Introduction

There is increasing rhetoric and concern today about government’s ability to perform, to actually get things done. Rapidly changing technology, significant economic constraints, and the increasing complexity of public challenges require public institutions to be, more than ever, adaptive and “learning” organizations. It is not surprising then, that performance measurement in the public sector once is again a timely topic.

Performance measurement can be an important tool for not only improving organizational focus and performance, but to provide transparency and accountability to the public. This is certainly not the first time performance measurement has become an important topic. The 1990’s saw the emergence of performance management as the latest organizational innovation, highlighted by then Vice-President Al Gore’s oversight of innovation in the federal government.

Unfortunately, for many, performance measurement became just another passing fad in the 1990’s, and many of the performance measurement systems and innovations were not sustained. The authors believe that it is an appropriate time to look at one of the experiments of that era and see how it may apply to today’s public needs.

Revisiting Oregon DEQ utilization of performance measurement effort begun in 1990’s

In the mid-1990’s, the authors were managers at the Oregon Department of Environmental Quality, overseeing regulatory field staff in a twelve county area in western region of Oregon. Reminiscent of today’s political landscape, there was lagging confidence in government (recall the shut-down of the federal government due to Congress not passing a budget), yet increasing expectations from the public about environmental protection.

The Western Region office of Oregon DEQ asked the managers of each of its six program areas (e.g. Hazardous Waste, Air Quality) to develop measures of success
that would be meaningful to staff as well as tangible to the various stakeholders we served. There was no set process. Each program area was allowed to come up with its own approach and measures. This was purposeful as the goal was to allow the program areas to focus on their highest priority areas, and utilize the flexibility and creativity within each section. In hindsight, this approach provided a unique opportunity to evaluate what factors influenced the success of a given performance measurement system.

The challenge was put this way: “What things would you talk about if you wanted to brag about how effective your program was?” This simple question led to an intensive, iterative process among the staff, their manager, and the Regional Administrator. The managers and staff, accustomed to occasional management initiatives (“the flavor of the month”) duly went through the routine of developing what they thought senior management wanted—in essence, they were searching for the “right rock”. The first iteration of these performance measures ranged from highly unrealistic, pie-in-the-sky goals to setting low bars—easy to jump over, yet not particularly meaningful in the long run. In addition, the list of measures was often long, difficult and resource intensive to track, and not particularly meaningful to staff. Rather than achievable and meaningful goals that reflected strategic priorities, the proposed performance measures were often viewed as “extra” work. The program teams were asked to go back and try again.

By the third iteration, the six sets of performance measures began to meet the “bragging” criteria discussed above: they were meaningful, measurable, and manageable. They provided a venue to describe in lay terms what success looked like. The most notable characteristic of the majority of the final measures was the staff chose priorities. These were clearly tied to key program activities, but staff had a clear had in setting the priorities and measures. The measures did not include every single activity they engaged in. Instead of their initial list of 20 or more measures, most programs settled on six to ten key measures. Not every program approached the exercise with the same level of initial commitment or enthusiasm, but a few examples are illustrative of the approaches and creativity used, and the successes and pitfalls of the various approaches.

**Environmental Cleanup Program**

Since this program responds to addressing contamination uncovered in the environment, the workload can vary depending upon when and how that contamination takes place and the economic drivers for returning property to productive use. Coming up with a measure of success that took this into account was a challenge. For example, the Environmental Cleanup program has a goal of getting responsible parties to clean up hazardous contamination. These cleanup efforts can take years, and a range of factors affect how quickly the cleanup goes, not all of which are in the hands of the staff.

This group approached the challenge by conducting a stakeholder analysis—what
they believed (and in some cases knew) that stakeholders thought about their performance and how could they improve not only the perception, but the actual results. There were several factors the group identified, but one theme that emerged that was of concern to the group was the perception by stakeholders that the cleanup program was “a black hole”—things went in but never came out.

The program staff and manager (one of the authors) determined that during the long cleanup process, many reports and proposals are submitted for approval by DEQ before moving to the next step. How quickly those reports were turned around could determine how quickly the cleanup could proceed. The turnaround time after receipt was something that DEQ had some control over. But these reports varied in size and complexity, so that a “responsive” time for review in one case might not be responsive in another.

Here’s the measure they came up with: “Meet the specified (by staff) turnaround time for review of work products at least 85% of the time”. And, “Renegotiate in advance 100% of the time a deadline that cannot be met”.

This measures couplet served two purposes. First, it opened the communication between the staff and the responsible parties about when the review would be completed. If this deadline didn’t meet the responsible party’s needs, then a dialogue ensued and another mutually agreeable date would be established. Then, if something got in the way of meeting a deadline (e.g., high priority spills or extended medical leave), the responsible party was notified as far in advance as possible about the deadline being in jeopardy and either a new deadline was established or other resources would be diverted to meet the goal.

Here was the result: In the four years following the establishment of this measure, the program exceeded performance (i.e. were above the 85% level) all but one quarter (81%), and were frequently above 90%. This was a great improvement on performance prior to establishing the measure, and the effect was immediate. Moreover, the program graphically tracked performance related to this measure and as one can see (Figure 1) the staff were paying attention to the measure. Without such a measure, we could expect to see a traditional bell curve on review times. With the measure, we experienced a significant spike on the day specified. Stakeholders noticed the difference and commented on the responsiveness of the regulatory team.

Another measure of success identified by the program focused on the substantive goals of the program: volume of soils and groundwater remediated. There was no threshold identified, as the variations in volume from quarter to quarter or even year to year depended upon factors over which the program had no control. However, knowing the number was important to staff, it was a meaningful measure of their work. These were tracked visually as well—imagine an image of a dry lake filling up with “cleaned up and restored” water over time. This measure easily met the cocktail party test, though later was changed to “Number of people and acres of
habitat protected by remedial actions.”

Air Quality Permitting Program

The Air Quality program at DEQ had a history of permit backlogs, i.e. companies would apply for an air quality discharge permit, or a renewal, and – with technological requirement, modeling, and public participation requirements - would often have to wait over a year to get the permit.

The Air Quality program adopted a goal of 100% elimination of the permit backlog, and that 75% of all new permits would be drafted and issued within 6 months of receiving a complete application. There was an incredible amount of energy, creativity and resourcefulness that went into tackling this permit backlog and the results were impressive. The Air program exceeded this goal, and received a Governor’s Award for excellence in 1995.

This is an example of a measure that yielded outstanding short-term results and recognition. However, this performance measure wasn’t fully integrated with the long-term strategic priorities and processes of the program, so that a permit processing backlog (albeit to a lesser degree) later re-emerged. To be both more sustainable and adaptable to changing conditions, the measure might have been re-cast to better integrate permit process and environmental priorities, keeping a focus on maintaining a minimal backlog.

The Air Program also adopted a goal of responding to citizen complaints within five working days of receiving them, which they met. Once again, this was a major improvement over previous performance. Dealing with environmental complaints was an activity for all programs in the region, though some did not establish targets, some had more aggressive targets (three days) and others had more lenient targets. So, although this target was met, the issue may have been more effectively handled in a larger, organizational and strategic context. Indeed, the Western Region programs evolved to having a common standard for complaint response. [As a side note, Oregon DEQ is currently evolved in numerous process improvement events (LEAN/Kaizen) that are comprised of staff across the agency and outside participant. One event focused on development of a uniform process for responding to environmental complaints from the public and the agency is currently standardizing the entire complaint process statewide.]

Underground Storage Tank Program

DEQ’s program to address leaking underground storage tanks utilized two goals that were already mandated by the U.S Environmental Protection Agency (EPA) for programs receiving federal funding: Initiate corrective action at 80% of tanks with petroleum leaks in a year, and completing the cleanup at 30% of the sites.

The quarterly data shows that once this performance measure was introduced and
adopted by staff in the Western Region office, corrective action was initiated at 85% of the cases within the allotted time every single quarter. Cleanup was completed in only 8% of cases in the first quarter, but that performance improved each quarter until by the 4th quarter nearly 40% of the cleanups had been completed. (See figure 2)

Hazardous Waste Program

In Oregon, as in most states, hazardous waste regulation has traditionally followed federal EPA guidance and protocols when regulating the storage and transport of hazardous waste. These federal Resource Conservation and Recovery Act regulations are prescriptive, painstaking detailed and the required inspections of industrial sites are highly labor intensive. Consequently, EPA required that Oregon’s Western Region only complete 22 inspections per year for the more than 200 facilities.

The development of performance measures for the program led the manager and staff to focus more on the overall goals of the program, which were to improve the management of hazardous waste and, ultimately to get people to reduce their hazardous waste. This ultimately led to a Generator Assistance Program, in which the state regulatory staff worked together to provide one-time-only technical assistance to companies in the region.

Here’s what happened: Over a six-month period in 1995, 139 facilities voluntarily received a much-less labor intensive technical assistance visit, that reviewed the hazardous waste management practices and then provided detailed recommendations for needed improvements. Follow up interviews and visits revealed initially that 65% made significant improvements to their waste management practices, and 52% made progress toward reducing their hazardous waste. In subsequent years, the program made improvements to the process to increase those figures to 80% and 60%, respectively.

The results—not only did the section meet its inspection targets, it proactively touched seven times the number of facilities that would have normally been visited by DEQ for a RCRA inspection. And, as the results demonstrated, hazardous waste management improved at the majority of these facilities.

What we learned

This foray into performance measures in the 1990s yielded some valuable lessons that the authors have since put to use in other organizational settings.

1. Invest the time up front. The time required to develop effective performance measures is critical and well worth the investment. Having employees take the time to wrestle with fundamental questions about how
success is defined helps align organizational effort for much greater focus and
productivity. While it may be tempting to short change this part of the process, a performance management system is only as good as the measures it is built upon.

In addition, we found that performance measures developed by staff, and systematically adjusted and fine-tuned in response to business needs and data availability, were more likely to have “staying power”, i.e. the level of understanding and commitment to the measures helped ensure a workable plan of implementation and thus the measurement and the performance itself were more sustainable.

2. Integrate performance measures with strategic priorities. Performance measures can be an incredibly effective tool for focusing and aligning resources. The important thing is to make sure they are focused on the right things—not what is too easy (becomes a check the box exercise), too expansive (they aren’t achievable) or too many (they lack focus). Although some measures clearly need to remain in place for trending purposes, be flexible and open to changing measures to meet the changing demands of staff, stakeholder and program needs.

A list of twenty-five performance measures is unlikely to change or improve organizational performance. In order to be effective as a sustainable management tool, the performance measure results must be clear, focused (usually no more than ten) and have at least tactical, if not strategic, meaning if they are to be viable.

Collecting data that has no demonstrable, tangible purpose tied to organizational goals can actually be disenfranchising to staff. Some of the failed measures had this as an inherent issue. Although the data was interesting (responding to complaints in five days) and led to subsequent process improvement, it lacked organizational context and meaning and did lead to some competition among sections (see below).

3. Bottom-up versus top-down development. It is imperative that staff and middle management need to truly buy-in to performance measurement. If this is done “to” people, there is a lack of ownership and cynicism will flourish. Both staff and managers need to recognize the value of the performance measurement process for it to “stick” and become part of the culture. The measures need to be built into individual performance goals, then section goals and ultimately strongly linked to organizational success.

In our early efforts in DEQ’s Western Region, we put targets into employee’s written work agreements. Although many of the elements of the early performance measurement work have not been maintained, several of these work agreement targets are still used today—they had the needed “staying power”. Another way to signal the importance of the performance management system is to visually display the results and own (and leverage) what is working well while using the areas of improvement as learning opportunities. The purpose of publishing results is not to shame individuals, but rather to acknowledge accomplishments and focus attention (in a positive, problem-solving way) on areas that need improvement, using these to grow your organization.
Hand in hand with integrating performance measures into organizational management structures, however, is ensuring that the emphasis is not on using performance measures as a punitive tool.

To do so ensures that “low bar” commitments will be made and there will be little buy-in to the overall process. A few managers reported that they did not set aggressive goals for fear of failure and some staff were reticent to have performance targets put in their work agreements for the same reason. Setting low bars to ensure personal success does little to improve organizational success.

This is not to say that underperforming (i.e. not meeting performance targets) by an organizational unit should not be addressed. On the contrary, performance measures offer a great opportunity for organizational problem-solving. Organizational cultures that embrace accountability do not look first towards the “people problems” to determine what is not getting done, but instead look towards the “process problems”. We definitely saw examples of the process problem approach (e.g., Hazardous Waste Program) and the results were truly transformational. What are the reasons we are not seeing the desired results? Is there a lack of training and tools to be successful? Was attention diverted to other activities for valid business reasons and targets not adjusted? Is the data truly measureable? Are we measuring the right thing? These are examples of effective process questions to ask if the desired results are not being met.

4. Managing intra-organizational competition. A natural outcome of measuring team performance is that teams will tend to compare their results with other teams. In our case, we tended to celebrate extraordinary team accomplishments, with the hope of inspiring the other teams. A little competition can be a good thing. However, one of the unhealthy aspects of the DEQ experience was the sometimes unhealthy competition that arose between some work units creating a sense of “winners” and “losers”. This signaled the need for groups to understand that their team is really the entire organization. And, it is the responsibility of each team member to make the organization as strong as possible. It is part of leadership’s responsibility to set the tone in this manner and quickly curtail any competitiveness that crops up.

Conclusion

The old adage of, what get’s measured get’s done, is alive and well! We learned that getting performance measures in place is a first step—getting the right performance measures in place is the correct step. Using existing priorities (the organization’s tactical and strategic objectives), external inputs (e.g., Legislative and Federal commitments; stakeholder requirements) existing data (e.g., permit backlogs) and staff knowledge are vital in identifying and delivering on the right programmatic measures. In summary, evaluating our efforts we found the most important keys to
success were that the performance measures needed to be (a) meaningful, (b) measureable, and (c) manageable.

In terms of meaningful, the performance measures need to have relevance to the staff, stakeholders and be clearly linked to organizational tactics and strategies. Without this, the measures became “extra” work versus integrated into the way in which an organization looks at continuous improvement and accountability, both within and external to the organization.

The next key was to ensure that the performance measure was actually measureable. We determined that there was less angst in trying to choose between an output measure (reducing permit backlogs) or output measure (improving hazardous waste management practices) than in trying to decide what and how to collect data. Much of the data programs elected to track had to be managed manually. This becomes unsustainable. Finally, setting the threshold for success is critical. We termed this the “Three Bears” approach was the most fruitful—not too high, not too low, but just right—in terms of the measures that were selected.

Finally, the measures need to be manageable. This is a case where less is more. The ability to focus on a few key measures is far more important than tracking numerous measures. This only serves to dilute effort and focus and may lead to a disproportionate amount of time being spend in tracking, measuring and reporting on a measure versus doing the work itself.

Delivering on commitments in a visible and transparent way not only helps an organization focus on the right things, but also is a big step forward in restoring confidence in government. For an organization to be proactively report on achievements demonstrates fiscal accountability and organizational integrity. An organization that can effectively deliver on performance objectives is one that is positioning itself for success now and into the future.

References:

The Oz Principle: Getting Results Through Individual and Organizational Accountability by Craig Hickman, Tom Smith and Roger Connors May, 2010

Change the Culture, Change the Game: The Breakthrough Strategy for Energizing Your Organization and Creating Accountability for Results by Roger Connors and Tom Smith, Jan, 2011.


Balanced Scorecard Institute


The Five Commitments of a Leader by Mark Leheney, 2008.