Performance Measurement: Producing Results at the Oregon Department of Transportation

LAURA WIPPER

In July 1989, a pilot project was developed to implement Performance Measurement at the Oregon Department of Transportation (ODOT). This program quantifies measures of efficiency and effectiveness for work crews and the department as a whole and equates these data on a common scale. Performance Measurement represents a change in philosophy. Rather than monitoring individual activities, the program focuses on results. Four key factors in the accomplishment of results are tracked and the outcomes are communicated on a regular basis. Efficiency measures gauge the volume of production and the cost, whereas effectiveness measures track quality and customer satisfaction. This new focus has seen increasing success as the 27 ODOT highway division work crews (approximately one-tenth of the total work force) participating in the pilot steadily improved productivity, culminating in savings amounting to $1.8 million over the last 6 months of 1991. The success of the pilot has not only led to full implementation of the program at ODOT but also has caught the eye of Oregon's executive department, which mandated the program for all state agencies.

State government in Oregon has evolved over the past century by adding commissions, boards, agencies, and, in turn, program upon program for what seemed important reasons at the time. Those reasons can become lost over time and needs can disappear, and yet activities and costs of programs often remain. Without a mechanism for ongoing evaluation, these factors can build inefficiencies along with a lack of effectiveness and accountability because of the absence of a clear mission, purpose, and focus.

The Oregon Department of Transportation (ODOT) had no readily visible signs of this malaise, yet in reality suffered from some of these symptoms. In 1988, the new state highway engineer, Don Forbes, asked the following questions, which did not always have answers at the time. How much does it cost to maintain the average lane mile? How accurate are construction contract estimates? Does the transportation planning process lead to accomplishment of department goals? What is the public's perception of the department? The search began for a method to provide answers to these questions and others and to quantify the efficiency and effectiveness of the highway division.

A program was chosen that had been developed at the Oregon State University Productivity Center by James L. Riggs and Glenn Felix. The program, called Performance Measurement, establishes measures of efficiency and effectiveness. The purpose is to improve performance by providing a tool to quantify and communicate results. It also provides data on which to base decisions for optimizing efficiency and effectiveness.

With the strong support of Forbes, now the ODOT director, Performance Measurement is currently in full implementation throughout ODOT. To date, more than 170 work groups have defined measures and are gathering data and receiving regular feedback on their overall performance.

Office of Productivity Services, Oregon Department of Transportation, 135 Transportation Building, Salem, Oreg. 97310.

WHY MEASURE PERFORMANCE?

A well-managed organization, be it public or private, needs to have a clear purpose and goals and objectives, base decisions upon data, provide regular feedback, and have some form of recognition for above-the-norm performance. The general state of our nation's economy suggests that many U.S. companies do not enjoy this type of management even under the powerful motivation of profits. Government agencies, too, suffer a similar lack. Initiatives for tax overhaul indicate that the public has lost confidence in government to operate efficiently and effectively. Over the decades, as layers of programs build up, a governmental organization can lose its focus without a regular, data-driven evaluative process in place. Performance Measurement provides that evaluative process for ODOT.

Performance Measurement clarifies the overall mission of ODOT and the purpose of its branches, sections, and units. It provides direction by presenting data against a backdrop of historical averages and historical potentials or goals. Presented in a matrix, seemingly disparate information can be converted to a common scale, which allows evaluation of the interaction between efficiency and effectiveness. This enables managers and staff alike to base decisions on data and to evaluate strategies for improvements to achieve the optimum balance between improved efficiency and effectiveness. This feedback is provided on a regular basis to help managers manage better at the program level and to communicate to those involved what is going well and what needs more attention. Because the focus is on programs and work groups, not individuals, teamwork is improved at all staff levels. The simple act of performance measurement alone usually prompts improvements because what is measured is what will surely get done.

KEY ELEMENTS OF PERFORMANCE MEASUREMENT

Results, Not Activities

Results are the point at which a product or service is delivered; activities are the actions that lead to delivery of that product or service. In the past, most forms of measurement at ODOT placed greater emphasis on forecasting and tracking activities—work load measurement. ODOT now places emphasis on results. Activity-based measurement only reinforces the accomplishment of activities; results-based measurement reinforces the accomplishment of results.

Group-Based Measures, Not Individual Measures

A key part of the process to develop performance measures is the involvement of the work group. Work groups are taught the concept of Performance Measurement and then facilitated in development of measures for their unit. In many cases, the individual members of the groups have minimal awareness of all functions of the group, so the discussion fosters a better awareness of the work group's priorities. Managers have reported improved work group cohesiveness following such discussion. Measuring the performance of individuals can be divisive, whereas measuring group-based results causes the members of the group to work better together to produce better results.

Performance Measures, Not Work Load Measures

Where work load measures capture just the number of activities, performance measures gauge results. When only activities are counted, desired results may not be produced because the focus is limited to the activities. This limited focus does not provide an environment to culture improvement strategies, whereas measurement of results does provide such an environment. As improvement strategies surface, they can be evaluated via the performance matrix.

Work Groups, Not Individuals, Develop Measures

The process of implementing Performance Measurement begins with a management team that develops broad guidelines. The work group then develops performance measures on the basis of their intimate knowledge of what they do and what they believe to be important. This ensures more accurate measures because the people who are actually doing the work know best what is being done.

Efficiency and Effectiveness Measures, Not Just Amount Done

Performance Measurement looks at both efficiency and effectiveness. Efficiency means doing the right things with the best use of resources. Effectiveness means doing the right things well and customer satisfaction with the product or service.

This program tailors measurement of quality to the product, service, and customer because quality holds different meanings for different people. For example, timeliness, accuracy, and availability of services equal quality for the driver and vehicle licensing functions of ODOT. Pavement condition and bridge sufficiency ratings are measures of effectiveness for not only highway maintenance, construction, and design, but also the department as a whole.

Credibility in State Government, Not Distrust of the Unknown

ODOT's goals, and those of other government agencies, and information about how well they are being achieved can be conveyed to the public via Performance Measurement. Budgets can be based on program performance and presented more effectively to the legislature because efficiency and effectiveness are demonstrated. This can also create a new role for government, which has not habitually played a proactive role in communicating exactly what it is trying to accomplish and how well it is doing.

The Visual Element

The performance matrix, a complex-appearing document, is actually how Performance Measurement keeps things simple. Once understood, the performance matrix will tell the user at a glance whether an entire organization's performance in key areas identified is improving or declining.

PERFORMANCE MATRIX

The matrix is not as complex as it initially appears. In fact, it can be understood in less than 30 min. In the sample matrix in Figure 1, Row A identifies emphasis areas of efficiency and effectiveness. Efficiency measures look at production volume and cost. Effectiveness measures look at such factors as timeliness, accuracy, and conformance to standards. A mandatory effectiveness measure is customer satisfaction, which is the customer's perception of products and services provided. Safety and work life quality are two more areas that should be included.

Row B identifies more specific key measures of performance important to the organization in each emphasis area. In the first column of Figure 1, the key measure is transactions per FTE (full-time equivalency). Row C contains the actual results achieved over the reporting period for each of the measures. In this sample matrix, the actual average transactions per FTE was 130.

Row D shows the potential results targeted to be achieved (in other words, a goal for each measure). Potential is based on either a historical best or an absolute goal such as 100 percent customer satisfaction or zero errors. The 10 is the level achieved when the goal is reached. In the example, the potential for transactions per FTE is 200.

			EFFICIENCY		EFFECTIVENESS				
A	Emphasis Areas		LABOR	COST	QUALITY		PERCEPTION	WORK FORCE	
B	Key Measures of Performance		Transactions Per FTE	Cost Per Transaction	Percent Delivered On Time	Percent Of Work Corrected	Percent Satisfied Customers	Work Life Quality Index	Safety
C	Actual Results		130	$2.30	90%	12%	80%	−10	0.11
D	Potential	10	200	$1.70	100%	0%	100%	100	0
		9	190	$1.75	98%	1%	98%	90	0.01
		8	180	$1.80	96%	2%	96%	80	0.02
		7	170	$1.85	94%	3%	94%	70	0.03
		6	160	$1.90	92%	4%	92%	60	0.04
		5	150	$1.95	90%	5%	90%	50	0.05
		4	140	$2.00	88%	6%	88%	40	0.06
		3	130	$2.05	86%	7%	86%	30	0.07
		2	120	$2.10	84%	8%	84%	20	0.08
		1	110	$2.15	82%	9%	82%	10	0.09
E	Baseline	0	100	$2.20	80%	10%	80%	0	0.1
		−1	90	$2.25	78%	11%	78%	−10	0.11
		−2	80	$2.30	76%	12%	76%	−20	0.12
		−3	70	$2.35	74%	13%	74%	−30	0.13
		−4	60	$2.40	72%	14%	72%	−40	0.14
		−5	50	$2.45	70%	15%	70%	−50	0.15
F	Level Achieved		3	−2	5	−2	0	−1	−1
G	Relative Weight		25	15	15	10	20	10	5
H	Earned Value		75	−30	75	−20	0	−10	−5
							Performance Index:		85

FIGURE 1 Performance matrix.

Row E lists baseline results or average, standard, or regularly expected performance based on historical averages. The 0 is the level achieved when average results are achieved. In this illustration, baseline for transactions per FTE is 100.

Because neither exactly average nor potential results are always achieved, a range of performance is also identified. Since performance, when measured, is more likely to be above than below average, ODOT's format contains 10 levels above baseline and only 5 below. The range between each level is determined by dividing the difference between baseline data and potential by 10. For transactions per FTE, 200 (potential) minus 100 (baseline) divided by 10 equals a range of 10 per level. This same range is taken in the opposite direction in the negative levels.

Row F is where the level achieved based on the actual results is shown. These levels are the common scale that can compare the interrelationships between measures that would otherwise be incomparable. The level achieved is reflected here because it is multiplied by the relative weight shown in Row G.

Relative weight in Row G is a method of weighting or prioritizing the key performance measures. By convention, all the relative weights in a matrix total 100. The assignment of relative weights is determined by the work groups once their measures have been developed. This process is somewhat arbitrary, but the measure of greatest importance is the measure with the greatest relative weight. Conversely, the measure with the lowest relative weight is the measure of lowest importance. In Figure 1, the labor efficiency measure, transactions per FTE, has the greatest weight, so it is of highest importance. The measure with the least weight and of lowest importance is a workforce measure, safety.

Row H shows the earned value of each measure, which is the result of multiplying the level achieved in each measure by its relative weight. For example, Level 3 was achieved in the transactions per FTE measure in Figure 1, which has a relative weight of 25, which equals an earned value of 75.

The performance index at the bottom of the matrix is the sum of the earned values for all measures contained in the matrix. This one number indicates overall how well an organization or work group juggled its priorities. A total of 0 means that the overall performance was average. A positive number means some degree of overall above-average performance. A negative number means some degree of overall below-average work. Because the relative weights must total 100, achieving potential in all measures would equal a performance index of 1,000; achieving level −5 in all measures would equal a performance index of −500, thus giving some relativity to the positive or negative degree of overall perfor-

mance. A performance index of 85 in the sample matrix indicates slightly above-average effort.

Various levels of achievement attained in each of the key measures contribute to an overall indicator. These measures can be evaluated individually to determine whether performance was below average in any specific area. When performance is below average in more than one area, the relative weights and the earned values can be examined to focus improvement strategies. In Figure 1, equal negative levels were achieved in two measures, cost per transaction and percent of work corrected. Cost per transaction would be the area of highest priority to improve because of its higher relative weight and greater negative earned value.

Analysis of the matrix in Figure 1 might reveal a work force working overtime to deliver increased products/services with a greater percentage on time. The negative side is a tired staff making more errors and working less safely. Increased timeliness counterbalanced by decreased accuracy accounts for average customer satisfaction.

IMPLEMENTATION PROCESS

Implementation begins with a steering committee consisting ideally of all senior managers or, at a minimum, the agency head, the budget officer, information services manager, personnel manager, and a performance coordinator. This group is taught the concept of Performance Measurement before going on to develop guidelines and performance measures that are very broad in scope.

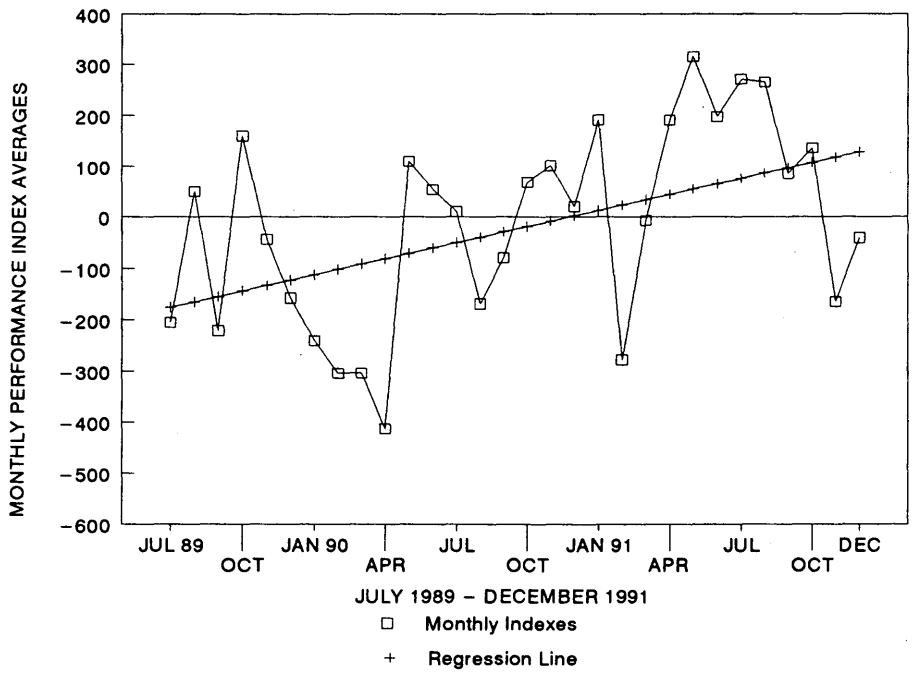
The midlevel management team participates in the same workshops as Performance Measurement progresses to the next level in the agency. This group develops measures that are still broad in scope, yet specific to that level in the agency while conforming to the guidelines and measures developed by the steering committee. The measures continue to get more specific as work groups learn Performance Measurement. Through workshops, they go on to develop their measures within the steering committee's guidelines.

At each level, the group decides what is important to measure within agency guidelines. This hierarchal approach allows data from all over the agency to feed into agency-level performance measures. For example, one motor vehicles division quality measure tracks timeliness, which is a measure of the percentage of transactions meeting service levels in 12 different service areas. The work groups then develop a measure to track the timeliness of the specific service offered by the group.

RESULTS OF PERFORMANCE MEASUREMENT

Figure 2 plots the overall average performance indexes of nine pilot work groups since July 1989. The overall average indexes fluctuate a great deal, but the trend is upward in improved performance, a shown by the regression line in Figure 2.

A recent comparison of the performance of seven additional groups with the performance of the original nine work groups shows virtually the same trend when compared on the same graph in Figure 3. Whereas this 6-month comparison (July through December 1991) might give an impression of declining performance, the expectation is for an overall improvement in performance similar to that in the original nine groups. When the performance of the original nine groups is viewed



Indexes are based on overall monthly performance indexes, not including safety measures

FIGURE 2 Performance index averages.

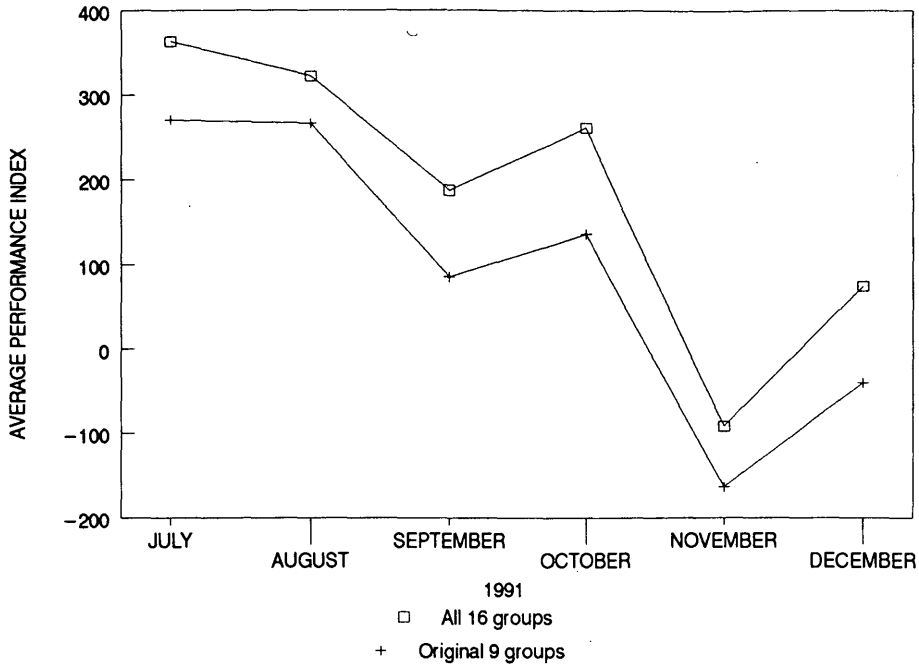


FIGURE 3 Average performance index (July–December 1991).

over a longer period, it fluctuates, but the long-term trend is upward (Figure 2).

Figure 4 shows dramatic improvements in the cumulative 6-month average indexes for the same nine pilot groups. The trend is again clearly upward but even more evident when comparing the same period, July through December, in 1989, 1990, and 1991.

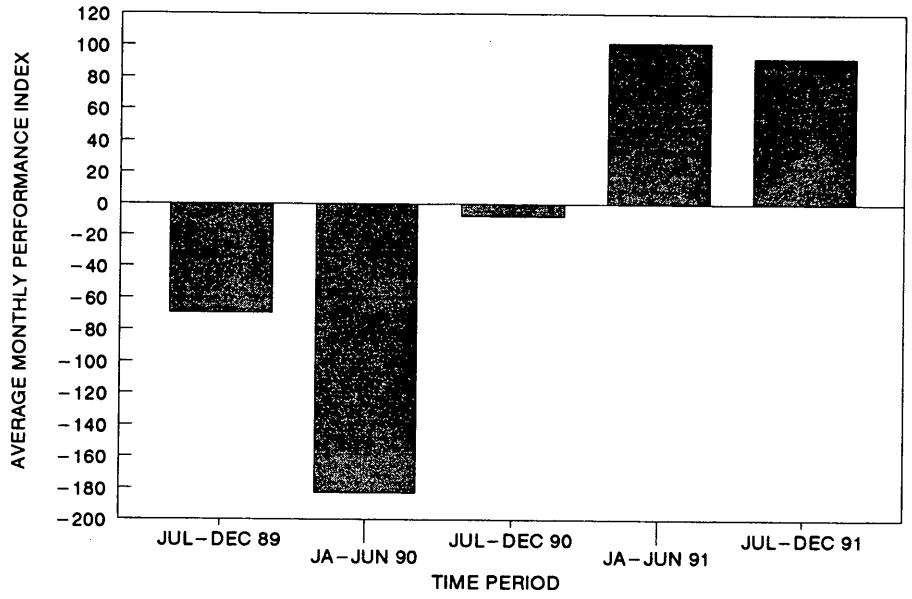
Figures 5 and 6 show performance trends for two specific areas: efficiency and effectiveness. Efficiency measures look at volume of output and cost, whereas effectiveness measures quantify product/service quality and customer satisfaction. The average total earned values of the measures of effectiveness hover around baseline performance in Figure 5. Figure 6 shows a familiar trend. Improved performance causes the graph of average total earned value for efficiency measures to be virtually the mirror image of the overall average performance indexes shown in Figure 2. The slight downward trend in average total earned value for effectiveness measures over the last 6 months of 1991 warrants further analysis to maintain the optimum balance between efficiency and effectiveness.

Figure 7 isolates a single measure of effectiveness, customer satisfaction, for the nine original work groups. Although efficiency has improved, customer satisfaction has hovered so close to baseline performance that it is essentially a straight line. One group, weighmasters, who weigh and inspect trucks and subsequently cite truck drivers when the truck is out of compliance, expected very negative results from a customer survey. This group was pleasantly surprised by the initial results and went on to increase the satisfaction rate by 50 percent in the next 6 months.

LESSONS LEARNED

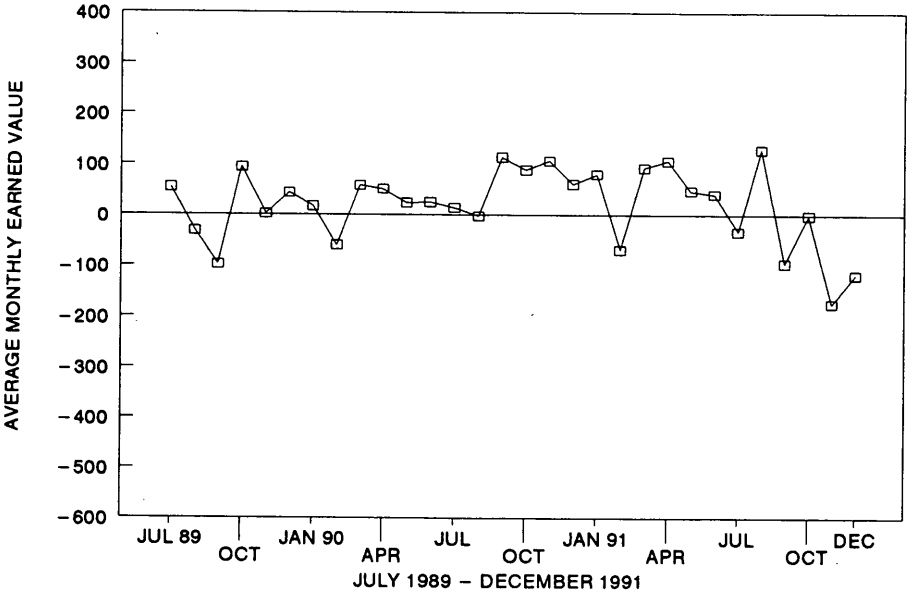
During the pilot phase of Performance Measurement, we learned that the program could beneficially affect results. In addition to seeing improved performance, four key lessons were learned to better implement the program department-wide.

1. An automated reporting process must be in place before agencywide implementation begins. Without automation, data gathering can become extremely labor intensive, making it difficult to produce timely reports. Once the measures have been developed and data gathering has begun, work groups are anxious to receive regular feedback. Confidence in the program can be lost if this part of the program is not performed.
2. Union representatives must be involved at every step of both a pilot and full implementation to learn the concept, the process, the reasoning behind steering committee guidelines and, above all, to realize that performance measures are based on results produced by a group and are not individually focused.
3. A communication and decision-making process must precede agencywide implementation. The steering committee must decide such things as the level of the agency responsible for review of the measures, baselines, and potentials; the frequency of review; and the criteria to be used to determine baselines.
4. All levels of management must be actively involved in the Performance Measurement process and kept informed.



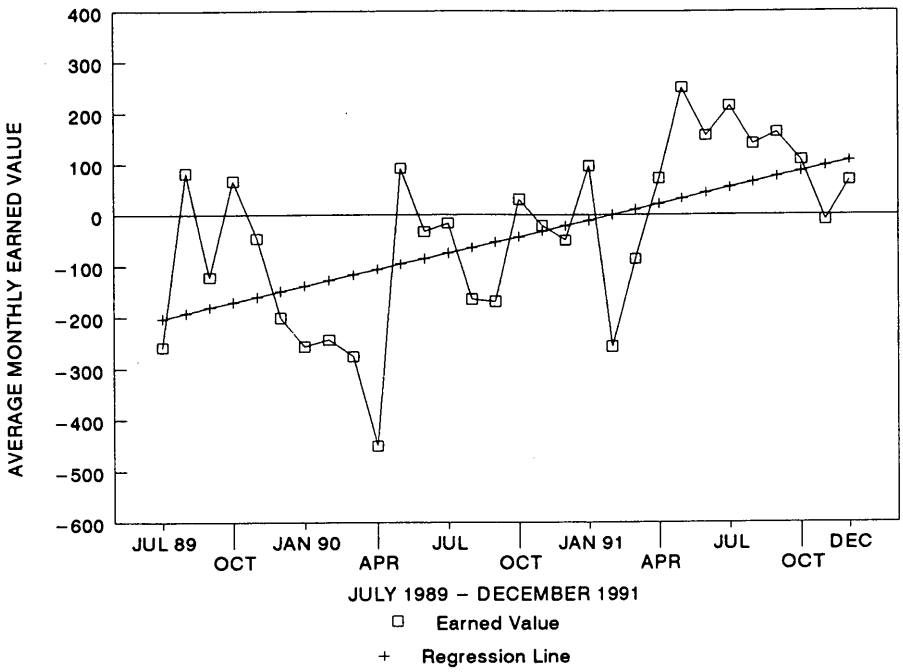
Indexes are based on overall average monthly performance indexes, not including safety measures

FIGURE 4 Six-month performance index averages.



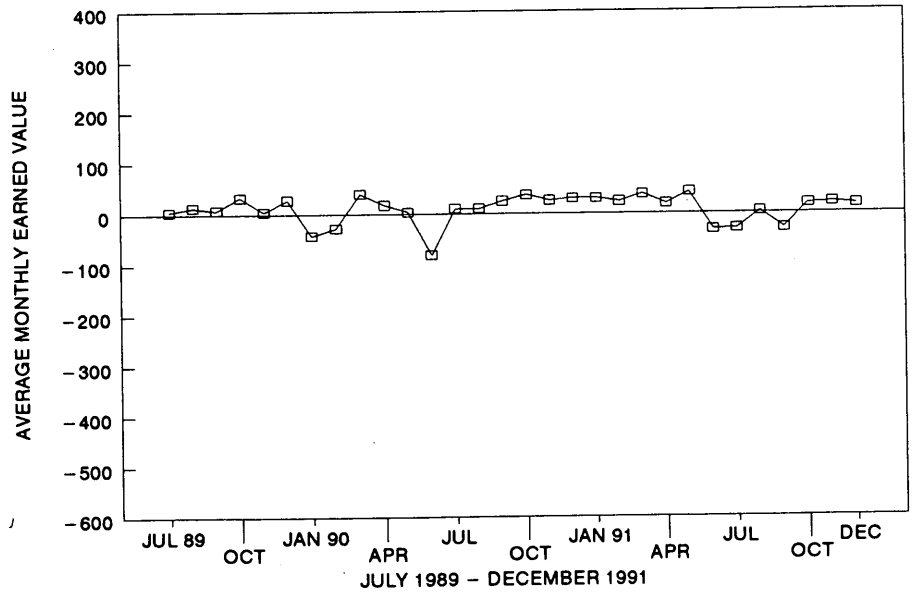
Earned value calculated by adding monthly effectiveness measures earned values

FIGURE 5 Average monthly effectiveness earned value.



Earned value calculated by adding labor efficiency measures earned values and materials measure earned value

FIGURE 6 Average efficiency earned value.



Earned value is customer satisfaction measure earned value for each group

FIGURE 7 Average customer satisfaction earned value.

In addition, senior management must understand, support, champion, and promote the program.

ODOT has been quick to incorporate these improvements into the program to streamline implementation as it continues through the agency.

CONCLUSIONS

From July through December 1991, 27 pilot work groups, amounting to 7 percent of the work force or 350 FTE, saved ODOT $1.8 million through improved efficiency and effectiveness. In addition, if success can be measured by what others imitate, Performance Measurement at ODOT can be considered a resounding success. What began as a pilot program within ODOT has become a full-scale initiative throughout state government in Oregon. The Oregon Executive Department recognized the value of the Performance Measurement program and mandated it for all state agencies. ODOT was instrumental in the success of this initiative by teaching representatives from over 115 state agencies the program concept and implementation. ODOT expertise assisted countless agency management teams with development of performance measures.

Oregon was awarded the "E for Effort Award" by *Financial World* magazine as a result of the magazine's annual evaluation of state government. The award is given to honor a state that has taken a leadership role in dealing with present issues facing state government. In the annual rankings by *Financial World*, Oregon has moved from 34th in 1990 to 17th in 1991 to 6th in 1992. The state's "trailblazing work in performance measurements" was the primary reason cited by the magazine for Oregon's movement into the top 10.

Beginning in August 1992, all agency budgets will be presented in the context of agency-level performance measures, thus providing a consistent platform to communicate the efficiency and effectiveness of programs throughout state government. Other agencies plan to do so, but to date, ODOT is the only agency implementing the program at all levels in the organization. We look forward to improved management of transportation programs via Performance Measurement.

Publication of this paper sponsored by Committee on Management and Productivity.