

Measure to Comply, Measure to Perform



**A GOVERNMENT
PERFORMANCE
WHITE PAPER**

**BY DR. SHELLEY H. METZENBAUM,
UNIVERSITY OF MARYLAND**

While every attempt has been made to ensure that the information in this document is accurate and complete, some typographical errors or technical inaccuracies may exist. Cognos does not accept responsibility for any kind of loss resulting from the use of information contained in this document.

This page shows the publication date. The information contained in this document is subject to change without notice.

This text contains proprietary information, which is protected by copyright. All rights are reserved. No part of this document may be photocopied, reproduced, stored in a retrieval system, transmitted in any form or by any means, or translated into another language without the prior written consent of Cognos Incorporated.

The incorporation of the product attributes discussed in these materials into any release or upgrade of any Cognos software product – as well as the timing of any such release or upgrade – is at the sole discretion of Cognos.

U.S. Government Restricted Rights. The accompanying materials are provided with Restricted Rights. Use, duplication for disclosure by the Government is subject to the restrictions in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013, or subparagraphs (c) (1) and (2) of the Commercial Computer Software – Restricted Rights at 48CFR52.227-19, as applicable. The Contractor is Cognos Corporation, 67 South Bedford Street, Burlington, MA 01803-5164.

This edition published February 2008
Copyright © 1989-2008 Cognos Incorporated.

Table of Contents

- Introduction** **3**
- Detecting problems and successes** **4**
 - Relative Frequency of Unwanted Events 4
 - Characteristics of Unwanted Events 5
 - Frequency and Characteristics of Demand. 7
 - Frequency and Characteristics of Beneficial Events 7
 - Performance Trends—Dates of Agency Action, Directional Change,
Rapid Acceleration, and Pattern Divergence 7
 - Outliers, Exceptions, Anomalies, and Unexplained Variations 8
 - Measurement Relationships—Pre-cursor and Predictor Indicators 9
 - Putting a Problem on the Action Agenda. 10
- Diagnostics and treatment** **11**
 - Practical Diagnostics 11
 - Data-driven Diagnostics 11
 - Data-driven Experiments 12
 - Mapping 12
 - Outliers and Exceptions 13
 - Repeat Unwanted Events 13
- Dissemination** **14**
 - Promote Better Practices 14
 - Competition, Demand Management and Matching Consumer Tastes 14
 - External Analytic Assistance 15
 - Dissemination Limits 15
- Summary** **16**
- About Cognos** **16**

About the author

Shelley H. Metzenbaum is a Visiting Professor at the University of Maryland's School of Public Policy, where she leads a consortium that brings state agencies together to develop better ways to measure and manage compliance and enforcement programs. Dr. Metzenbaum holds a Ph.D in public policy from the Kennedy School of Government, Harvard University, and is the former head of that school's Performance Management Project. She writes, teaches, and consults on performance management and the use of performance goals and measures as powerful management tools for government.

Introduction

Performance measurement and reporting is on the rise in governments all around the world. Too often, however, it is executed as a mindless response to a measurement mandate from a legislative body, elected executive, budget office, or other funding source. When governments measure simply to comply with these mandates or to keep up with trends, when they only compile data into annual reports to remain accountable, they miss the rich opportunity of measurement to inform decisions and improve performance. And, they often spend money unnecessarily.

Public goal-setting and performance reporting is valuable: goal-setting compels government agencies to clarify their priorities, and measurements make the public aware of what they have accomplished. But public organizations can go beyond this limited use of goals and measures. They can use measures to identify problems that need attention and successes that warrant replication, and they can disseminate those findings to spread effective practices. When they do, the information they gather is useful, not just a reporting obligation but a tool.

At the same time that measurement mandates have increased, many agencies have also installed more integrated, agile information systems. These systems make it feasible for anyone in an organization from almost any location to access vast quantities of data almost immediately. They let organizations slash marginal costs of high-volume activities such as billing or information provision. They allow Internet transactions, saving money and time.

But too few governments use their agile systems and the rich store of information contained within them to their full extent. This paper describes how some government agencies have gone beyond using data simply to improve operational efficiency and report performance. These agencies have, in addition, learned to find stories in their data to achieve better societal outcomes at a lower cost and enlist expertise and analytic talent from outside their agencies. Adopting the analytic attitude suggested in this paper will vastly improve returns on government's investment in measurement and data systems.

Detecting problems and successes

Detection describes the practice of scanning for societal problems and program-related ones. It also refers to the hunt for program successes worth replicating. Detection helps answer simple questions such as when, where, what, who, how many, how often, how long, and how serious. Governments can begin this search by looking for patterns such as repeat occurrences, steady trends, geographic clusters, and anomalies. Interesting stories begin to emerge from these patterns and variations that prompt further questions. This section illustrates how some organizations organize and analyze information in their databases to aid their detection efforts.

Relative Frequency of Unwanted Events

Many agencies start the search for performance-improving insights by focusing on harmful events, risks, or problem conditions they seek to reduce. Knowing relative problem frequency and consequence helps set priorities, allocate resources, and tailor intervention tactics to the problem.

This approach has been effectively used by a federal marine safety program that seeks to reduce fatalities and injuries on the water. Initially, the safety office counted only ship and boat inspections and enforcement actions as indicators of program performance. When a law was passed requiring the program to shift its emphasis to outcomes, the agency switched to counting fatality rates. It discovered that tow boats have the highest fatality rate of any other commercial vessel, and therefore made tow boat safety an agency priority. Counting the relative frequency of unwanted events helped the agency work with the industry to understand the kinds and causes of accidents, address the causes, and eventually cut the fatality rate in half.

Other agencies similarly benefit by shifting attention from tallying program activities, such as inspections, to counting the frequency and severity of unwanted events. Public health agencies can count health problems, such as asthma incidents, to determine when, how often, where, and in whom they are occurring. Regulatory agencies can tally permit violations and unwanted incidents. A service agency might count response and processing times that exceed a maximum threshold. Measuring unwanted events helps agencies detect emerging problems. Programs can then test different treatments, count the number of unwanted events before and after the treatments, and determine if the treatments effectively reduced the number and severity of incidents.

The following chart shows how a school system is tallying test scores to learn from them. The green circles show areas with a performance gain, while the red circle indicates an area with a performance decline. Identifying educational topic areas in which students most frequently incorrectly answer test questions draws attention to areas needing improvement. The school system can conduct cross-teacher analysis to determine whether patterns are similar among all teachers (suggesting a need for curriculum improvements) or unique to one or a few teachers (suggesting the need for teacher training or team teaching if the pattern persists). With additional analysis of the test results not shown here, the system can also detect which students are having trouble in each subject area so they can be given special help.

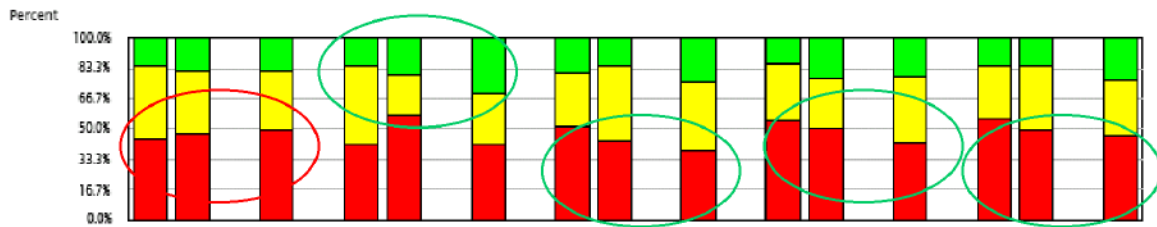
OK Cancel

Diagnostic Elementary School Mathematics Thirds by Strands FY03 vs FY04 - (Drill Down) - % - PRDIM0211

School Type/School/Class-Teacher: WsES 0351,
 Current Grade: 04, Student Group: Standard Curriculum (Yes and No).

NCLB: Gender: All Genders, Ethnicity: All Ethnicities, ESE: ESE(Yes and No), ESOL: ESOL(Yes and No), FRL: Free/Reduced Lunch(Yes and No)

■ Low
 ■ Medium
 ■ High



	Alg. Thinking				Data Analysis				Geometry				Measurement				Number Sense			
	FALL		WINTER		FALL		WINTER		FALL		WINTER		FALL		WINTER		FALL		WINTER	
	FY03	FY04	FY03	FY04	FY03	FY04	FY03	FY04	FY03	FY04	FY03	FY04	FY03	FY04	FY03	FY04	FY03	FY04	FY03	FY04
Low	44%	47%	/0	49%	41%	57%	/0	41%	52%	43%	/0	38%	55%	50%	/0	42%	56%	49%	/0	47%
Medium	40%	35%	/0	32%	43%	22%	/0	28%	29%	42%	/0	37%	31%	28%	/0	37%	29%	36%	/0	30%
High	15%	18%	/0	19%	15%	21%	/0	31%	19%	15%	/0	25%	14%	22%	/0	21%	15%	15%	/0	23%

Counting the frequency of unwanted events also helps agencies avoid the danger of directing resources from the highest risk problems to events that grab attention because they are unusual or “squeaky wheels.” In addition, counting unwanted events lessens the chance organizations will overlook risky situations that have become so common they are no longer recognized as dangerous. When hospitals began systematically counting medical errors, for example, they discovered numerous problems of high consequence that had previously received little attention. One of the most notable was a high number of operations on the wrong side of the body. Once unwanted events are recognized, it is far easier to devise effective means for cutting their frequency.

Characteristics of Unwanted Events

The search for performance-improving insights is greatly enhanced when events are not just counted but characterized to note who, where, why, when, and how long. In studying oil spills, an environmental protection agency noticed that most spills occurred at night. However, the office conducted all of its inspections during the day. When it shifted some inspections to the evening hours, the number of oil spills dropped, an easily noted effect because the agency routinely counts oil spills and notes the time of day, day of week, and week of year they occur. By studying its data, the agency was able to detect a problem, test a solution, and determine its effectiveness.

Highway safety agencies track the characteristics of traffic accidents, such as the age of the driver, location (intersection, number of lanes), type of vehicle, alcohol use, and seat-belt use. Characterizing accidents helps identify areas needing priority attention, prompts causal analysis, and suggests solutions tailored to problem specifics. The chart below naturally triggers the quest to understand variations in state fatality rates.

4 Traffic Safety Facts 1999 — Pedalcyclists

Table 3. Pedalcyclist Traffic Fatalities and Fatality Rates by State, 1999

State	Total Traffic Fatalities	Resident Population (thousands)	Pedalcyclist Fatalities	Percent of Total	Pedalcyclist Fatalities per Million Population
Alabama	1,138	4,370	3	0.3	0.69
Alaska	76	620	2	2.6	3.23
Arizona	1,024	4,778	26	2.6	5.44
Arkansas	684	2,551	6	1.0	2.35
California	3,669	33,145	112	3.1	3.38
Colorado	626	4,056	5	0.8	1.23
Connecticut	281	3,282	3	1.0	0.91
Delaware	160	754	1	1.0	1.33
Distrot of Columbia	41	519	1	2.4	1.93
Florida	2,275	15,311	121	5.2	3.11
Georgia	1,646	7,788	22	1.5	2.22
Hawaii	98	1,185	1	1.0	0.84
Idaho	278	1,252	4	1.4	3.19
Illinois	1,466	12,138	28	1.9	2.31
Indiana	1,013	5,943	14	1.4	2.36
Iowa	480	2,889	6	1.2	2.09
Kansas	527	2,654	7	1.3	2.64
Kentucky	814	3,881	10	1.2	2.52
Louisiana	924	4,372	29	3.1	6.63
Maine	181	1,253	1	0.6	0.80
Maryland	600	5,372	6	1.0	1.16
Massachusetts	414	6,175	6	1.4	0.97
.....

Data revealing the characteristics of the most frequent and serious accidents can also help one agency enlist cooperation from others. For example, unsafe railroad crossings can be a leading cause of automobile accidents. Making crossings safer requires numerous agencies including highway, rail, local public works, and local public safety officials to work together. Specific data documenting the most dangerous crossings and the most dangerous kinds of crossings makes it less contentious when national safety agencies seek to persuade local agencies to fix a problem. Specific data also supports inter-agency coordination.

Agencies can benefit from looking at not only *what* kinds of unwanted events occur, but also *at who* is associated with them. Fifth-grade teachers at a primary school discovered the benefits of studying the *who* characteristics of performance data. They noticed that the students of one fifth-grade teacher did better in science than students of other fifth-grade teachers. To improve student learning and test scores, the teachers opted to teach as a team so that all fifth-grade students would study science with the one teacher whose students excelled in that area.

Being able to predict unexpected events helps minimize or manage them appropriately. A government contract agency tracks unexpected events associated with military contracts, noting important characteristics of the events. In one case, historical analysis revealed that a certain type of unexpected repair had historically taken about six hours. When a contractor submitted an over-and-above work request estimating that an additional 20 hours would be needed, the agency was able to negotiate a far lower repair cost.

Agencies stuck in crisis response mode may benefit by counting and characterizing in-house crises, which tend to consume much senior management and public attention.

Frequency and Characteristics of Demand

Counting and characterizing service demands, especially for high-volume government operations, can be beneficial. Many websites and call centers count and categorize the questions they receive to identify frequently asked questions (FAQ) and then pre-package boilerplate responses. This can greatly reduce response time, especially when agencies can shift callers to websites to get responses to the FAQ. Many agencies also track common timing patterns, such as more phone calls on Monday, more contracts at the end of the fiscal year, or more building permits in the spring. This data helps organizations schedule staff, market off-peak times, or redesign processes to ward off backlogs.

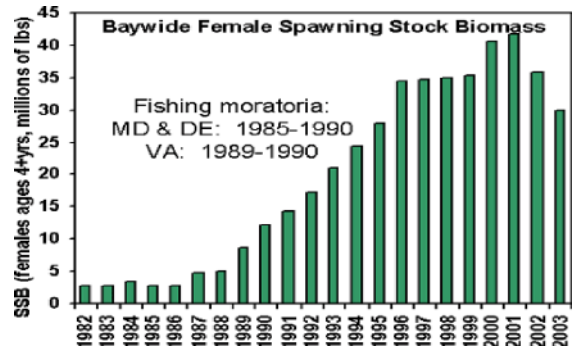
Frequency and Characteristics of Beneficial Events

The focus of measurement need not always be on problems. Insights can come from counting and characterizing positive events and upward shifts in societal conditions. Purchases of healthy food, increases in student test scores, drops in teen pregnancy rates, and increases in air and water quality are all beneficial events.

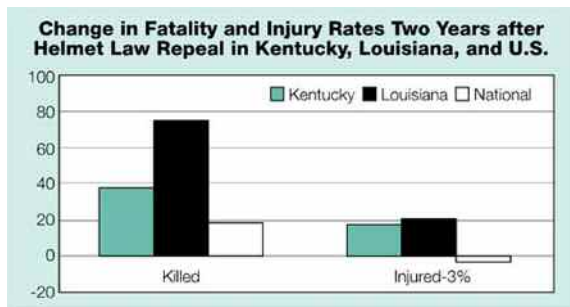
Knowing the characteristics of beneficial events can guide the search for effective solutions. An environmental agency catalogues and characterizes water quality improvement success stories, providing valuable models for replication. An educational not-for-profit organization identifies schools with high-minority and high-poverty populations that demonstrate high performance. These successes are worth studying to uncover contributors to their success.

Performance Trends—Dates of Agency Action, Directional Change, Rapid Acceleration, and Pattern Divergence

Tracking the dates of relevant policy or private sector actions alongside outcomes sheds light on the possible effects of those actions. In the chart below, the steady upward trend in the health of spawned fish between 1986 and 2001 compares favorably to earlier levels. The government-produced chart, communicating the trend to the public, suggests that fishing moratoria work and are worth replicating in other communities. The same chart also shows a downward trend in 2002, which should trigger an inquiry to understand the cause.



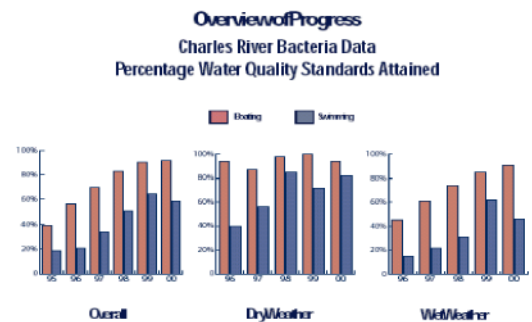
A national highway safety agency's risk-based analysis on motorcycle accidents links the number and cost of motorcycle fatalities to local laws on helmet use. Tracking the laws lets the agency learn from natural experiments that occur when laws change. So when two states repealed their motorcycle helmet laws, the safety agency was able to use the data it collected to determine how helmet use affects the fatality and injury rate.



This analysis helped the national agency and other interested parties make the case to elected officials for policies requiring safer practices, for incentive grants, and for sanction-linked regulatory authority to motivate other levels of government to adopt safer practices.

Monitoring the rate of change also informs the search for possible problems or successes. In one case, a media outlet compared students' average test scores by grade in each school to test scores for the same set of students the following year. The paper found 400 schools out of 7700 statewide with test scores three standard deviations away from test scores for the same students the previous year. The rate of change raised suspicion of a cheating problem.

Looking for divergences in trend lines, sliced by *who*, *what*, *why*, *where*, and *when* characteristics, can also reveal problems and solutions. For example, storms affect water quality because they stir up contaminated sediments and send run-off from oil-laden roads and fertilizer-fed lands into bodies of water. The following graph shows that solutions have been found to reduce most of the dry weather contamination problems and some wet weather problems at a particular river. Supplementary analysis revealed that the bulk of the remaining wet weather problems occurred at just five out of 37 monitoring locations. The data-driven refinement of the problem description allowed the agency to refine its solution and concentrate on problems in those five areas.



Outliers, Exceptions, Anomalies, and Unexplained Variations

Organizations often focus on performance averages, but much can be learned by studying data at the upper and lower end of a distribution.

Performance extremes are often simply statistical artifacts. An exceptionally brilliant or incapable child may skew a class or even a school's average performance scores upward or downward. Other times, odd amounts are simply data entry errors. Statistical techniques such as multi-year averaging and outlier elimination have been developed to counter such distortions.

Variations can sometimes be as instructive as norms, however. Many agencies measure timeliness as the percentage of payments, responses, or deliveries made in less than a target time. But what about agency actions that exceed targeted deadlines? How long did they take to resolve? Studying these cases to understand why they occurred often helps reduce recurrences.

The preceding chart showed how the water quality of the Charles River improved notably over a six-year period, especially on dry days. These improvements happened because an agency official paid attention to measurement variation. He noticed a downstream site with a lower water quality reading than an upstream one, not explained by a permitted waste water discharger between them. Why? To answer that question, he and some local sewer officials walked the pipes and detected the problem: an illegal hook-up to the storm sewer was pouring untreated sewage directly into the river. The facility was penalized and forced to fix the problem immediately.

Could similar problems exist elsewhere? The agency came up with a creative approach: it encouraged local officials along the Charles River to lift manhole covers for the storm sewers on dry days to look for running water. This revealed numerous illegal dischargers; appropriate actions were taken and outcomes improved significantly.¹

Measurement Relationships—Pre-cursor and Predictor Indicators

Finding a relationship between measurements can help an agency manage rare but high consequence events such as a terrorist attack or a nuclear accident. It can also help agencies manage slowly emerging problems such as cancer or ecological damage. By linking an easily measured precursor or predictor indicator with a rare or delayed negative event that is harder to pin down, an agency can anticipate and manage such outcomes.

An environmental agency decided to analyze complaints to assess their value as a predictor indicator. This is interesting because government agencies often treat complaints as a nuisance. Some even discount complaints, based on the perception that most pertain to minor or local problems, and that they vary with how well complaint lines are advertised. Some agencies, however, appreciate the problem-detection potential of complaints and have made them a core component of their operations.

¹ Metzenbaum, "Measurement that Matters: Cleaning Up the Charles River," in Donald F. Kettl, ed., *Environmental Governance: A Report on the Next Generation of Environmental Policy* (Brookings Institution Press, 2002), pp 58-117.

The agency found that, for some programs, most notably the waste management and air quality program, complaints were a useful problem detection tool. Complaints detected a significant proportion of total violations. Many of these violations would never have been found through agency inspections, which focus on permit holders. Complaints identified facilities that were operating without a permit or were infrequently inspected.

Identifying and managing pre-cursor events can also prove a powerful performance improvement tool. An aviation agency measures and manages tarmac incursions to lessen the likelihood of threats to aircraft, whether accidental or intentional. Fire codes now require the installation of smoke detectors, because analysis showed how often and by how much smoke preceded fires. A health agency has started to monitor obesity, realizing how frequently it is a pre-cursor to many serious illnesses and thus a problem worth controlling in itself.

Putting a Problem on the Action Agenda

Detecting a previously unknown or sudden change in a problem often nominates it for public attention, but it does not necessarily imply that an organization should immediately fix it. Citizens, their elected officials, and government agencies must first consider how important the problem is relative to other known problems. They may also want to understand the cause of the problem first if they only measured the symptom, so a more precise treatment can be developed. Similarly, finding effective solutions does not necessarily warrant replication. The decision to apply a known solution depends on other demands on the organization, the cost of the treatment, values-based choices of the political leadership, and the external environment.

Measurement is powerful both for reporting performance and for detecting problems and successes. Numerous tracking and analytic techniques can help agencies in their search. Detecting patterns and anomalies does not always indicate that a problem exists or that it needs attention, but it does trigger focused follow-up questions. It narrows the search for underlying causes and better solutions, the topics discussed in the next section.



Diagnostics and treatment

If detection involves the use of data to answer questions such as when, where, who, and what, diagnostics is the pursuit of answers to questions such as why and how. This section illustrates various diagnostic approaches, ranging from conversation and on-site investigations to more formal analytic methods. Diagnostics is used here not as a precise, scientific term, but as a reminder to those building their measurement and information systems that using data to find problems and successes is only a first step toward performance improvement.

Practical Diagnostics

Diagnostic methods can be as simple as an official walking the pipes after detecting an unexplained water quality reading, as described earlier. Detection of an illegal hook-up led to detection of similar problems. Once the problem was understood, a solution became readily apparent. Implementing it quickly improved water quality.

When ready causal explanations are not available in collected data (it is, of course, too costly to collect all possibly relevant data all the time), government organizations can drill down in less technical ways. To find explanations for high performance gains, schools may ask students, their families, and their teachers why they think it happened. This may reveal a common factor or two that invites further analysis.

One school system is working with high-performing teachers of ethnically and economically diverse students, identified through measurement, to discover aspects of their teaching talents that might be replicable. By videotaping the classes of these high-performing teachers and talking with them to understand what they do, the school system hopes to find effective methods other teachers can adopt. The system uses its data to detect success, then conducts in-depth observations to identify “treatments” other teachers may want to adopt.

Data-driven Diagnostics

Other times, diagnostics entail far more data analysis: drilling down into existing data sets to search for possible causal factors. Just as a school system can identify students having problems, it can identify students making progress.

After drilling down to identify students with sustained performance gains, it can follow-up with additional data analysis and interviews to determine if a student’s progress is linked to a specific teacher, teaching method, computer use, or additional assistance outside the classroom. These findings, in turn, can guide future decisions about how best to serve other similar students.

Most Reading Gain

04 03

Area: DC		FY03 - FY04 FCAT 555 Reading Developmental Score Gains													Date: 9/18/05								
School: Current Grade Level: 06		100% of 167 All Students in Levels - LV1, LV2, LV3, LV4, LV5 (167 Students)																					
		Start Direction: Descending - Field Sorted: FCAT 555 Developmental Score Gain																					
ST#	Last	First	M	FCAT 555		FCAT 555		FCAT 555		FCAT 555		FCAT 555		App	Retain	P-#	Race	LFP	DOB	Gender			
				2003	2004	2003	2004	2003	2004	2003	2004	2003	2004										
1																							
2																							
3																							
4																							
5																							
6																							
7																							
8																							
9																							
10																							
11																							
12																							
13																							
14																							
15																							
16																							
17																							
18																							
19																							
20																							
21																							
22																							
23																							
24																							

FY03 - FY04 FCAT 555 Reading Developmental Score Gains - 9/18/05 Page 1

When gathering accident data, a highway safety agency gathers information about three distinct stages of each unwanted event: accident characteristics, pre-accident information (such as causal factors), and post-accident data (such as accident consequences). This data supports a risk-based analysis of accidents and strengthens the case for policy changes. In the following chart, postaccident data shows how much higher hospital costs are for vehicle occupants not wearing seat belts, strengthening the case for seat-belt use policies.

**Exhibit 7.
Average Inpatient Charge by Safety Belt Use for Inpatient and All Crash-Involved Passenger Vehicle Drivers in the CODES States**

Group	Safety Belt Use		Increase for Not Using Belts
	Used	Not Used	
Inpatient Victims	\$9,004	\$13,937	55%
All Crash-Involved Drivers	\$110	\$562	408%

Data-driven Experiments

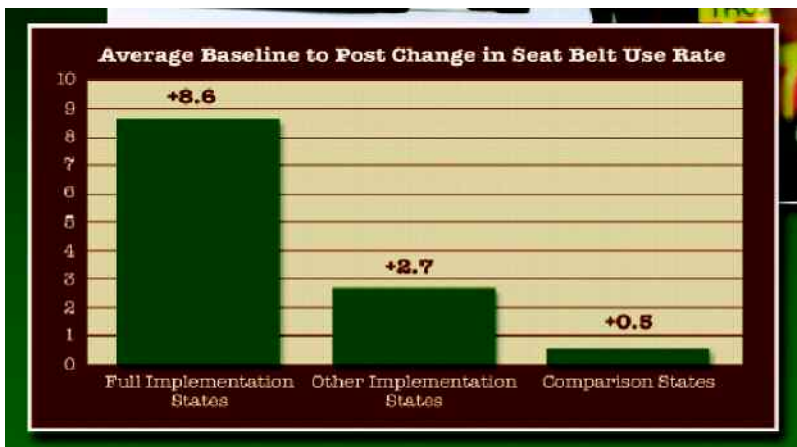
The same national safety agency also conducts controlled experiments, sometimes involving local governments. It enlisted local agencies to test a model program to increase seat belt use by issuing tickets for not buckling up. This program had effectively been used by a local government and in another country. The agency asked several local communities to implement the buckle-up campaign designed by the national agency, and asked a few communities to do nothing, acting as a control group. It also allowed some communities to test their own method.

It funded communities in all three groups to measure behavioral change in the same way, using observers at busy intersections before and after the campaign to count seat belt use. As the results shown below indicate, the experiment's findings helped the national agency make a strong argument to local governments to adopt the agency's suggested approach.

A few further analytic techniques are helpful to diagnostic and treatment efforts.

Mapping

Police have long appreciated the value of mapping crimes to see crime clusters. Push pins on wall maps have evolved into computer-drawn pictures, but the story is the same. Mapping crime lets police deploy their workforce appropriately and detect crime sprees likely to be linked to specific perpetrators. Crime is not the only location-linked problem that can be tackled more effectively with maps. John Snow demonstrated the value of mapping for diagnostic purposes in 1854, for example, when he mapped cholera cases in London and noticed that most clustered around a single water pump. He had the handle removed, and cholera cases dropped steeply.



More recently, a government agricultural service helped a local farmer by mapping fungus on his land, leading to the detection of an underground conduit that was spreading the fungus. Removal of the pipe eliminated the problem, resulting in less fungicide use and higher yield.

Outliers and Exceptions

Analyzing outliers and exceptions can also generate insights. Was a permit deadline missed because the permit was put into the wrong pile, arrived at the agency when one of the people responsible for its review was absent, or was lacking a key piece of information? Drill-downs to understand outliers can help agencies understand both production and societal problems.

Repeat Unwanted Events

Agencies also benefit by counting repeat unwanted events, such as repeat violators. A city police department, which has used many of the analytic tools described above with great success over the last decade to drive the crime rate down, recently began focusing on repeat criminals. Numerous environmental and safety agencies track repeat violators, in some cases denying them permits to expand or open new locations.

Diagnostic methods may be as simple as walking the pipes, or they may entail detailed data analysis. Whatever the method, the key is to keep asking why and how until satisfactory answers are found that help an agency encourage the good and discourage the harmful.

Dissemination

Well-disseminated measurement information can promote the uptake of more effective and cost-effective practices, within and beyond an organization. It can also improve program quality through competition, demand shifting, and a better match between consumer and government service providers. In addition, it can enlist external analytic help to identify causal factor and effective treatments.

Promote Better Practices

After problems and solutions are found, a key challenge remains: getting the solutions adopted. Once the marine safety agency office discovered that shifting inspection times to evenings reduced oil spills, it did not just sit on its finding. It shared it with other field offices so they could look for the problem and, if found, test the solution in their communities.

It is often important to gain the attention of people beyond an organization's boundaries and build their understanding of a problem in order to win cooperation to reduce it. The highway safety agency, discussed earlier, packages its data-derived findings to win over audiences whose behavior it seeks to change. It pairs easy-to-understand facts with marketing materials. Its website includes sample letters for law enforcement officials to send to school principals, sample morning announcements for principals to use at their schools, and a sample proclamation for local government leaders.

Competition, Demand Management and Matching Consumer Tastes

Although government agencies are thought of as a monopoly, they can improve service delivery through healthy competition, demand shifting, and better matches to consumer needs. The key is analyzing data and disseminating it to encourage these dynamics. Motor vehicle registration offices are often detested because of their long wait times. One state now posts real-time wait times on its website for all local offices. Visitors can check the wait times and decide which office to visit. Dissemination of additional analysis that showed average wait times by type of transaction, day of week, and hour of day could further shift peak load demand, lessen stress, and instill a bit of healthy competition among offices.

Parents have long made home purchasing decisions based on the best available data they can find about schools. Unfortunately, this data currently reveals more about parental income, highly correlated with test scores, than about the teaching skills of the schools. School performance would be better reflected by reporting scores controlled for key background variables so parents could more clearly distinguish the value added by the school from that linked to the students' parents. Moreover, schools can organize their measurements to reveal which students fare well at which schools, helping parents pick schools matched to their children's learning styles. Parents of slow learners may want a school that does well teaching the lowest quartile of students, while those of average learners may want a school that teaches the middle two quartiles well. Similar analysis could help parents of children who do not speak English well.

External Analytic Assistance

Government agencies lack the resources to conduct all possible useful analysis. Moreover, they do not always understand the relevant questions, and often have political difficulties conducting and sharing analyses that would be interesting. They can, however, organize and disseminate data they collect in ways that invite others to conduct their own audience-tailored analyses.

Many agencies make their data available for others to organize, analyze, and package. Transportation data is used by various levels of government and consultants to support infrastructure planning. When an education agency was mandated to ensure that no student drop below a certain performance threshold, it gathered all student data from every state and normalized each student's score relative to the state median. This enabled a non-profit organization to build its own website that detects high-performing, high-minority, high-poverty schools. This in turn encouraged further research to identify what those schools are doing right.

As agencies build their performance measurement and information systems, data dissemination needs to be seen as a strategic opportunity. Too few governments think about how and when to communicate their findings so that government managers can make more informed decisions that ultimately result in more effective and cost-effective programs.

Dissemination Limits

Dissemination of data and analysis, especially performance information, is a potentially powerful tool that can lead to dramatic improvements in societal outcomes. At the same time, performance measurement can be used in performance-debilitating ways, including punishing those with poor performance. Even when incentives in

the form of punishment or reward are not explicitly linked to individual or organizational performance results, people fear they will be. Agencies should adopt several specific practices to minimize this problem.

First, they should explicitly adopt an accountability system focused on learning and performance improvement, not punishment. Second, individual performance data should not be publicly disclosed. Third, rewards should not be linked to relative performance except for explicit opt-in competitions designed to set new performance records.² Fourth, great effort should be directed to providing performance feedback that is constructive but non-threatening.

² For a discussion of what this sort of accountability system looks like, see Metzenbaum, "Creating Effective Performance Accountability: The Five Building Blocks and Six Essential Practices" forthcoming from the IBM Center for the Business of Government.

Summary

Government agencies need to tap into their data to discover what is working and what is not. Agencies will thrive if they adopt an attitude of inquiry, not embattlement, toward performance measurement, always asking who, what, where, when, and why, rather than starting with the question: *Did I meet my target?*

When agencies make measurements readily available at all levels of the organization, when tools are provided to support widespread analysis by more people across the organization, and when an attitude of interactive inquiry prevails, great things begin to happen. Previously unrecognized problems are seen. Known problems are characterized, and more precise, effective, and cost-effective treatments are found.

Finally, when governments think more intentionally about how best to communicate performance information, they benefit society. Well-designed measurement systems provide relevant information to decision-makers when they need it. They inform policy decisions and consumer choice, and ultimately improve program quality.

Catchy, count-informed, campaigns attract the attention of target audiences, while persuasive, well-packaged performance measurement speeds needed behavioral changes. And accessible, easily queried databases enlist external analytic assistance and expertise, leveraging government resources to improve public value.

About Cognos, an IBM company

Cognos, an IBM company, is the world leader in business intelligence and performance management solutions. It provides world-class enterprise planning and BI software and services to help companies plan, understand and

manage financial and operational performance. Cognos was acquired by IBM in February 2008. For more information, visit <http://www.cognos.com>.

For more Information

Visit the Cognos Web site at www.cognos.com

Request a Call

To request a call or ask a question, go to www.cognos.com/contactme. A Cognos representative will respond to your enquiry within two business days.

GLOBAL

Cognos Incorporated
3755 Riverside Drive
P.O. Box 9707, Station T
Ottawa, Ontario
Canada K1G 4K9

ASIA/PACIFIC

Cognos PTY Limited
Level 2 110 Pacific Highway
St. Leonards, NSW 2065
Australia

EUROPE

Cognos Limited
Westerly Point
Market Street
Bracknell, Berkshire
UK RG12 1QB

NORTH AMERICA

Cognos Corporation
15 Wayside Road
Burlington, MA
USA 01803