements that resulted from or that incorporate advanced monitoring technologies; (2) regulated sources using advanced monitoring to measure their own emissions; (3) facilities electronically reporting CWA national pollutant discharge elimination system (NPDES) data to authorized states and tribes and EPA (stated as a percentage of all regulated facilities); and (4) attempts by the public to use compliance transparency tools.14

B. Evaluating EPA’s Proposals

How useful would these new approaches and measurements be? Although one may well question whether their adoption will truly transform environmental protection as Assistant Administrator Giles suggests, the sorts of techniques advocated in Giles’ piece will, if fully implemented, constitute a clear advance in promoting environmental compliance. Ms. Giles is entirely correct that the technology for monitoring environmental pollution has recently been improved significantly. Requiring its use by pollution sources, and increasing transparency regarding releases of pollution from identified sources (as well as ambient pollution levels), may well lead to environmental improvement. Moreover, the adoption of clear, straightforward environmental regulations will benefit regulated entities, regulatory agencies, and the general public alike.

On the other hand, Ms. Giles’ call for allowing third-party compliance validation “in some cases” appears to raise at least as many questions as it answers. What are the “some cases” in which third-party validation would be deemed appropriate? What third parties would be eligible to do such validation? What credentials would they be required to have? Who (if anyone) would be responsible for checking their credentials? Who would pay validators for their services? If this approach is taken, I hope the Agency will provide clear and satisfying answers to these questions. Even though third-party validation may be workable in theory, EPA and the states will need to take much care to assure that third-party validators will be truly independent, qualified, and objective. The possibility that their findings will be influenced by the pollution sources whose compliance status they are evaluating must be squarely recognized, and eliminated systematically, if this approach is to succeed. Obviously, the likelihood of pro-source bias will increase if third-party violators are paid by—and thus, possibly, beholden to—the very sources whose compliance record they are responsible for evaluating.

The precise ways in which EPA will evaluate the success of any Next Gen compliance measures it adopts are
now unresolved. Nonetheless, the metrics that the Agency included in its Strategic Plan do provide a useful starting point. As EPA now appears to recognize, these measures must supplement, and not supplant, some of the outcome, output, and organizational measures identified above. The elimination of all of those more traditional measures would leave the public, and congressional overseers, without information as to how EPA is responding to reported instances of noncompliance. Instead of promoting transparency, and encouraging vigorous enforcement, the Next Generation enforcement measurements set forth in EPA's Strategic Plan—standing alone—would shroud EPA and state enforcement efforts in a dark cloud of secrecy.

However, assuming that they will be additions to already established enforcement yardsticks, the Next Generation metrics that EPA identified in its recent Strategic Plan seem promising. One hopes that the Agency will one day also use its rulemaking authority to require advanced monitoring and electronic reporting by most, if not all, regulated pollution sources, rather than only compel the use of these technologies through enforcement settlements. In the future, EPA may wish to require facilities to report electronically on their compliance with all federally mandated environmental requirements, rather than limiting electronic self-reporting to the CWA NPDES data mentioned specifically in its draft Strategic Plan. Moreover, EPA would do well to measure the federal government's rates of success in litigated enforcement actions and, as an indicator of the efficiency of EPA's use of its enforcement resources, it may want to consider measuring the levels of pounds of pollutant discharges abated through enforcement activities per dollar spent on enforcement work by the Agency. Those cavils and caveats aside, however, EPA's first brief attempt to suggest Next Generation compliance measures seems a sensible place to start evaluating the successfulness of the Agency's Next Gen initiatives.

**V. Conclusion: Which Measurements Are Best?**

As the foregoing discussion has shown, no possible approach to measuring the effectiveness of government environmental enforcement programs is free of flaws and drawbacks. The various outcome, output, organizational, and next-generation metrics described above each have unique strengths and evident weaknesses, and the likely impacts of certain types of violations (e.g., failures to file monthly reports) may be very different from violations that involve effluent discharges in excess of required limitations. A judicious combination of these sorts of metrics thus seems needed. However, it remains to be determined precisely which enforcement yardsticks, taken together, will best promote an accurate objective assessment of environmental enforcement successfulness.

In an ideal world, EPA and the states might regularly employ all of the environmental enforcement measurements mentioned above, with the modifications and additions I have recommended, to present as full and accurate a picture of the accomplishments, impacts, capabilities, and shortcomings of government enforcement efforts as is possible. However, this is not an ideal world. Keeping accurate records of enforcement outputs, outcomes, organizational metrics, and Next Gen measurements consumes staff time and resources. Since the 1990s, both federal and state environmental agencies have had to face budget cuts, hiring freezes, and (in some cases) employee furloughs and layoffs. These limitations require that EPA and state enforcers make difficult choices among the various evaluative tools that are available to them.

I believe there is room for reasonable people to differ as to which mix of environmental enforcement metrics to rely upon. Numbers alone can never fully describe the warp and woof of environmental enforcement work and, notwithstanding rhetorical claims to the contrary, no one can truly prove that any combination of measurements will provide the best balance of thoroughness and cost-effectiveness.

In my own view, notwithstanding continuing resource limitations, EPA and the states should continue to make use of the range of output measurements that have been employed for some time to measure enforcement success. Over time, perhaps, as advanced next-generation self-monitoring and measures are implemented, the number of on-site inspections made by environmental agencies may become less important gauges of their enforcement effectiveness. In the short term, however, they—along with the other ongoing measures of enforcement activity mentioned above—remain vital indicators of enforcement program effectiveness. Those output measurements will be far more valuable, however, if EPA uses workload modeling and other analytical tools to classify various types of enforcement actions according to their relative complexity and resource-intensiveness, and then reports on numbers of enforcement actions taken in each of the categories identified in its case classification matrix, as well as in the aggregate.

EPA and the states will also do well to continue to measure the quantity of pollutants that will not be released into the environment as a result of enforcement agreements and orders, and the health-cost benefits of enforcement-compelled cutbacks in emissions of air pollutants. Rather than merely report on overall quantities of pollutant releases that will be abated, however, environmental agencies should make an effort to speciate (i.e., categorize by type) the types of pollutants abated due to enforcement work. When they do so, Congress and the public will better understand the relative toxicity and human health impact of the masses of pollution that will be controlled as a result of enforcement work.

EPA and the states should also routinely make public the amounts of money allocated each year for the enforcement aspects of their Agency’s work, and the average and cumulative numbers of years in service, as enforcement professionals, of their staffs. Notably, the latter statistic will not always be telling. In some cases, “experienced” staffs...
may be experienced at doing substandard work. Nonetheless, enforcement staff years in service will often provide at least an approximate measure of an environmental agency’s enforcement capabilities.

Finally, as its Next Generation compliance efforts pick up steam, EPA should supplement its other enforcement recordkeeping with the kinds of metrics the Agency mentioned in its FY 2014-2018 Strategic Plan. Those measurements will provide useful insights into the effectiveness of that promising reform of the Agency’s enforcement and compliance work.

Other enforcement measures, although useful, may be too resource-intensive to employ in the short and medium term. Thus, though they may be revealing, surveys of employee morale and working conditions, and measures of the promptness with which agencies close out initiated enforcement cases, may need to be held in abeyance for now.

Resource limitations are real, and environmental agencies will likely be compelled to live within the constraints they impose for some time to come. Nonetheless, the transparency of Agency enforcement work remains a matter of paramount importance. Just as EPA intends to publicize the compliance records of regulated pollution sources, the Agency must also openly disclose as much information as possible that casts light on the effectiveness of its (and the states’) enforcement programs. While they impose costs and have imperfections, diligent recordkeeping and regular open disclosure of a broad and varied combination of environmental enforcement indicators are essential to the public’s right to know. These must receive high priority as environmental agencies approach their futures.