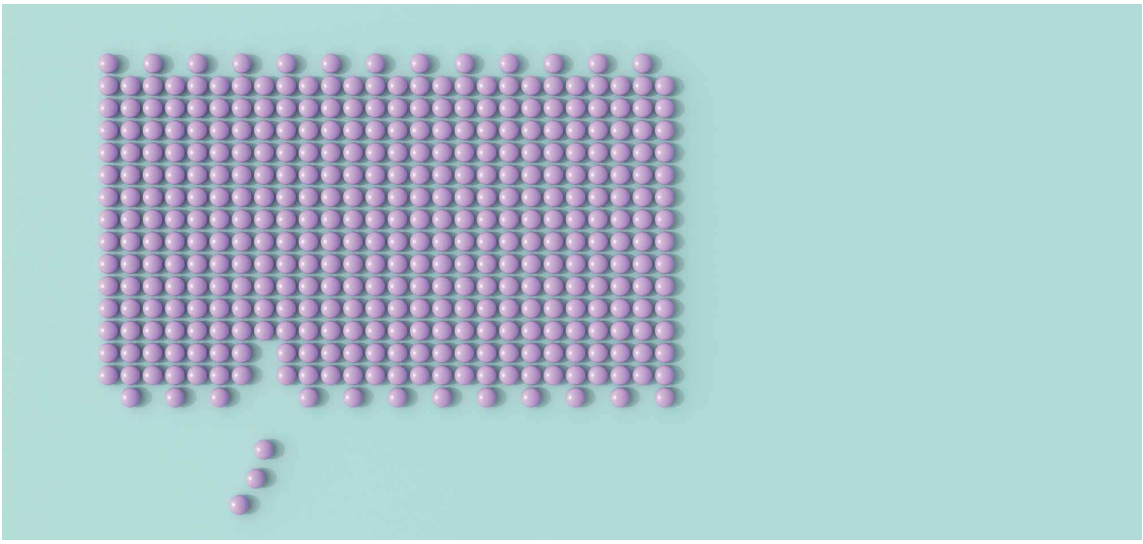
Harvard
Business
Review



REGULATION

Why Compliance Programs Fail —and How to Fix Them

by Hui Chen and Eugene Soltes

FROM THE MARCH–APRIL 2018 ISSUE

Millions of fraudulent accounts at Wells Fargo. Systemic deception by Volkswagen about its vehicles' emission levels. Widespread bribery at Petrobras that damaged both the government and the economy of Brazil. While those corporate scandals made headlines in recent years, countless others failed to penetrate the global consciousness. According to the Association of Certified Fraud Examiners, almost half of all fraud cases are never reported publicly, and a typical organization loses close to $3 million in annual revenue to fraud. Furthermore, of the nearly 3,000 executives interviewed for EY's 2016 Global Fraud Survey, 42% said they could justify unethical behavior to meet financial targets. Clearly, malfeasance remains deeply entrenched in private enterprises today.

The ubiquity of corporate misconduct is especially surprising given the staggering amount firms spend on compliance efforts—the training programs, hotlines, and other systems designed to prevent and detect violations of laws, regulations, and company policies. The average multinational spends several million dollars a year on compliance, while in highly regulated industries—like financial services and defense—the costs can be in the tens or even hundreds of

millions. Still, all these assessments deeply underestimate the true costs of compliance, because training and other compliance activities consume thousands of valuable employee hours every year.

Many executives are rightly frustrated about paying immense and growing compliance costs without seeing clear benefits. And yet they continue to invest—not because they think it's necessarily productive but because they fear exposing their organizations to greater liability should they fail to spend enough. Employees, too, often resent compliance programs, seeing them as a series of box-checking routines and mindless training exercises. In our view, all this expense and frustration is tragic—and avoidable.

We're both acutely aware of the perceptions and challenges surrounding compliance. From November 2015 until her resignation in June 2017, Chen served as the sole (and first-ever) compliance consultant at the U.S. Department of Justice (DOJ), advising prosecutors in evaluating the compliance efforts of companies under investigation. Soltes, in his research at Harvard Business School, has studied the obstacles general counsels and compliance officers face in ascertaining how well their programs work and explaining the benefits to others in their organizations. It's obvious to us that the value of compliance must be made clearer to company leaders and employees alike.

The answer, we believe, lies in better measurement. At its core, the idea is as simple as it is crucial: Firms cannot design effective compliance programs without effective measurement tools. For many firms, appropriate measurement can spur the creation of leaner and ultimately more-effective compliance programs. Put simply, better compliance measurement leads to better compliance management.

How We Got to This Point

To appreciate how compliance evolved into its current state, let's consider how such programs began. Following a stream of corporate scandals in the United States in the 1970s and 1980s, industry groups banded together and adopted internal policies and procedures for reporting and trying to prevent misconduct. Those efforts helped assuage legislators who had sought to more heavily regulate and penalize firms for dishonest practices. Self-policing appealed to business leaders as a way to avoid the cost and disruption of additional regulation. It also eased the investigative burden on regulators, and many people believed it would successfully deter wrongdoing.

Attracted by the perceived benefits, in 1991 the U.S. Sentencing Commission (USSC) amended its guidelines and offered firms substantially reduced fines if they could show that they had an "effective compliance program." A series of memoranda from senior officials at the DOJ soon followed, urging prosecutors to consider the effectiveness of a firm's compliance program when deciding on criminal charges. Those efforts were intended not only to encourage better monitoring by companies but also to recognize that firms can become victims of rogue employees. Other civil regulators, including the Securities and Exchange Commission, the U.S. Department of Health and Human Services, and the Environmental Protection Agency, also adopted this carrot-and-stick approach to compliance.

An industry quickly sprouted to provide compliance training programs, hotlines for whistle-blowers, and risk assessments. Not having a compliance program became a liability too significant for any major firm—even a foreign firm that simply utilized U.S. banks—to ignore. This potential liability has steadily increased as other countries, such as the United Kingdom, Brazil, and Spain, have enacted laws that take compliance into consideration in enforcement actions.

For many company leaders, compliance programs are protection against worst-case scenarios, akin to an expensive insurance policy. Employees may be asked to sign lengthy codes of conduct attesting that they know their firm's polices; additionally, they may sit through training programs on topics such as privacy, insider trading, and bribery. Yet individuals often pay only enough attention to these generic classes to pass the 10-question quiz at the end. Even at firms spending millions of dollars annually on their programs, compliance often lacks substance.

When the DOJ brought criminal charges against Morgan Stanley employee Garth Peterson in 2012, the prosecution documents noted that Peterson had received seven compliance training sessions and 35 related reminders to eschew the very conduct—bribing a government official—that he ultimately engaged in. But those compliance initiatives had little influence on Peterson because he viewed them as pro forma. "You can have programs and e-mails," he said, "but if people just delete them [or] if people have to do teleconferences but…instead of actually listening, all you have to do is say, 'Garth Peterson's on the phone,' [then] they check the box that says he's complied. And then you either quietly hang up, or you just put your phone aside and you do your other work."

The DOJ recognized that firms might be spending a lot and creating all the components of compliance programs but actually producing hollow facades. In its 2008 revision of the "Principles of Federal Prosecution of Business Organizations," the department specifically calls for prosecutors "to determine whether a corporation's compliance program is merely a 'paper program' or whether it was designed, implemented, reviewed, and revised, as appropriate, in an effective manner." The same year, in a case against Siemens in which a record-setting $800 million penalty was paid to the U.S. authorities, the prosecution repeatedly called out the inadequacies of Siemens's "paper program."

Over and over, prosecutors have recognized that firms with ineffective compliance programs don't deserve credit for their supposed efforts. However, it was often challenging to distinguish substantive programs from those that were merely window dressing, since evaluating a program required considerable time and expertise. The DOJ's decision not to prosecute Morgan Stanley in the Peterson case, for example, was seen as validating the firm's approach to ensuring compliance, which included numerous training sessions in addition to the standard hotline and the usual employee certifications of the firm's code of conduct. Yet Peterson claimed that the government was "lying to the public and saying that they [Morgan Stanley] had this wonderful compliance program, when in fact the government knows that it wasn't getting into people's heads, which is what really matters."

The DOJ retained Chen in the fall of 2015 to address the challenges of evaluating the actual effectiveness of firms' compliance efforts. Right from the start, she observed something amiss with many of the programs she examined. Companies routinely produced large binders of policies and procedures and counted the number of controls in their financial systems. And yet they offered no evidence of having tested those policies, procedures, and controls, nor did they track how many breaches they had experienced. A company might cite its long-standing internal whistle-blower program, for instance, but not have data on the program's rate of usage by employees. Firms also routinely reported how many times they had trained wrongdoers on the very topic of their misconduct, apparently blind to the irony of defending their compliance efforts that way.

In response to her mandate to focus on effectiveness, Chen drafted an extensive list of questions for prosecutors to consider when assessing compliance programs. The questions covered a wide range of compliance areas, including training ("What analysis has the company undertaken to determine who should be trained and on what subjects?"), individual accountability ("Were

managers held accountable for misconduct that occurred under their supervision?"), and leadership ("What compliance expertise has been available on the board of directors?"). The DOJ publicly released the questions in February 2017 in a document titled "Evaluation of Corporate Compliance Programs."

> Companies routinely had policies and procedures but did not track breaches.

The document was not intended to be used as a checklist; instead, as it stated, it listed "some important topics and sample questions that the Fraud Section has frequently found relevant in evaluating a corporate compliance program." Indeed, all evaluations would continue to be individualized. Nonetheless, as Soltes observed in his interactions with managers and corporate attorneys at the time, firms quickly began to appropriate the document as a manual on constructing an effective program. In particular, managers believed that if they could provide an answer to each question, they could assure themselves that they were meeting the DOJ's expectations. Even more worrisome, Soltes saw firms selectively picking data to support the notion that their practices were effective, rather than recognizing that some were clearly falling short.

For example, one question in the DOJ document asks firms how they evaluate the quality and effectiveness of their training. A survey by Deloitte and *Compliance Week* suggests that the most common way is to measure completion rates and to deem training effective if enough employees —perhaps 90% or 95%—finish it. However, that metric reflects neither the quality of a training (how appropriate and valuable the content is) nor its effectiveness (how much employees actually learn and put into practice).

Firms rely on completion rates not because doing so has been shown to be the "right way" to measure success but because their objective is merely to demonstrate to regulators that they've accomplished the task—they can check that training box. While some firms surely provide their employees with effective instruction about following the rules, we have seen many more delude themselves into believing that their training is satisfactory simply because it's been completed.

One of the main reasons that companies keep investing more and more in compliance is that they do not have the right measures and thus cannot tell what works and what doesn't. At many companies, strengthening compliance has become synonymous with hiring more compliance

managers, buying more-sophisticated software, and creating more policies, even when those moves are redundant and wasteful or just don't deliver results.

How Compliance Metrics Go Astray

According to Deloitte and *Compliance Week,* only 70% of firms even try to measure the effectiveness of their compliance programs. And of those that do, only a third are either confident or very confident that they are using the right metrics. In early 2017 the U.S. Department of Health and Human Services convened a meeting to develop metrics to help health care organizations better judge their compliance programs' effectiveness. The group produced a report detailing more than 550 different indicators. The report acknowledged that a given organization would need only a subset of those, tailored to the firm's specific business or risk profile. Still, with so many metrics to choose from, ascertaining which would be appropriate in which instances remains challenging and beyond the grasp of most firms.

In seeking to assess program effectiveness quantitatively, firms tend to make the same mistakes. Here are the common pitfalls:

Incomplete metrics.

The DOJ and USSC guidelines expect effective compliance programs to hold individuals accountable for violations. The DOJ evaluation document, for example, asks: "Has the company ever terminated or otherwise disciplined anyone for the type of misconduct at issue?" and "Have the disciplinary actions and incentives been fairly and consistently applied across the organization?" To demonstrate individual accountability, firms often list the employees who have been terminated or denied promotions and bonuses as a result of compliance-related transgressions. Yet such statistics aren't enough to substantiate that a firm rigorously holds employees accountable since they don't indicate the number of employees who were *not* disciplined. A firm that disciplines five employees because five people behaved improperly during the year is very different from one that sanctions five employees out of the 50 who violated company policies. We've seen firms punish lower-level employees or those with less potential, while protecting high earners or senior executives. So the simple statistic on the number of sanctioned employees can be incomplete and misleading.

Invalid metrics.

Although a wide range of data may be collected on the various facets of a compliance program, only a subset of that data actually correlates with the impact of a program. For example, in response to the DOJ question asking how the company has measured the effectiveness of its training, firms often focus on the percentage of employees who've completed the training, as we noted earlier, or the number of hours they've spent doing so. Those are entirely the wrong metrics to use. Completion rates may be relevant for a firm to track for other purposes, but a meaningful measure of effectiveness must be directly tied to a clearly articulated outcome—for example, employees' demonstrated understanding of policies and procedures, their acquisition of useful skills for confronting anticipated scenarios, or a change in their behavior.

How Effective Is Your Compliance Program?

When the U.S. Department of Justice prosecutes a company, it evaluates the effectiveness of the organization's compliance program. Below are key topics and sample questions that the Fraud Section considers, excerpted from its 2017 document titled "Evaluation of Corporate Compliance Programs."

Senior and Middle Management

- How have senior leaders, through their words and actions, encouraged or discouraged the type of misconduct in question?
- What types of information have the board of directors and senior management examined in their exercise of oversight?

Autonomy and Resources

- What has been the turnover rate for compliance and relevant control function personnel?
- Have there been specific transactions or deals that were stopped, modified, or more closely examined as a result of

As another example, to back up the assertion that management has a "strong" commitment to compliance, firms may cite the number of pro-compliance communications that top executives issue. However, such a metric is invalid if employee surveys show a lack of trust in management and a belief that whistle-blowers face retaliation.

Mistaking legal accountability for compliance effectiveness.

Compliance policies serve important legal functions, but forcing them into legal frameworks may limit their ability to positively influence employee behavior. Take this question: "How has the company assessed whether these policies and procedures have been effectively implemented?" Firms often respond by showing that employees signed a statement that they had read and understood the company's policies and codes of conduct. While such a signature may provide legal grounds to fire someone who violates a rule, it does not demonstrate that an employee has converted knowledge about policies into everyday work practices. How many times do

compliance concerns?

Policies and Procedures

- How has the company assessed whether [applicable] policies and procedures have been effectively implemented?
- How has the company evaluated the usefulness of these policies and procedures?

Risk Assessment

- What methodology has the company used to identify, analyze, and address the particular risks it faced?
- What information or metrics has the company collected and used to help detect the type of misconduct in question?

Training and Communications

- How has the company assessed whether its employees know when to seek advice and whether they would be willing to do so?
- How has the company measured the effectiveness of [employees'] training?

Confidential Reporting and Investigation

- How has the company collected, analyzed, and used information from its reporting mechanisms?
- How high up in the company do investigative findings go?

Incentives and Disciplinary Measures

- What is the company's record (e.g., number and types of disciplinary actions) on employee discipline

we all reflexively assent to the legal terms of an agreement, especially those that we have no power to negotiate? Employees may sign an acknowledgment of corporate policies without actually having read or understood the terms. Moreover, the policies may be hard to grasp because they are written in language that is legalistic, technical, or just plain dense. There could also be an implicit understanding within the firm that the policies don't really have to be followed or that best practices can be improvised. Thus, counting employees' legally binding assents to policies is not an appropriate way to quantify the effectiveness of a compliance initiative.

Self-reporting and self-selection bias. Compliance managers often rely on surveys to assess the performance of their programs. For instance, to gauge employee comfort with reporting mechanisms, a firm might ask: "Do you know when to seek compliance advice? Are you willing to do so?" The challenge with surveys is that self-reporting and self-selection by the respondents may bias the results and lead managers to draw incorrect conclusions. Employees who have observed dishonest behavior, for example, may be reluctant to "out" their colleagues and may choose not to answer related survey questions, which will skew the results toward employees who have not observed wrongdoing. Similarly, people in senior positions and those who actually do engage in misconduct may be less inclined to

relating to the type(s) of conduct at issue?

- Have the disciplinary actions and incentives been fairly and consistently applied across the organization?

Continuous Improvement, Periodic Testing, and Review

- Has the company reviewed and audited its compliance program, including testing of relevant controls, collection and analysis of compliance data, and interviews of employees and third parties?
- What types of relevant audit findings and remediation progress have been reported to management and the board on a regular basis?

Third-Party Management

- How has the company monitored the third parties in question?
- How has the company incentivized compliance and ethical behavior by third parties?

participate. Thus, bias in the data collected needs to be accounted for when interpreting the metrics.

Linking Compliance Initiatives to Objectives

So how *do* you create models that can credibly evaluate the impact of a compliance program? The first step is recognizing that such programs actually have multiple purposes. As laid out in numerous memoranda by senior officials at the DOJ, the three main goals are to prevent misconduct, to detect misconduct, and to align corporate policies with laws, rules, and regulations. Each component of a compliance program should be linked to one of these objectives. For example, training serves to prevent misconduct, whistle-blower hotlines are designed to detect it, and codes of conduct are intended to align employees' behavior with company policies and external regulations. Although it's possible that one compliance initiative will overlap with or impact another, clearly identifying the goals of each will help managers create more-meaningful metrics.

Consider a confidential hotline for whistleblowers. Its objective is to improve the timely detection of wrongdoing. To understand whether it's achieving this goal, several pieces of information are needed, including whether the hotline works ("mystery tester" reports), whether people actually use it (usage data), how they use it (data on types of calls received), the firm's responsiveness to allegations (response time, investigation completion time, investigation results, communication of results), and whether employees feel comfortable contacting the hotline (periodic surveys of employees' sentiments). Each of those metrics captures a different dimension of the initiative's efficacy.

However, tracking those variables independently is insufficient, because it doesn't allow managers to identify which ones are responsible for particular outcomes. For instance, a "hot" hotline might reflect a rising number of problems or just a high level of employee comfort with calling. To get clarification, managers can apply multivariate regression analysis. Regression models allow an investigator to examine the impact of one variable while holding the others constant. In this case, to ascertain whether an increase in calls indicates an increase in compliance breaches, we would seek to hold the following other factors constant: the availability of the hotline, people's comfort in using it, its operational performance, and the number of potential callers (people who have access to it). Designing appropriate regression models takes time and experience, but it is the most reliable way to know whether to be reassured by or concerned about shifts in call volume.

Let's take another example—compliance training, the objective of which is to prevent misconduct by helping employees internalize rules and regulations. Assessing how well employees understand what's expected of them after they complete training is, by itself, insufficient to establish the training's effectiveness. A high degree of understanding could reflect the positive influence of the instruction they received, but it could also simply reflect employees' baseline knowledge. Therefore, firms must assess what employees know both before and after training. If there is little change, the training may be unnecessary, or it may need to be refined to more fully engage people and make better use of their time.

Of course, the goal of training is not only to improve employees' understanding of the rules but also to instill and perpetuate appropriate behavior. Again, a regression model can help firms understand the link between training sessions and changes in employee behavior. By controlling for the other factors that may contribute to policy violations, we can test whether the individuals who undergo training become more or less inclined to break the rules.

As these examples demonstrate, firms should use empirical data generated from their compliance programs to gauge how well a program is meeting its objectives. Again, we stress that firms need to do more than simply track metrics independently. They must focus on creating models that measure the desired output while controlling or excluding other factors.

In the past, firms trying to show the effectiveness of their programs might have been able to offer metrics that were not well aligned with compliance objectives, but the standards and stakes are changing. Prosecutors, courts, and regulators increasingly seek more-rigorous evidence. This

means that firms must have the capacity to back up compliance claims with better data and models—a process that's possible only when the capabilities to accurately measure a program's performance are in place.

Compliance Engineering

Some companies may be willing to invest significant time and resources in compliance and ethics programs because they see them as critical to the organization's long-term success. But we're pragmatists. We understand that with all the other competing demands on a firm's limited resources, the ever-present regulatory and liability concerns often become the rationale driving compliance efforts. Yet this focus on the regulatory aspect is exactly why it's critical to get serious about measuring outcomes. As compliance programs continue to be more closely scrutinized, those that cannot show meaningful results will fail to meet the stronger regulatory standards being applied today. To put it more bluntly, if the best that can be said for, say, an anti-corruption training course is that employees finish it, prosecutors, courts, and regulators are not going to give a company credit for having an effective program.

While many firms continue to see ensuring compliance as a legal exercise, it is really much more a behavioral science. That assertion may make attorneys uncomfortable, but for compliance programs to have real impact, managers need to test what works and what doesn't. This will require firms to engage in some experimentation and innovation. Codes of conduct should articulate policies that are core to a firm's success. And hotlines should exist not only to record reports of wrongdoing but also to help employees resolve predicaments before they make a bad move. By developing better measures of effectiveness, firms can adopt more ambitious and innovative programs that really do curb improper behavior.

Given all the complex regulations governing business today, it's no wonder that companies struggle to understand and meet their legal and ethical obligations. It would be convenient if there were a one-size-fits-all yardstick that could show if a compliance program is on track or not. But simple univariate metrics will not adequately capture a program's effectiveness. Successful compliance engineering requires some creativity, some testing, and careful model design to appropriately measure outcomes.

Companies worldwide are already spending a fortune on compliance. Let's make sure that all those resources are being spent well. Better measurement can help managers identify redundant or ineffective initiatives that can be replaced or eliminated—and ultimately reveal opportunities

to make programs more effective.

A version of this article appeared in the March–April 2018 issue (pp.116–125) of *Harvard Business Review*.

Hui Chen, formerly the compliance expert at the U.S. Department of Justice, is an ethics and compliance consultant to government regulators and companies worldwide.

Eugene Soltes is the Jakurski Family Associate Professor of Business Administration at Harvard Business School, where his research focuses on corporate misconduct.

This article is about REGULATION

⊕ **FOLLOW** THIS TOPIC

Related Topics: BUSINESS LAW | DATA

Comments

Leave a Comment

POST

0 COMMENTS

JOIN THE CONVERSATION

POSTING GUIDELINES

We hope the conversations that take place on HBR.org will be energetic, constructive, and thought-provoking. To comment, readers must sign in or register. And to ensure the quality of the discussion, our moderating team will review all comments and may edit them for clarity, length, and relevance. Comments that are overly promotional, mean-spirited, or off-topic may be deleted per the moderators' judgment. All postings become the property of Harvard Business Publishing.