

## PERFORMANCE MEASUREMENT

---

Steven J. Wilcox  
New York State Department of Transportation

### ABSTRACT

Transportation maintenance organizations have collected data for some time, but rarely has that data been linked to the strategic interests, or management information needs of the leadership or management. Frequently that is due to the lack of clarity of what those interests are, making it impossible to align the organization to satisfy them. Performance measurement in New York State's Transportation Maintenance Division is an attempt to create a comprehensive and strategic framework for management. The framework begins to clarify the organization's mission, services that support that mission, and the characteristics of the service delivery process necessary for success.

Particularly important for agencies interested in exploring the idea of organizing around public service is to begin to define service expectations from the perspective of customers. By linking budgets and the quality assurance systems to customers, service priorities and levels can begin to be established cooperatively. Only then can meaningful performance measures be developed for the organization's various roles and responsibilities, that separate success from failure, from the standpoint of the people being served.

### WHY MEASURE PERFORMANCE?

The process of organizing to measure the performance of an entire agency or some subdivision at that agency is arguably the most important element in results-oriented management. This involves examining why the organization exists, clarifying and reaffirming its purpose, and determining whether it is objectively meeting that mission as simply and elegantly as possible.

Implicit in defining an organization's purpose is understanding who it serves. Consideration must also be given to those characteristics processed by the organization and its services in satisfying the need of the user of these services, and in maintaining a healthy, effective organization.

Armed with the knowledge of why and for whom the organization exists, leaders can then determine exactly what services it provides that fulfill its mission, and then

the processes for delivering those services. Assessing the service delivery processes, and any other processes providing direct support, will distinguish value-creating parts of the organization from others that no longer directly or indirectly serve the organization's purposes.

At this point questions of why, what, how and for whom have been answered, and that knowledge can be used in the development of strategy, priorities and goals. Once it is clear how the organization fits together and what the roles and responsibilities are for its strategic, tactical and operational levels, one can begin to measure the organization's capacity, behavior and health. Having an organization that is well-aligned in support of its mission, and where expectations and roles are clear, performance measurements can be set up to find how well the mission is being accomplished, to show that it continues to stretch its capacity, and that through its people, the vision of demonstrably better public service can be realized.

### BACKGROUND

In 1993, the Transportation Maintenance Division of the New York State Department of Transportation began an effort to improve management effectiveness. A literature search suggested several forces driving organizational improvement efforts - among these are total quality management, reengineering, reinventing government, ISO 9000, and performance management.

These performance improvement approaches were studied to learn which aspects could benefit a public-sector, service-based organization like the Transportation Maintenance Program. Relevant concepts began to be sewn together and tailored in a manner that appeared most sensible for the program's circumstances. The ultimate goal that we refer to as *Performance Management* was to create a mission-driven, comprehensive, and well-aligned of management system.

As managers of the Department and Maintenance Program changed, so did the scope, priorities, and support for managing its performance. Though these changes have created "roller-coaster" effects and stalled application of the system, they have also provided the opportunities to clarify important precepts -- such as consistent leadership

support, and clarification of roles within the management hierarchy -- necessary for the successful implementation of an approach to management that significantly departs from the more comfortable circumstances engendered in the traditional command and control bureaucracy.

The original implementation plan consisted of a top-down statewide approach. The intent of this was to align the organization with a consistent statewide mission and strategy, and to gain management support before application to use at the next level down into the organization. As leadership changed, so did direction. A pilot based in the Hornell region in southwestern New York was chosen, based at the operational (working) level of the organization. The benefit of this change has been to work through several iterations of performance measures to obtain a set that operational managers find useful and relevant, and to develop graphical methods for expressing them. The weakness of this approach has been that without agreed-upon purpose, strategy, tactics, or key result areas supported by the leadership there is only an assumed direction to align the system to. Managers also are not being held accountable for measuring their performance causing it to be developed and used on a time-available basis. This results in performance measures being "nice to know," but not "necessary to use." New leaders are currently evaluating a return to the original approach.

### A COMPREHENSIVE APPROACH TO PERFORMANCE MANAGEMENT

From the late 1980's, government agencies across the country, at every level, have been reevaluating how they deliver service to the public. Many government organizations across the country have produced radical improvements in the efficiency and effectiveness of their service delivery. Our intent has been to benchmark those agencies in various aspects of organizational transformation, and sew them together into a purposeful, comprehensive, and complementary system of management. The pieces viewed as critical to a comprehensive performance management system include: performance measurement, quality assurance, a service delivery system, and performance budgeting. These pieces needed much integration and overlap, as well as, the development of supporting pieces, such as a clear mission; service definitions; customer-oriented outcomes; key result areas; process management; clarification of strategic, tactical, and operational roles; and balanced, continuous improvement.

The models that we attempted to integrate included: Oregon DOT's Performance Measures [1], concepts of variation and continuous improvement from Total

Quality Management [2], Texas State Government's Strategic Budgeting System [3], ISO 9000 Quality Assurance System Standards [4], and the key result areas from the Balanced Scorecard described by Robert S. Kaplan and David P. Norton in a series of Harvard Business Review articles.[5]

### ORGANIZING FOR MEASUREMENT: PURPOSE, PROCESS AND PEOPLE

To measure performance anywhere in the organization a framework must be established clarifying 1) what results the organization is aligning to achieve, 2) why those results and with what priority, 3) how the various parts of the organization work together to achieve those results, and 4) the roles of the various levels of management in steering the organization to those ends.

#### Purpose

In the course of maintaining highways the organization does a lot of "stuff" -- pavement stuff, bridge stuff, roadside stuff, and snow and ice stuff. Establishing a mission, the answer to the question of *why* the organization exists, replaces "stuff" with purpose. The Transportation Maintenance Program has established four fundamental reasons why we do stuff: *Mobility, Appearance, Preservation or Safety* ("MAPS" for short, making it easy to remember). Following definition of the four fundamental purposes of the organization it was possible to survey management, asking for each of the roughly 100 direct tasks performed by maintenance workers, which of the four purposes was being met by doing that task.

The survey revealed eleven natural groups of activities defining the services that the program delivers to meet our mission. The services are defined by the feature being maintained, and what maintenance purpose. These services are: Winter Mobility, Mobility Restoration, Traffic Guidance and Control, Roadside Appearance, Rest Area Appearance, Pavement Preservation, Bridge Preservation, Pavement Safety, Bridge Safety, Roadside Safety, and Safety Appurtenances.

Outcome expectations can then be set and budgeted for these services. For example, Roadside Appearance is a service provided primarily to maintain an attractive roadside. The maintenance program performs four tasks to that end: mowing, litter collection, landscape maintenance, and removing dead animals, debris and encroachments. Statewide task history from our management information system details the historic costs and effort put into this service. If program managers want

to negotiate increased funds for a higher service level, a set of expected outcomes can be defined in terms of grass, trash, and carcasses. Administrative and overhead costs are included in the service cost making the cost a bottom-line price. The alignment with the mission also becomes clear. The purpose of litter-collection, for example, is improved appearance of the roadside, satisfying the overall mission of managing the highway's appearance.

### Process

This answers the question of how the organization is put together, so that the desired results are produced as efficiently and effectively as possible. It orients the organization toward work flow and away from the static structural hierarchy.

Process flowcharting clarifies the steps necessary to successfully deliver a service. Receiving a service with the desired characteristics is all that the public cares about (characteristics include quality, cost, timeliness, etc.). Poor process design can generate inefficiency, high cost, and poor quality which are why so many organizations are undergoing process reengineering.

Process flowcharting also can serve to separate line functions from support functions. An important point to recognize here is that support functions are only as successful as their ability to supply the line function—they must be measured against those ends. For example, the Equipment Management function supports the Transportation Maintenance function and must be measured as if they supplied road salt, consultant service, or photocopier repair. All too often, support functions are measured as ends in themselves. Using the Equipment Management example, they could, through the type of equipment that is supplied, result in the line's work being determined by the support group, and not by the strategic service priorities. So rather than measure cost per vehicle maintained, it is more sensible to measure, for example, equipment cost per acre mowed demonstrating their impact on mission fulfillment. They would be fulfilling their unique purpose of maintaining equipment, but without a link to the line's mission, there is no guarantee that the equipment necessary is available to the line to meet expected service levels.

It is important to remember that the practice of continuous improvement that is fundamental to Total Quality Management (TQM) assumes that the processes being improved are already well designed. Further, strong bureaucratic structures that have been common in the public sector were designed to maintain the status quo - in an effort to control poor performance, they also discourage doing anything differently even if it is an

improvement. Hierarchical approval systems discourage innovation and improvement by requiring approvals through lengthy and document laden justification processes. This is why many organizations are leaving the bureaucratic machine model with its roots in Frederick Taylor's scientific management model and are moving to a Japanese style values-based normative-control model.

### People

For persons or groups to be effective, their roles and expectations must be absolutely clear. Roles are strategic, tactical, or operational. All three are critical, but all three are separate. Roles can become confused, particularly in organizations with poorly conceived decentralization or empowerment.

The strategic role involves those in the organization in a position to set a course. Typically, this is the senior executive management team. It is crucial that those with strategic responsibilities be connected to the market being served by the organization. For a public transportation maintenance organization, that market consists of 1) system users, 2) transportation interest groups, 3) politicians, and 4) agencies whose interests involve transportation. The primary responsibility of strategic management is to clearly understand where the organization needs to go, and to develop strategies, set priorities, and establish policies consistent with that direction.

The tactical role translates strategy into a form that can be carried out. This requires ability to assess organizational capacity to achieve strategic outcomes, plan and deploy resources, and perform the quality assurance function to determine the effectiveness of various tactics. Typically, this includes central office program managers and regional functional group directors.

The operational role delivers the work that fulfills the organization's mission in terms of the priorities of the organization. Those priorities and service quality requirements must be clear and supported to do this in a manner that aligns with the strategic priorities of the organization.

## CUSTOMER REQUIREMENTS, QUALITY ASSURANCE AND MISSION-BASED BUDGETING

### Customer Requirements

Public service is the essence of what government agencies provide, but only exceptional agencies are organized from a service perspective, and very few have an institutional

connection with the public. The greatest obstacles to overcome in reinventing the culture of public agencies are 1) serving the budget, 2) serving the structure, and 3) serving the boss.

For a true public-service orientation, organizations must be clear what services they provide to the public to fulfill their missions, ask the public what they want from those services, and also ask how much they are willing to pay. With that knowledge, agency leaders can then look into their organizations to deliver those services.

The widening division between government and the people is a consequence of government's lack of dialogue with those they serve. It is arrogant for a program administrators to assume they know what their customers want better than the customers themselves. Inevitably, tactical program managers get caught up in providing elegant, high-tech solutions to problems (and non-problems) because frequently the only formal dialog with the external citizenry is with the consultants or vendors.

Customer focus groups, customer councils, perception surveys, and town meetings are options that leaders have available to understand expectations of their constituencies. It is difficult to engage the public if the organization has not clarified what services it provides to fulfill its mission.

To understand services, it is important to know more than what is being serviced (pavement, bridge, roadside) but also why and to what ends, that services are aimed (mobility, preservation, safety, appearance). If you ask someone what is expected from the roadside he might ask, "The roadside *what?*" If you ask about roadside *appearance* or roadside *safety* you then have the basis for an exchange. Services are linked directly to tasks, and as a starting point the organization's task history can be used to learn what the average annual direct and bottom line costs, and labor commitment have been required for each service. Roadside appearance consists of mowing, litter pick-up, removal of dead animal/debris, and landscaping maintenance. Roadside safety consists of tree and brush removal, slope maintenance, and chemical weed control (of vegetation growing along guide rail). (It is unmanageable to define results for 100 different maintenance tasks and the permutations of why each is done.)

Defining broad categories reduces these to eleven services that have a mission-based organizing principles and results. This also provides overall outcome expectations for front-line managers and supervisors without tying their hands with individual task requirements. Roadside-appearance outcomes can be tailored at the local level to local conditions. Perhaps a rural area needs to place increased emphasis on mowing and dead-animal removal, while an urban area would focus

more on litter pick-up and maintaining landscaping plantings. Both approaches are aimed at a pleasing roadside appearance.

Tailoring these approaches to public expectations completes the picture because a clear (rather than a guessed-at), organizational expectation is being met. Measuring organizational performance in delivering on public expectations fulfills the other half of the equation of what (paraphrasing the philosopher Robert Nozick) creates a sense of importance in human beings - doing something that counts, and having someone appreciate (by accounting for the fact) that it was done. Civil servants must know that their work contributes positively to society (by fulfilling a mission), and, favorably or unfavorably, that it is noticed (by being measured). People would rather know that they are doing a poor job than be ignored, unless there is no perceived relevance to their work.

This accounting through performance measures must consider both the level of performance relative to the organization's capacity, and the quality of service delivery. Capacity issues tend toward measures of efficient uses of fiscal and human resources, while quality issues are oriented toward the customer's satisfaction with service, developing internal human resource quality, and assuring the technical quality of the service being delivered.

### Quality Assurance

The international standard for quality assurance systems is the International Standards Organization's 9000 Series (routinely referred to as ISO 9000). The quality assurance system is under development and uses the *framework* established in the 12-step 9003 Standards -a model for quality assurance in final inspection and test.

ISO 9003 was chosen over 9001 and 9002, which manage quality throughout the production and installation process (9001 begins as early as design and development), based on the nature of maintenance activities being well defined in their design, and the complexities of managing quality continuously for an array of 100 discrete maintenance activities. As the organization becomes more results-oriented, statistical sampling of the quality of services that have been delivered will require operational managers to continuously control work quality.

The precepts established for the quality assurance system are as follows: use the framework of ISO 9003 (that the aim is not certification, but consideration of all elements necessary for a comprehensive quality system), develop customer-oriented standards, organize around the eleven mission-based services, measure (quantitatively) the quality of services being delivered and develop a statistical

sampling plan. The organization is currently at the point of determining the mechanisms to define customer expectations for the eleven services. When completed, the quantitative results of the quality-assurance reviews will be included as performance measures.

### Mission-Based Budgeting

To create a holistic management approach of complementary components, budgeting is a key consideration. Public organizations throughout the country have developed clever performance-measurement systems, quality systems, or budgeting systems. A belief underlying the approach in New York is that all of these components must be thought through in designing and developing a total system, though they do not need to be (and should not be) brought on line at the same time.

The premise that the budgeting system works from its ability to define what services the organization has to offer and how much of each service a dollar will buy. The control element holding the agency accountable then must be a results-oriented quality assurance review. For budgeting purposes, this review should be performed outside of the organization being evaluated.

Developing a service-based budget should also simplify the decisions as to what services to contract out, performed with state forces, or privatized. Using tasks as a service basis also lays the groundwork for activity-based costing for value comparisons necessary if service-provider decisions are to be based primarily on economic considerations.

### BALANCING A FAMILY OF MEASURES: KEY RESULT AREAS

"No measure stands alone" is becoming a familiar phrase in the vocabulary of performance measurement. Management is no longer aimed solely at increased productivity or lower costs. High productivity or low cost at the expense of quality, customer satisfaction, innovation, learning, and continuous improvement can be damaging to an organization's market position in a world economy no longer defined by the industrial production of commodities.

The characteristics describing excellent service delivery and sound organizational health shape the categories of performance that require measurement. Efficiency and effectiveness are the simplest means of classifying measurements and are frequently heard in political rhetoric.

Searching for the categories to paint a comprehensive picture of the organization's performance has led to an approach discussed by Robert S. Kaplan and David P. Norton in a series of Harvard Business Review articles detailing what they term "The Balanced Scorecard."<sup>[6]</sup>

### THE BALANCED SCORECARD AND KEY RESULT AREAS

The critical premise underlying the balanced scorecard is that organizational success is a consequence of manager attention being devoted to the entire family of factors that are critical for high-quality, profitable service delivery. The factors that must be both balanced and continuously improved upon are generally referred to as *Key Result Areas*. Though more than a dozen such areas have been identified, the balanced scorecard focuses on four that are critical for understanding and diagnosing the total organization. The four key areas consist of views of the organization from the following perspectives: 1) *financial*, 2) *internal business*, 3) *customer*, and 4) *innovation & learning*.

The Financial Perspective asks "how do we look to shareholders?" In the public sector that translates to taxpayers and the control agencies. Frequently this is the only perspective anyone looks at, generally from a gross input standpoint, and rarely from a value creation or performance basis. In developing financial measures for the Transportation Maintenance Program, managers are asked to develop measures that they could use to find whether their part of the organization is making best possible use of fiscal resources entrusted to them in delivering services. At the operational level, these measures have focused on appropriate use of overtime and on controlling support costs.

The Internal Business Perspective asks "how must we excel?" From a maintenance operations perspective, this involves efficient use of labor, equipment, and materials; providing quality service results; minimizing rework; and rapid response. At the operational level, measures of labor efficiency have been developed and a quality-assurance system is under development. Though there is agreement that rework and response time are important also, meaningful and useful measures have thus far proven to be elusive.

The Customer Perspective asks "how do our customers see us?" What does the public think of services delivered by the Maintenance Program? Commonly there are two aspects to this question. The first is the priority of service *importance*- how do our priorities align with those of the traveling public? The second concerns organizational *performance* in delivering services, how

good a job do we do? One such survey was linked to individual counties as performance measures. The goal was to define this practice through a combination of winter and summer seasonal perception surveys.

The Innovation and Learning Perspective asks "can we continue to improve and create value?" This is the perspective that is generally the most difficult to define in the public sector due to the bureaucratic systems controlling most agencies. Bureaucracy operates on a basic principle that is at odds with *kaizen* (continuous improvement). An aim of bureaucratic systems is rocklike stability (rational division of work by function, authority and structure to persons who work in a controlled manner according to strict standards, rules and procedures)- hardly a system where innovation can flourish. Contrast this with the philosophy of total quality management advocated in Deming's fifth point for management: "improve constantly and forever every process for planning, production, and service," or his sixth point: "institute training on the job," or his eighth: "drive out fear," or his twelfth: "remove barriers that rob people of pride of workmanship . . .," or his thirteenth: "institute a vigorous program of education and self-improvement for everyone." [7]

To highlight these values innovation and learning are included as the fourth area for performance measurement. At the operational level, innovation itself has been difficult to measure directly. To develop a bureaucratic paper trail effectively stifles (formal) innovation. Rather than attempts have begun to measure innovation directly, videotaping practices and creating teams of peers to review those practices and share ideas, and have been received enthusiastically. The underlying belief is that innovation is best encouraged through peer exchange and by creating an atmosphere of competition for improvement.

Measures for learning and employee perceptions of the performance of an array of management areas have been developed for the operational level. These examine training hours as a percentage of total time available. The second measure looks at how purposefully the organization is managed through a survey of employee's perceptions of six key management practices: performance expectations, effective authority, teamwork, performance evaluation, rewards and recognition, and responsibility for results. The survey is completed by the manager and those supervised by the manager to determine both the subordinates absolute level of the perception of performance, and the relative differences in perception between how the manager and the manager's direct reports. Surveys by level and location have provided interesting results indicating where communication breaks down, the differences in perception of how the manager's view themselves and how others perceive them, and the

differences between problems created by the manager as opposed to those created by the bureaucracy.

## BUSINESS CYCLES, TRENDS AND THE OREGON MATRIX

One of the first lessons learned in developing performance measures was that transportation maintenance in New York functions on two distinct business cycles - summer and winter. This became clear through the iterative process of developing measures. Measures of overtime, labor efficiency, quality, and customer satisfaction were difficult to set goals for because of the forces that drive the weather-dependent winter season, as opposed to the much more planable summer season. The principles in organizing a winter operation also are much different from summer, as is the task mix. By separating the two major business cycles, measures much more closely model the operating conditions facing managers. Consistency of data models also improves significantly. The next step was to format the measures, reflect the balance of the key result areas, and provide a foundation for continuous improvement.

### The Oregon Objectives Matrix

A common complaint voiced by managers concerning measurement is that every organization's operating environment is unique, so that comparisons between counties or regions are not legitimate. One key principle that resulted is that a county residency, or a region is not different from itself. A measurement system must provide a foundation for continuous improvement by measuring current against past performance for the individual organizational unit being measured. The format that allowed for this (as well as for balancing key result areas) is the Oregon Objectives Matrix.

This was conceived by James Riggs and Glenn Felix at Oregon State University in the 1980's [8], and was adopted by Oregon DOT as part of their ground-breaking performance measurement effort.[9] Each performance measure is identified at the top of the matrix, and level achieved for the measurement period is listed along the side as an index ranging from a low of -5 to the measure's goal level of +10, with zero being the historical average performance level for each individual measure. Each measure can then be weighted, the index level achieved can be multiplied by the weight, and the results can be summed to provide an overall performance index ranging from -500 to 1000. Figure 1 shows the application of the matrix and graphically displays the results for the first 16

### WINTER PERFORMANCE MATRIX

Wyoming County  
SFY: 1995 - 96 PAY PERIOD: 15 - 23

Level of Performance	Fiscal Performance			Internal Process		Customer Satisfaction		Innovation/Growth	
	PP 16-2 % Overtime Hours of Total Regular Hours	Efficiency Measures % Direct Overtime Hours of Total Overtime Hours	Measures % Support Cost of Total Costs	Labor Efficiency	Quality Assurance Survey Result	Effectiveness Public Perception Survey Result	Measures % Training Hrs of Total Hrs	Employee Survey Result	
Goal = 10	5.00	100.00	20.00	150.00%	100	3.5	7.00	144	
9	5.85	98.31	22.38	142.86%	96	4.5	6.53	136	
8	6.70	96.62	24.76	135.71%	91	5.5	6.06	129	
7	7.55	94.93	27.14	128.57%	87	6.5	5.59	121	
6	8.40	93.23	29.53	121.43%	83	7.5	5.12	113	
5	9.25	91.54	31.91	114.29%	79	8.5	4.65	105	
4	10.10	89.85	34.29	107.14%	74	9.5	4.18	98	
Baseline = 3	10.95	88.16	36.67	100.00%	70	10.5	3.71	90	
2	17.98	84.08	44.08	87.50%	60	12	2.86	75	
1	25.00	80.00	50.00	75.00%	50	13.5	2.00	60	
Score for Period	18.90%	93.30	30.03%	132.23%		7.6	0.9	82	
Level Achieved	1	5	5	7		5	1	2	
Weight	10.00%	10.00%	5.00%	20.00%	20.00%	15.00%	10.00%	10.00%	
Total	0.1	0.5	0.25	1.4	0	0.75	0.1	0.2	

<b>Performance Index</b>
3.3

## 1995-96 WINTER PERFORMANCE Wyoming County PP 15 - 23

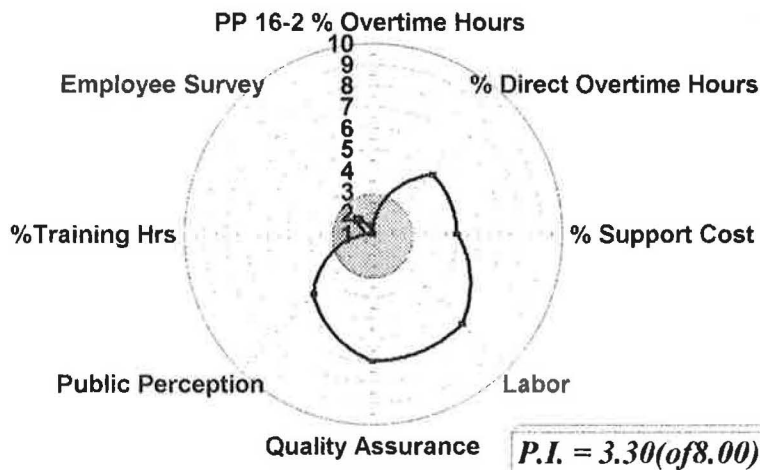


FIGURE 1 Application of the Oregon objectives matrix for a residency.

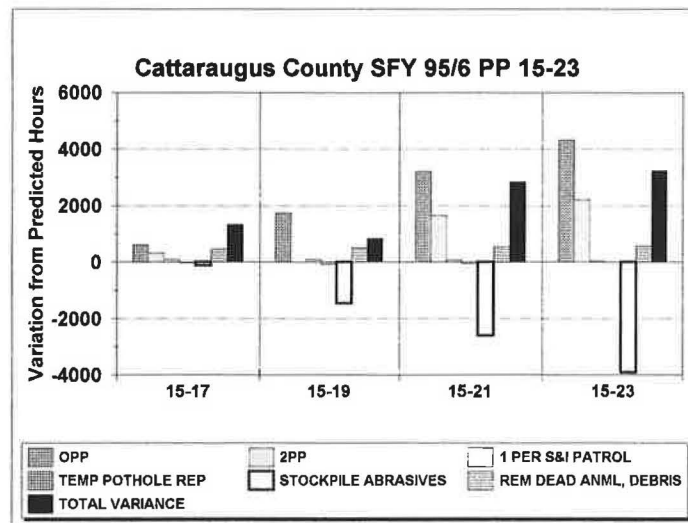


FIGURE 2 Trends in variation for key winter tasks.

weeks of the 1995-96 winter season for a county in western New York.

### Trends

Information is useful only in a form that can create a contextual meaning. Snapshots of data are dangerous in the wrong hands, since there is no context to make them meaningful. Information from a single measurement cycle (one matrix) provides only a point on a piece of paper. What is important to look for are measurement trends, typically of at least three periods tending in specific direction (upward or downward). These could be trends in overall performance based on an overall performance index, or trends for individual measures. Figure 2 shows trends in year to date task variation for key winter activities as compared to what would be predicted if they worked each task at their own 3-year average productivity rate. Understanding the variations that occur in performance of the organization provides a key for managers using measurement as a diagnostic control mechanism.

### VARIANCE, PARETO ANALYSIS, AND THE RELATIVITY OF MEASUREMENT

Consistency and accuracy are the two critical elements of a reliable database. While accuracy is very important for understanding such absolutes as volume or cost, consistency also is important from the standpoint of relative improvement. Getting everyone in an organization to measure an acre of mowing to the same area can be a difficult task; much can be gained by the organization measuring consistently, even if units being measured vary significantly. If one unit's actual measurement of an acre is consistently 20,000 square feet and another's is 40,000 square feet relative improvement or failure can be ascertained, even though accuracy of the total combined acreage is a meaningless number. To evaluate variation in performance trends, it is critical to maintain a consistent database, and accuracy is much less important. The implication for maintaining the integrity of a database is that task definitions, task scope, and units of measurement for individual tasks need to be fixed. Every change or refinement will undermine the integrity of a measure, until sufficient history and experience with the task is established to recreate a context.

Consistent measurement over time will allow for an understanding of the variance in performance of the area being modeled by the measure. This is the same principle

found in measures of the stock market like the Standard and Poor's 500 Index. This concept in conjunction with Pareto analysis is particularly useful for measuring something like labor efficiency.

Creation of a graph to track trends in task variance begins with a Pareto analysis of the previous three years of labor hours. This involves creating a list of tasks by hours consumed from highest to lowest, and including only those that add up to 80% of the total hours for the three years. For counties involved in our pilot, this amounts to 14 to 20 tasks in the summer and fewer in the winter. This immediately reduces the roughly 100 "trivial many" tasks to the "vital few" involving less than 20 tasks. Tasks beyond those making the Pareto cut typically average less than 2% of total hours, so a focus on these would provide very little impact on aggregate labor efficiency. (The same analysis could be performed from a fiscal standpoint by substituting costs for hours to determine which tasks drive financial performance.)

For each task passing the Pareto cut, three-year average hours is divided by three-year average accomplishments to determine hours per accomplishment over the three-year period. Then current year-to-date accomplishments are multiplied by the three-year hours/accomplishment to predict the labor hours required if the task was performed at the same rate. Subtracting actual year-to-date hours from predicted hours then generates the *variance* from predicted hours. This is done for every Pareto task, variances summed to provide total variance. This is similar to profit or loss, except that the currency is hours. Total predicted hours divided by actual hours, expressed as a percentage, becomes the measure of labor efficiency.

Task variances graphed for each measurement period then demonstrate performance trends. Another useful feature of this graph is that hours gained or lost are directly related to the number of hours used, so that improvement efforts can be focused on activities being performed most often where productivity improvements will create the most hours for performing additional work.

This concept has been extended for use at the regional level to create a different context for trends in task variance. (In New York, eleven DOT regions are subdivided into county maintenance residencies.) A Pareto analysis is performed at the regional level from the aggregate data of all county residencies in a region and a regional three-year average rate of hours per accomplishment is established for those tasks. Then variance for each task passing the Pareto screen is calculated, based both on the region's hours per accomplishment and the individual residencies. The result demonstrates whether a residency is improving or failing,



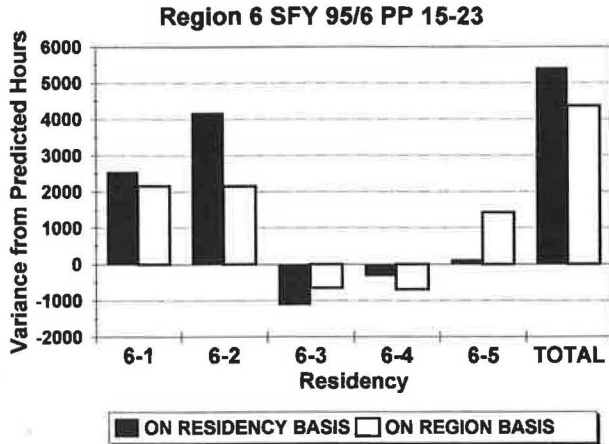


FIGURE 3 Total variation in task performance for residencies in the Hornell region for the 1995-96 winter season.

compared to their own past performance and against average performance within the region. For example, a residency's performance for a task could improve against its own past performance by a significant percentage, and still be significantly below average performance in that region. For the region as a whole to improve, low performers in key regional tasks must be evaluated to attempt to bring them up to the region's rate. Figure 3 demonstrates which residencies are improving or losing overall productivity compared to their own 3-year average rate (on residency basis), and compared to the average rate for the region (on region basis). An important assumption here is that the region is homogeneous.

To test this hypothesis, the coefficient of determination is calculated comparing a residency's performance to the region's. If the coefficient is close to 1.0 it can be assumed that in general the hypothesis is true. The same calculation can be performed for a residency, comparing its current performance to its three-year average to determine whether the data are behaving consistently. If not, the same variance chart can be used to address the data in detail by pointing out tasks that are highly variable. An analysis of the residencies in Figure 3 had coefficients greater than 0.90 when compared to the regional average showing that the residencies behave very much like the region as a whole with the exception of Schuyler County. Further analysis demonstrated that though productivity did not relate strongly to the region it did not relate strongly to its own history indicating inconsistent data collection, not necessarily indicating that conditions in that county being different from any other county. Figure 4 shows a residency that is under performing both in comparison to both their own historical performance and the region's. Their efficiency for the period at 97.43% compared to themselves and 94.47% compared to the region resulting in a loss of 325

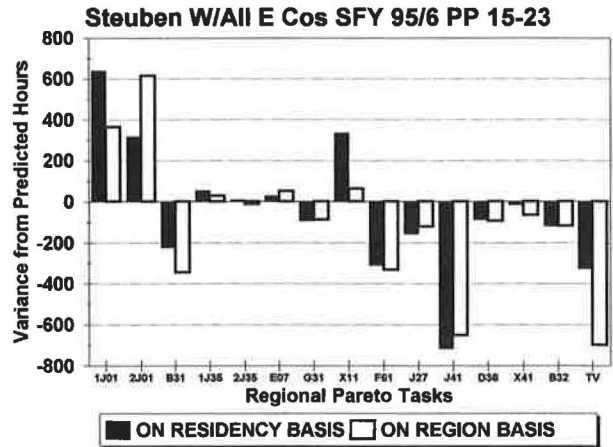


FIGURE 4 Winter task variance for a residency.

hours on the residency basis and 698 hours on the region's. The tasks driving down overall performance are primarily tasks B31- Temporary Pothole Repair, F61- Maintain Guide Rail, and J41- Mechanical or Manual Clearing of Snow and Ice.

Upper and lower control limits are also set as part of the variance charts to inform managers which tasks are varying beyond the limits of normal error. For the residency in Figure 4 the tasks falling outside control limits are 1J01 and 2J02 - One and Two Person Plowing, and J41. Task X11 - Maintain Signs exceeds the upper control limit on a residency basis only.

At the statewide and regional levels similar analysis is performed, except that rather than use Pareto analysis to determine which tasks to focus on, tasks are separated into each of the eleven mission-based the individual services that are delivered statewide. The analysis at this level is cost-based rather than time-based. From this a theoretical figure can be calculated of dollars earned or lost due to variation in performance.

## CONCLUSIONS AND LESSONS LEARNED

The most important lesson from trying to implement a performance-measurement system is that this process is not about measurement. The consequences of the measurement-development process are measures, but no real understanding of the mission of the organization, what services support that mission, and what roles people in the organization play in delivering service.

Without strong, knowledgeable and committed leadership, performance measurement cannot be successful. The organization must understand clearly what its strategic interests are, who its stakeholders and customers are, and service priorities and characteristics are essential in shaping services to the requirements of the organization's customers. This is a significant departure

from the traditional internally focused, and reactive organization.

Though the effort is difficult, and emotionally as well as intellectually challenging, it is worthwhile to take advantage of the opportunity for using performance measurement to bring meaning back to the term *public service*, and to simplify and clarify the organization's internal workings to make work purposeful, understandable, manageable and meaningful.

## REFERENCES

1. Wipper, L., "Performance Measurement: Producing Results at the Oregon Department of Transportation", Transportation Research Board, Washington, D.C., 1993.
2. Cohen and Brand, *Total Quality Management in Government: A Practical Guide for the Real World*, Jossey-Bass Publishers, San Francisco, 1993.
3. Martin, R., "Texas Tomorrow: Strategic Planning and Performance Budgeting", Presentation to NYS-GFOA Conference, Albany, N.Y., 1994.
4. Hutchins, G., *ISO 9000: A Comprehensive Guide to Registration, Audit Guidelines and Successful Registration*, Oliver Wright Publications, Essex Junction, VT., 1993.
5. Kaplan and Norton, "The Balanced Scorecard - Measures that Drive Performance", "Putting the Balanced Scorecard to Work", "Using the Balanced Scorecard as a Strategic Management System", Harvard Business Review, Boston, January -- February 1992, September -- October 1993, January -- February 1996.
6. Kaplan and Norton, *The Balanced Scorecard: Translating Strategy into Action*, Harvard Business School Press, Boston, 1996.
7. Deming, W., *Out of the Crisis*, Cambridge University Press, Cambridge, MA, 1986.
8. Felix, G., "Performance Measurement In Oregon State Government: Using The Productivity Matrix", Howard Publishing Company, Tualatin, OR, 1991.
9. Wipper, L., "Performance Measurement: Producing Results at the Oregon Department of Transportation", Transportation Research Board, Washington, D.C., 1993.