

Using a Performance Budgeting System: Lessons from the Texas Experience

Joe Adams
Research Associate
John C. Stennis Institute of Government
Mississippi State University

Introduction

One of the most touted management improvements during the past few decades has been the introduction of performance measures for government entities. Performance measures are often viewed as a kind of “bottom line” for the purpose of providing accountability in government.

In addition to scores of books and articles, numerous conferences, presentations, and training opportunities are available on the use of performance measures, but very few of them provide information or evidence indicating that using performance measures produces better performance, particularly in government settings. Most examples of empirically successful performance measurement systems have been taken from the private sector, where performance is generally defined in financial terms. As many governments adopt progressive business models, performance measures have become more prevalent. For many government entities, performance measures have become standards of organizational success in the absence of market indicators of success such as profits or dividends.

This chapter describes how the Texas system operates. It also identifies lessons from the implementation of the system over the past decade as well as offers recommendations for improvement in several areas.

The Texas performance budgeting system has become highly institutionalized within the agencies and the Legislature. As a result, many of the system’s practices have grown organically in response to political, fiscal, and institutional forces, many of which stem from the biennial budget cycle itself.

Many factors in the political process have limited the adoption of a more coherent, systematic approach to target setting and interpretation. But these need to be addressed if the system is to become more results oriented and less compliance oriented. The key issue for improved budgeting and performance management is better identification of performance drivers for costing and more sensitivity to the potential manipulation of performance data by agencies to make themselves look better. For example, while most agencies complied with the performance targets set in appropriations law in recent years, only a little more than half of the targets showed improved program performance over time, suggesting that the target-setting process may not be very rigorous.

For agency managers, having multiple and sometimes opposing performance indicators leads to confusion regarding agency objectives and priorities. For the public and for policy makers, performance measures would be more meaningful and understandable if there were greater standardization across agencies as well as a consistency between state and federal indicators. Also, in some instances, huge portions of the state budget have no performance indicators at all, especially in the area of general administration.

As government devolves and privatizes many of its programs, more attention needs to be given to the practice of using performance measures as contract management instruments. Moving to more performance-based contracting will require a rethinking of performance standards, from a *post hoc* compliance system to an early warning management system that provides leading indicators. Such a rethinking will also require a better understanding of internal agency operations, which is an area largely ignored by the Texas performance budgeting system. Whatever direction policy makers take, navigating the current system and building on its lessons is essential to getting to the other side.

The Texas Legislative Budget Process

Because of post-Civil War abuses of power by the governor, the 1876 Texas Constitution vests most key authorities in the Legislature, then restricts it to meeting but once every two years, for a 140-day session. The executive branch is composed mainly of boards and commissions, to which the governor has appointment, but not removal, authority. As a consequence, the Texas budget system is largely driven by the Legislature, which, when it is not in session, delegates its authority to the Legislative Budget Board (LBB), a 10-member joint standing committee chaired by the lieutenant governor, co-chaired by the Speaker of the House, and composed of key legislators. The LBB oversees the budget development and monitoring process, and develops a recommended legislative appropriation for all state agencies. It provides the Legislature with a recommended state budget at the beginning of each legislative session, which occurs in early January of odd-numbered years.

The Office of the Governor’s Budget, Planning, and Policy Division is largely advisory, and generally works collaboratively with the LBB at the staff level. The Governor’s Office also provides its budget proposals at the beginning of each session, using the same format as the LBB. In the 2003 regular session, Governor Perry issued a zero-based proposal for the FY 2004–2005 biennium, with no performance measures; however, it contained a new “informational objects of expense,” which harkens to a more traditional budgeting approach. Speaker Craddick and Lieutenant Governor Dewhurst endorsed Governor Perry’s general approach, and it was reflected in House Bill 1 and Senate Bill 1 introduced by House Appropriations Chairman Heflin and Senate Finance Chairman Bivins. The introduced bill, however, retained key performance measures.

Performance Measures in Texas State Government

The State of Texas is a well-known leader in the development and application of performance measures in government. Beginning in 1991, the state embarked on a comprehensive strategic planning process for all state agencies. House Bill 2009, 72nd Legislature, Regular Session, 1991, established the process, requirements, and time frame for the first planning cycle, with the pledge of then-Governor Ann Richards to make agencies more accountable and customer focused. Governors George W. Bush and Rick Perry continued the effort, with an emphasis on commitment over compliance.¹

Compliance is an issue in Texas. State government is highly decentralized, with over 200 state agencies and institutions. Moreover, most of them have their own governing boards, most of which are appointed by the governor, although a few—such as the Attorney General, Comptroller of Public Accounts, Department of Agriculture, the Railroad Commission, and the courts—are headed by elected officials. Decentralization is an ingrained feature of Texas government, so much so that it remains the only state in the Union with no centralized human resources department.

House Bill 2009 imposed on state agencies a new way of doing business, requiring a strategic planning process and the development of performance measures. Since then, the Legislature has amended this

statute in subsequent sessions, but the general process has been thoroughly institutionalized.

It is important to understand what a massive undertaking the system was to build. The system encompassed more than 200 entities that employ about a quarter million people. In theory, each individual's work is tied in some way to a performance measure. The performance budgeting cycle is an intensely busy activity. It is also an iterative process that builds on structures set by the Legislature during previous biennia (see Table 9.1 on pages 430–431).

Each biennium, in June of even-numbered years, all state agencies and institutions of higher education are required to submit an updated five-year strategic plan, including specific targets for their desired outcome measures. That document, which is usually prepared just before the beginning of the even-numbered fiscal year, provides the basis for the agencies' Legislative Appropriations Request (LAR), which is structured in a "pattern" resembling the agency's strategic plan, complete with goals and strategies, along with the funding request for each strategy.

Instructions accompany each step in the process, from changes to an agency's strategic plan to the submission of its LAR, as well as the performance reporting periods following the enactment of the General Appropriations Act (GAA). Even before the start of the process in even-numbered years, teams of analysts from the Legislative Budget Board (LBB) and Governor's Office of Budget and Planning (GOBP) work on drafting concise instructions, while state leaders consider broad policy guidelines for agencies to follow in preparing their LARs and subsequent reports. As a result, instructions may vary from one biennium to the next as policy makers focus on this or that specific concern. Travel budgets may be an item of particular interest one year, employment or full-time equivalent (FTE) position levels may take more interest the next year.

In general, one or two analysts from the LBB and at least one analyst from the GOBP are directly involved in shepherding the development of an agency's strategic plan and reviewing requests from the agencies. For the most part, the two agencies work in tandem during the strategic planning process until the first budget hearings. LBB and GOBP analysts co-chair the initial budget hearings held in the fall, just prior to the biennial legislative session. This is the agencies' first chance to publicly present their requests for the next biennium. At this point, the LBB and GOBP go their separate ways. Shortly after the hearings, analysts at the LBB make their recommendations to the board, which will be vetted and refined. The agency's requests and the LBB's recommendations will be published in the form of the Legislative Budget Estimates (LBE). And a draft bill will be introduced containing the same information. Meanwhile, the GOBP begins a parallel vetting process to produce the governor's budget recommendations, based on the same structure that both entities agreed upon in approving agency strategic plans.

Acronyms

ABEST II	Automated Budgeting and Evaluation System of Texas
GAA	General Appropriations Act
GOBP	Governor's Office of Budget and Planning
LAR	Legislative Appropriations Request
LBB	Legislative Budget Board
LBE	Legislative Budget Estimates
PBBS	performance-based budgeting system
PRWORA	Personal Responsibility and Work Opportunity Reconciliation Act
SAO	State Auditor's Office
TANF	Temporary Assistance for Needy Families
TCWEC	Texas Council on Workforce and Economic Competitiveness
TDHS	Texas Department of Human Services
TWC	Texas Workforce Commission
WIA	Workforce Investment Act

Table 9.1: Texas' Strategic Planning, Performance Budgeting, and Performance Monitoring System, Two-Year Cycle

Even Years	Planning	Budgeting	Monitoring
Jan.	Statewide goals established; Agency Strategic Plan and Information Resources Strategic Plan Instructions issued		1st Quarter Performance Report due
Feb./March	Agencies request changes to budget structures		
April		Agency Budget Request and Information Resources Operating Plan Instructions issued	2nd Quarter Performance Report due
May/June	Agency Strategic Plans and Information Resources Strategic Plans due		
	Budget structure changes approved		
July		Budget requests due	3rd Quarter Performance Report due
Aug./Sept.		Budget hearings held	
Sept. 1		Fiscal Year begins	
Oct.		Budget recommendations developed	4th Quarter Performance Report due
Nov.		Budget recommendations developed	Quality Assurance Team Report issued
Dec.		Budget recommendations developed	Annual Performance Assessment Report issued

Table 9.1: Texas' Strategic Planning, Performance Budgeting, and Performance Monitoring System, Two-Year Cycle (continued)

Legislative Session	Odd Years	Planning	Budgeting	Monitoring
	Jan.		Budget recommendations submitted to Legislature	1st Quarter Performance Report due
	Feb./March		Budget mark-up	
	April		Budget mark-up	2nd Quarter Performance Report due
	May		Budget approved	
	June		Budget sent to Comptroller for Certification; budget sent to Governor for signature	
	July			3rd Quarter Performance Report due
	Aug./Sept.		Agency Operating Budget and Information Resources Operating Plan Instructions issued	
	Sept. 1		Fiscal Year begins	
	Oct.		Agency Operating Budgets due	4th Quarter Performance Report due
	Nov.			Quality Assurance Team Report issued
	Dec.			Annual Performance Assessment Report issued

Source: Texas Legislative Budget Board

As shown in the example of the Aircraft Pooling Board bill, strategies constitute the line items in the appropriations bill. As such, they effectively limit the governor's ability to exercise the line-item veto, because the governor cannot modify a line, but only veto it or not. In this particular instance, the governor decided to veto the appropriation for the entire agency.² Although the original strategic plan requires joint approval by the

Legislative Budget Board and the Governor's budget office, strategies can be combined, eliminated, or added at the pleasure of the Legislature during the legislative session, along with any performance measures that the members may decide to add or eliminate. Strategies within the larger agencies can be as large as \$500 million and contain multiple programs, or a few thousand dollars in very small agencies.

In general, an agency will have three to five substantive strategies, sometimes referred to as "direct strategies," as well as one or more strategies labeled "indirect administration" for functions shared among strategies, such as accounting, human resources, information technology, reporting, and overall administration in the higher executive offices. Some larger agencies and institutions of higher education may have as many as 20 to 30 strategies, though many of these additional line items may not contain a high percentage of the agency's funding and do not generally have any performance measures associated with them. Some smaller agencies do not have an indirect administration strategy at all, allocating those costs among the direct strategies, while agencies with indirect administration strategies must show their cost allocations as related to direct strategies. These points become increasingly important when the calculation of efficiency measures are considered, because few, if any, cost-per measures include the amounts spent in indirect administration.

What may be even more surprising is that indirect administrative strategies typically range from as little as 5 to as much as 30 percent of total appropriations. However, mere percentages can mask significant differences between agencies. While 30 percent of a \$4 million agency's budget may seem large, it is worth remembering that 5 percent of a billion-dollar budget is an even larger amount. Because of economies of scale, larger agencies tend to spend a lower percentage on indirect administration than smaller agencies.

The Legislative Appropriations Request, or LAR, also contains reported performance for the past two years, projections for performance measures for the coming year, and additional targets for output and efficiency, as well as projections for explanatory measures. This information is loaded into the LBB's Automated Budgeting and Evaluation System of Texas (ABEST II), a constantly evolving computer system that has been web-enabled for agency reporting. ABEST II provides the backbone for information used to publish the LBB's recommendations in the Legislative Budget Estimates, in the House, Senate, and conference committee versions of the appropriations bill, as well as in the General Appropriations Act.

Since the establishment of the performance measurement system, results have been tied to incentives and penalties in the GAA. For example, agencies exceeding 80 percent of the established key performance targets were eligible to implement, with permission from the LBB, salary increases for agency employees.³ On the other hand, an agency's failure to perform

A BILL TO BE ENTITLED AN ACT

appropriating money for the support of the Judicial, Executive and Legislative Branches of the State government, for the construction of State buildings, and for State aid to public junior colleges, for the period beginning September 1, 2003 and ending August 31, 2005, authorizing and prescribing conditions, limitations, rules and procedures for allocating and expending the appropriated funds; and declaring an emergency.

BE IT ENACTED BY THE LEGISLATURE OF THE STATE OF TEXAS:

ARTICLE I

GENERAL GOVERNMENT

Sec. 1. The several sums of money herein specified, or so much thereby as may be necessary, are appropriated out of any funds in the State Treasury not otherwise appropriated, or out of special funds as indicated, for the support, maintenance, or improvement of the designated agencies of general government.

AIRCRAFT POOLING BOARD

	For the Years Ending	
	August 31, 2004	August 31, 2005
A. Goal: EFFICIENT AIRCRAFT TRAVEL		
To provide safe, efficient aircraft transportation upon request to state officers and employees traveling on official state business.		
A.1.1. Strategy: FLEET OPERATIONS	\$ 2,174,296	\$ 774,296
Operate a fleet of currently scheduled aircraft to meet users' needs.		
Output (Volume):		
Number of Hours Aircraft Are Flown	2,400	2,320
Number of Flights	2,262	2,188
B. Goal: AIRCRAFT MAINTENANCE		
To provide quality and economical maintenance, fuel and oil, hangar space, and line services for state-operated aircraft.		
B.1.1. Strategy: AIRCRAFT MAINTENANCE	\$ 2,172,640	\$ 2,172,640
Operate a state-of-the-art maintenance and avionics repair facility.		
Output (Volume):		
Number of Work Orders Completed	660	650
C. Goal: INDIRECT ADMINISTRATION		
C.1.1. Strategy: INDIRECT ADMINISTRATION	\$ 216,286	\$ 216,286
Grand Total, AIRCRAFT POOLING BOARD	\$ 4,563,222	\$ 3,163,222
Method of Financing:		
General Revenue Fund	\$ 1,400,000	\$ 0

as expected could bring a negative response. In April of 1994, the LBB used reported performance measures during the first six months of the biennium as the basis to make reductions in appropriations for seven agencies.⁴ In addition to budget execution, which is a power jointly shared with the Governor's Office, the GAA grants the LBB authority to raise or lower performance targets during the interim between legislative sessions.

Reporting Requirements

Not all of the agencies' performance measures appear in the GAA. In fact, only a fraction of the total number of measures may appear in the appropriations bill. Those that do are designated "key" measures. The remaining measures are called "non-key" measures.

The designation of key and non-key measures is generally left to the staff of the LBB as recommended to the committees, and may be changed at any time under special provisions contained in the GAA. In some cases, the Appropriations Committee may direct the LBB staff to add measures, as in the case of new programs or strategies, or they may eliminate entire strategies and their measures, as in the case when those strategies are unfunded. Agencies are supposed to integrate these changes in the next planning cycle. They are expected to incorporate new legislation and new performance measures into the appropriate structure when they revisit their plans during the subsequent biennium. In the meantime, they must begin implementation of the legislation enacted, usually by September 1, three months after the session. And that is not much time to make significant changes in programs if agencies need to adopt rules or add new departments. Performance-measure data gathering is probably the last item on management's "to do" list after the session.

In recent years, analysts have been instructed to reduce the number of measures in the appropriations bill. At around 800 pages, such a request is intuitively understandable. However, cutting out many of the measures can make the system appear less coherent. As new measures enter the system during the legislative process and apart from the strategic planning process, a breakdown in the logic of the performance-budgeting nexus occurs. There tends to be a loss of elegance in being able to multiply outputs by efficiency measures to arrive at the line item from what is reported on the pages of the appropriations bill. If one of the measures is not made key, the logic of the appropriation is hidden, literally unseen to readers. Efficiencies may not tie to outputs and line items, making it impossible to calculate the necessary dollar amount of a line item from the projected performance measures. As a result, the system becomes increasingly obscure to uninitiated readers of the GAA.

Having noted that the state's internal process for the development of performance measures is fraught with a number of twists and turns, the matter is further complicated by the state system's place in the larger context of accountability.

Layers of Performance Measures: Accountability to Whom?

Performance measures have trickled down to almost every level of government—federal, state, and local. While Texas began its efforts to develop performance measures in the early 1990s, the federal government also began to establish performance standards for federal agencies with the passage of the Government Performance and Results Act in 1993. Although the details of that effort have been widely studied, it is important to understand the impact of federal performance measures on the system in Texas. In some cases, there are as many as four layers of accountability.

One of the major results of the simultaneous development of performance measures in Texas was that state agencies that administer federally funded programs inherited a dual performance reporting role, one to the Texas Legislature and another to the cognizant federal agency.

This may seem like a mundane point, but the fact that federal and state fiscal years, quarters, program years, and grant periods frequently do not mesh became an unending reporting nightmare, particularly in the areas of welfare reform and workforce development. Oversight committees and legislative staff continually request explanations or ad hoc reconciliation of competing performance figures. At times, a frenzy of questions swamp agency personnel because of built-in confusion resulting from competing reporting cycles, definitions, and formats. Such fruitless exercises bleed resources away from analyzing trends or examining processes.

Examples of Layered Measures

Some of the most significant pieces of federal legislation affecting state performance reporting systems were the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996 and the Workforce Investment Act (WIA) of 1998. Both pieces of legislation made sweeping changes in the implementation of their respective program areas. With PRWORA came new funding in the form of block grants to states. Temporary Assistance for Needy Families (TANF) replaced Assistance to Families with

Dependent Children, and contained numerous changes in reporting and tracking requirements for welfare recipients. Likewise, WIA replaced the Job Training Partnership Act of 1982. Both of the predecessor programs possessed large quantities of data, case records, and associated operating procedures. (For a discussion of indicators used in the WIA program, see chapter seven.)

At the state level, Texas took the lead in anticipating legislation in both areas, consolidating workforce programs and establishing a framework—similar to the federal government's for welfare reform—in House Bill 1863, enacted in 1995. The resulting transfer of responsibilities between the Texas Department of Human Services (TDHS) and the newly formed Texas Workforce Commission (TWC) included the wholesale transfer of over 300 performance measures. Eventually, the total number of key and non-key measures was reduced over the course of the next four years to 180, then to 154, and again to 104. These changes required that most measures undergo a process of redefinition, on the fly, as new collection, data warehousing, and reporting systems were being developed. As scores of measures were being systematically migrated to new collection and reporting systems, major changes were occurring in associated accounting systems. A census of reporting requirements at the TWC completed in 1998 revealed that the agency maintained 154 state performance measures, including 75 key performance measures reported to the LBB, as well as over 100 additional measures to federal agencies. Needless to say, there was little time for much in the way of evaluation during the construction of the new reporting systems.

During this same period, the Texas Council on Workforce and Economic Competitiveness (TCWEC), a trusted agency within the Governor's Office, was charged with developing "core measures" with which to evaluate the state's workforce development system from a broad, multi-agency perspective. This ongoing effort resulted in a number of highly aggregated measures that were not yet fully defined in 1995. The same year, the GAA required that the Texas Workforce Commission report several of these measures to the LBB and incorporate them into its strategic plan. In subsequent years, policy makers asked that the measures be disaggregated by program, which multiplied the number of potential variances to be explained.

To compound matters, TWC was engaged in block-granting child care, TANF, and WIA programs to local workforce development boards through contracts that included and tried to integrate many of the same measures (by federal funding stream) as reported to the Legislature and the federal government.⁵ Local boards complained about the relevance and number of contract measures, given the fact that previously segmented programs were supposed to be integrated in one-stop centers. Their concerns were validated in a State Auditor's report, which stated:

State law requires local workforce boards to administer programs through a one-stop service network. Therefore, the Commission administers its workforce programs primarily as members of a single workforce system rather than as individual programs. Because it administers workforce programs in this manner, the Commission should assess overall performance as a whole as well.⁶

Despite the modifications in the law, and the development of a new delivery system, narrowly defined performance indicators continue to be based on federal funding streams. Instead of substituting systemwide measures for program measures, systemwide measures were piled on top of program measures, adding to the number of items competing for attention.

The overall result of the piling on was twofold. First, instead of becoming more focused, like a microscope, the addition of multiple layers of related measures became a kaleidoscope of differing perspectives, endlessly fascinating yet analytically baffling and ultimately distracting. Second, keeping old program-specific measures along with the new aggregate measures meant that the entire system, while attempting to become integrated and coordinated, was driven back into programmatic segmentation by the attempt to track federal dollars and federal measures.

Unfortunately, the measurement system contradicted state policies about integration. In a sense, performance measures took on a life of their own. Institutional inertia kept out-of-place remnants from the bygone era alongside measures for a new system. Reluctance to drop measures without the full consideration of the Legislature, which connected some programmatic measures to appropriation riders, and congressional action between state biennia caused the state system to be at an odd tangent with changes in federal policy. Until the two regain alignment, the state's measurement system threatens to supersede state policy on program integration.

Exemptions from Performance Measures

While some agencies may seem overwhelmed with measures, and some program areas have multiple indicators for performance, others have no performance measurement at all. As mentioned earlier, most agencies of any size have a strategy for "indirect administration." It is a separate line item in the GAA, sometimes split into distinct functions such as "information systems," "central administration," "regional administration," and "other support services." Amounts for indirect administrative strategies range from as little as 5 to as much as 30 percent of an agency's total appropriations, or from as little as tens of thousands of dollars to as much as \$50 or \$60 million. In fact, as much as \$79.9 million was appropriated for indirect administration in one year for the Texas Department of Criminal Justice, which

included a large information technology project.⁷ None of these strategies have performance measures.

Likewise, none of the system offices for the major university systems have performance measures. According to LBB performance assessments, public community and junior colleges have 400 to 500 measures that they report to the LBB, but they are not listed in the GAA. Appropriations for those institutions totaled \$789 million in FY 2002.⁸

While the system is designed to assist in the appropriations process, there are significant gaps. Even if only 5 or 10 percent of the budget lacks any attached performance, that is still a significant amount of money appropriated without any obvious performance requirements.

The Target-Setting Process

In the end, the elegant notion of finely tuned outputs multiplied by precisely calculated efficiencies to provide total costs withers during midnight decisions on pending items late in the legislative session. After months of slowly plowing through each agency's request, with questions that are generally focused on specific objects of expense—computers, travel, FTEs, and the like—an explosion of decisions erupts in the twilight of each session as each chamber passes its version of the appropriations bill and the conference committee meets to resolve the differences.

Things move fast late in the legislative session and new information becomes harder to obtain in a timely manner. At these times, best guesses are thrown into the mix as decision makers grasp for more information or a sense of "what would happen if." Sometimes performance measures are made in the hallways, where Legislative Budget Board analysts sometimes huddle with agency directors to piece together a performance measure for a new program. In the haste to fashion an amendment, a flurry of activities occurs. Neatness and consistency are the first casualties of these last-minute modifications to the appropriations bill. Start-up and indirect costs may be accidentally stranded. Fast promises and assumptions are made about implementation and capacity to absorb funding, and other hastily developed pieces about spreading the funding between years are thrown into spreadsheets or on the backs of briefings papers, while inside the committee room another member of the committee scrawls a rider specifying that funding is contingent upon a special quarterly report on performance. This is the reality of the performance-based budgeting system at work.

All of this ends up in the General Appropriations Act to be implemented in another three months. The agencies will work out the details in the months to come, with LBB analysts nipping at their heels and asking for

status reports. Meanwhile, data collection on the new measures is being started as well.

After several legislative sessions, tracing the origins of a specific measure and the reason for its existence may be a significant challenge, even to experienced insiders.

Events like these are outside the control of even the best planning process. LBB budget analysts work within broad guidelines set by the director and senior management, which are frequently coordinated with colleagues at the Governor's Office of Budget and Planning. With the development of the performance-based budgeting system, the responsibilities of analysts have evolved to combine budget and performance analysis. Many analysts serve both roles for small to medium-sized agencies. For large agencies with complex funding streams and large sets of measures, it is not uncommon for pairs of analysts to be assigned, with one having primarily a budgetary focus and the other a focus on performance measures.

As a general rule, LBB analysts recommend targets that meet or exceed an agency's maximum actual performance during the previous two or three years. In rare instances, when funding is expected to decrease for a particular program or strategy, they may reduce the performance targets for the associated outputs and outcomes; however, analysts are reluctant to increase efficiency measures, such as cost per unit or days per output, even in the face of stridently argued agency requests for increased targets. When such recommendations are made, the analysts make it a point to say that the agency has requested such an increase. These practices are designed to put the burden of proof on the agency for budget and performance changes. Nevertheless, the actual process becomes much more fluid once the legislative session begins.

During the appropriations hearings, agencies will typically repeat a slightly more refined version of their presentations made earlier to the LBB and GOBP. In some cases, better information has been received or developed that may change an agency's request. Sometimes there are technical reasons for changes, as in a situation where caseload projections may be revised or amended as a result of new information. Other times it may be the result of instructions from policy makers who balk at program costs. At times, the Governor's Office or legislative offices may have signaled that a change is deemed appropriate for a new initiative or may direct that an agency develop funding options for consideration. In the legislative process, subcommittees and individual members of appropriations committees are frequently assigned the task of working out precise details with specific agencies. Sometimes these deliberations result in a new strategy or a new set of performance measures.

Another source of changes in the budget structure for an agency is new legislation. Usually toward the end of the legislative session, the Appropriations Committee may consider new funding, contingent upon the enactment of pending legislation. Enactment of new legislation may create a new

program, abolish an existing program, change the focus of an existing program, or shift program responsibilities between agencies. Entire agencies may be abolished, created, or replaced. Everything is subject to change.

As mentioned earlier, these changes are often made under severe time constraints, when measures are proposed, developed, and adopted within a matter of just a few days or even hours. An analyst may be assigned the task of developing new measures and targets for those measures without the benefit of historical data or comparable benchmarks. Sometimes there is no information other than the intent of the legislation. In fact, the process may result in the collection of entirely new data. The result of this process is that targets are sometimes the result of truncated analysis and incomplete information. Rough estimates must suffice in these situations to meet the needs of policy makers, whose intent is to maintain accountability.

Defining Success: Does the System Work?

The development of the measurement system in Texas is not a tightly conceptualized or neatly packaged affair. The Texas performance-based budgeting system (PBBS) was designed to meet the needs of the Texas Legislature. Prior to the establishment of the current system, the Legislature used a zero-based budget fashioned around objects of expense that gave little information about the services that would be provided with specific appropriations. Performance measures provided a means of identifying the use of the funds in a way that objects of expense cannot.

Decision rules regarding performance have never been clear and seem to be determined on an ad hoc basis. Strong performance is not necessarily rewarded, nor is poor performance necessarily punished. Though the door is open both directions, the Legislature reserves discretion to do as it chooses, for its own reasons, independently of what performance measures may indicate. Indeed, the case for additional funding is a multifaceted enterprise, a matter of general policy more than technicalities. Nevertheless, discussions in the House Appropriations and Senate Finance Committees often focus on such specifics as changes in caseloads, matching federal funds, and revenues, if any, collected by the agencies. Typically, more time and discussion is devoted to an agency's business practices, such as its capital spending, adherence to travel limitations, personnel caps, and the like, than to its performance measures.

Given this context, is the Texas performance budgeting approach successful? The answer depends on how success is defined and who defines it. "Success" for Texas' performance budgeting effort could be defined as (1) compliance with the performance targets set in law, (2) improvements in performance by agencies over time, (3) satisfaction with the performance

measurement system by its key users, and/or (4) the accuracy of the performance information being reported.

Compliance

During the strategic planning process and the subsequent development of the agency's Legislative Appropriations Request, agencies go through an assortment of forecasting exercises to complete the portions of their submissions dealing with requested performance targets. Over the course of three to six months, numbers are refined and adjusted based on the most recent information available. Trends are closely examined, economic conditions are taken into account, and so on. In the planning process, meetings are held with the Governor's Office of Budget and Planning and the Legislative Budget Board to determine which measures will be included in the agency's strategic plan. Some yet-to-be determined subset of these measures will appear in the GAA. "Key" measures are the measures that will be tracked by the LBB and used to make determinations about the agency's performance in subsequent reports to the Legislature. The measures not included in the GAA, "non-key" measures, are to be maintained and used by the agency as well. All measures, key and non-key, are listed in the agency's strategic plan.

Many of the key performance measures established in Texas' PBBS are designed to monitor compliance with statutory provisions or to meet specific standards established by the Legislature. For example, in the case of regulatory agencies, grouped in Article VIII of the GAA, many measures indicate whether complaints, license renewals, or inspections have been performed within statutory time frames. Yet adherence to these standards, however stringent, does not necessarily produce better performance. In fact, many outcome measures are targeted at 100 percent compliance, which, if consistently met, would yield no improvement whatsoever.

The notion of agency compliance is also embodied in the LBB performance assessment, where agencies are scored on the basis of how many key measure targets, which are listed in the GAA, the agency has met or exceeded (within 5 percent). For example, an agency that reaches 94 percent of the targeted value on each of its performance measures would be given a score of zero percent, because it did not attain a reported performance within 5 percent of the target.

On the other hand, an agency that reaches 95 percent of the targeted values on each of its key performance measures would be given a score of 100 percent. With only 1 percentage point difference between two measures relative to their targets, one counts and the other does not. Thus, one agency could appear to have a perfect record of performance while another appears to have been derelict, with only a tiny margin of difference between the two agencies' actual performance.

Agency compliance is embedded in the Legislative Budget Board's performance assessment of each agency—how many key-measure targets are met or exceeded within 5 percent of the established target? Between 2000 and 2003, for example, there were over 2,600 key measures and for these, 69.5 percent of the targets were met or exceeded.⁹ While there are methodological problems with the usefulness of compliance as a measure of success, compliance is a key factor in the minds of legislative users.

Performance Improvement

Second, did performance actually improve over time?

Meeting a target is not the same as improving performance! In a series of 10,249 paired observations of performance measures over a 10-year period between 1993 and 2002, only 52 percent of the key performance measures reflected in the Texas state budget process showed improvement over time. While the general trend has been positive, the general lack of cause-and-effect relationships among the measures makes it difficult to determine where and why success or failure occurred.

During the past three biennia, two-thirds to three-quarters of all LBB recommendations for performance targets did not involve any change in performance between the targets during the next biennia (see Table 9.2). Put differently, the majority of LBB recommendations did not raise the bar from the first to the second year of the biennium.

Table 9.2: Legislative Budget Board Recommendations Containing No Change in Performance

Recommendations for the Biennium	Number of Performance Measures*	Recommendations with No Change within Biennium	Percent Having No Change within Biennium
2000–2001	1,956	1,276	65%
2002–2003	1,984	1,334	67%
2004–2005	2,013	1,426	71%

Source: Legislative Budget Board, Legislative Budget Estimates, 1999, 2001, and 2003.

* Excludes explanatory measures and missing targets.

Many LBB recommendations increased performance from the previous biennium, but in practice the LBB generally sets the mark for the next biennia at the better of the preceding two years.

Why the reluctance to increase targets? As previously mentioned, failure to perform at the level set forth in the GAA could result in negative consequences. Given the responsibility of all parties to defend their recommendations and requests, conservative estimates are more likely to prevail. The LBB analyst would have to defend a recommended increase because the agency would probably question the recommendation in committee hearings. The most likely answer to the question of why the reluctance to increase targets is that there are no institutional incentives for waging those battles, either on the legislative or administrative side.

The Texas performance budgeting system relies on four types of measures. Were there differences in performance improvement among the four types?

Although the Texas system applies common sets of measures to institutions of higher education, to the courts, and to regulatory agencies, very few measures are directly comparable to other indicators outside their functional area.

To evaluate whether performance improved or not, it is necessary to identify the types of performance measures that Texas agencies report and how they work. There are four types of measures in the Texas system.

First, *outcome measures* are developed for each agency's goals and objectives. These measures are typically designed as percentages, sometimes as percentage change or as a percentage of an overall objective. In general, better performance is indicated as the values of the measure increase; however, some outcome indicators are designed to show improvement when the values decrease. Outcome measures are reported annually, not quarterly as output and efficiency measures are.

Output measures constitute the second kind of measure. These measures are generally identified as the number of things produced by an agency. Examples include the number of clients served, complaints resolved, and licenses issued. Workload and production indicators are frequently mixed together in the state's output measures. In general, output measures show improvement as the reported values increase. In practice, exceptions to this rule are extremely rare.

Efficiency measures generally deal with cost per unit of output or time per unit of output. In some cases, measures that might seem to be likely efficiency or output measures are designated as outcome measures, such as administrative cost as a percentage of total cost or the amount of revenues collected by an agency. Quite often, efficiency measures combine unit and time in calculating cost, as in the case of monthly cost per client. Still, there is little standardization except in the area of regulatory agencies, where standard measures have been developed.¹⁰

Explanatory measures constitute the fourth and final type of measure used in the Texas system. These measures are sometimes included in the GAA to provide information on certain external factors, beyond the agency's control, such as the unemployment rate or the number of jurisdictional complaints received. Less than 5 percent of all the measures in the GAA are explanatory. They are not used to evaluate agency performance.

In a further analysis of the 10,249 paired observations, these four types of measures varied in the degree to which the Legislature required performance improvements over the 10-year period. For example, the performance targets for 49 percent of outcome measures were increased, 57 percent of output measures were increased from one year to the next, and 43 percent of efficiency measures were increased. Explanatory measures did not contain performance targets.

Table 9.3: Examples of Types of Measures

Type	Measure	Examples of Measures
Outcome (Results/Impact)	Identifies the actual impact or public benefit of an agency's actions (results or impact)	<ul style="list-style-type: none"> • Percentage of Clients Rehabilitated • Percentage of Entities in Compliance with Requirements • Percentage of Licensees with Validated Complaints
Output (Volume)	Counts the goods and services produced by an agency (volume or workload)	<ul style="list-style-type: none"> • Number of Clients Served • Number of Inspections Conducted • Number of License Applications Processed
Efficiency	Identifies the cost, unit cost, or productivity associated with a given outcome or output	<ul style="list-style-type: none"> • Average Cost Per Client Served • Average Cost Per Inspection • Average Time to Process License Applications
Explanatory/ Input	Shows the resources used to produce services and displays factors that affect agency performance	<ul style="list-style-type: none"> • Number of Clients Eligible for Services • Number of Entities Subject to Inspection or Regulation • Number of License Applications Received

Source: State Auditor's Office, *Guide to Performance Measure Management: 2000 Edition*.

User Satisfaction

Third, how satisfied were the users of the performance budgeting and measurement systems?

In March 1998, staff of the State Auditor's Office (SAO) conducted two surveys, with the assistance of staff from selected state agencies, House Appropriations, Senate Finance, the Governor's Office of Budget and Planning, and the LBB. The surveys were designed for two distinct sets of users of the Texas performance budgeting system. One survey was conducted to examine the perceptions of members of the legislative committees that use performance measures the most: the House Appropriations Committee and the Senate Finance Committee. The other survey was directed at state agency personnel and their perceptions of the performance measurement system. Survey results reveal very distinct differences in perspective between the two groups.

In general, legislative members were much more satisfied with the performance measurement system than agency personnel. This may not seem surprising, but the margin of difference is worth comment. Of the 17 members responding, 65 percent indicated that the system "always" or "almost always" improves accountability, which points again to the use of the performance system to maintain compliance with legislative intentions.¹¹ Similarly, 69 percent of respondents indicated that the quality and types of information have either "always" or "almost always" improved with performance-based budgeting. However, only 53 percent of respondents affirmed that there was an appropriate number of measures in the GAA and in the system.¹² One legislator responded with the comment, "We measure too many things."¹³ Another notes, "Many measures are beyond the control of agencies."¹⁴ Nevertheless, most of the members responding found the system to be informative and useful in the appropriations process.

By contrast, only 43 percent of 171 agency respondents indicated that the system "always" or "almost always" provided accountability to the Legislature and the public.¹⁵ Slightly more, 47 percent of 172 respondents indicated that performance measures were "always" or "almost always" used to manage agency performance, though large agencies were more likely than small agencies to provide such a response.¹⁶ Fewer still were enthusiastic about established performance measures as a management tool. Only 36 percent of respondents indicated that performance measures provided an early warning system for problems within the agency.¹⁷

Clearly, enthusiasm for performance measures differs markedly between policy makers and agency administrators. Representative Henry Cuellar cited the Texas State Auditor's Office in his dissertation when he noted, "There is a belief by some agencies that performance-based budgeting is used to punish, but not to reward."¹⁸ Given the historical record of reducing appropriations and the reluctance to appear generous during difficult budget

years, one wonders if such a belief might be justified. To date, no agency has ever been given authority to increase salaries under provisions of the GAA.

While interesting, these surveys leave many unanswered questions. Respondents provided a few examples of their use of performance measures in making decisions related to specific procedures, operations, or funding, but no general consensus is apparent about the way to use performance measures. Some legislators say that they should be used for rewards and penalties, others to prioritize appropriations, and others as instruments to make agencies "think through the process and put added emphasis on the things the Legislature deems important."¹⁹

When asked, members were able to cite specific examples of when they had applied performance measures in their decision making. In some cases, appropriations for specific agencies were mentioned, such as the Office of the Attorney General, the Department of Public Safety, the Texas Department of Health, and the Department of Protective and Regulatory Services.²⁰

In general, the survey of members of the Texas Legislature found that they were satisfied with the performance-based budgeting system, though many comments indicated that they would like to see some specific improvements to the measures in the areas of simplicity, relevance, savings, and improvement.²¹

Reporting Accuracy

Finally, how accurate are performance reporting systems?

In addition to statutory provisions requiring quarterly and annual reporting cycles, performance measures are also subject to audit by the State Auditor's Office (SAO). Agencies with appropriations in excess of \$10 million are also required to employ an internal auditor, who may also perform audits of performance measures. Moreover, criminal penalties may apply in cases where data have been falsified.

During the past decade, the SAO has audited performance measures in 200 state agencies. As the following excerpt from the SAO's *Guide to Performance Measure Management: 2000 Edition* suggests, audits result in one of five distinct findings:

Measures are designated as either "certified," "certified with qualification," "factors prevented certification," "inaccurate," or "not applicable." These categories are assigned based on a combination of the adequacy of the controls over a measure and the results of testing a sample of source documents. Following are explanations of the five certification categories:

- A measure is **Certified** if reported performance is accurate within plus or minus 5 percent and if it appears that controls to ensure accuracy

are in place for collecting and reporting performance data.

- A measure is **Certified With Qualification (CQ)** when reported performance appears accurate but the controls over data collection and reporting are not adequate to ensure continued accuracy. A measure is also certified with qualification when controls are strong, but source documentation is unavailable for testing. A measure is also certified with qualification if agency calculation of performance deviated from the measure definition but caused a less than 5 percent difference between the number reported to ABEST and the correct performance measure result. Findings may be issued for these measures if qualifications are significant.
- **Factors Prevented Certification (FPC)** is given if documentation is unavailable and controls are not adequate to ensure accuracy. Factors prevented certification is also given when there is a deviation from the measure definition and the auditor cannot determine the correct performance measure result. Findings are issued for these measures.
- A measure is **Inaccurate** when the actual performance is not within 5 percent of reported performance, or there is a greater than 5 percent error in the sample of documentation tested. A measure is also inaccurate if the agency's calculation deviated from the measure definition and caused a greater than 5 percent difference between the number reported to ABEST and the correct performance measure result. Findings are issued for these measures when a complex or system-wide problem exists.
- Certification for a measure is **Not Applicable** when performance is justifiably not reported for a given year. This category is rarely used. A measure is usually put into this category if it is new and information is not yet available for reporting.

Results of the performance measures audit are published in a public report to be used by the LBB, GOBP, and the Legislature.²²

Table 9.4 on page 448 summarizes the results of the performance measure audits conducted by the SAO since the inception of the performance-based budgeting system a decade ago.

As Figure 9.1 on page 449 illustrates, the control systems for an agency's reporting system for performance measures may be quite complex. Documentation may be as simple as paper files or as sophisticated as a large mainframe computer. Extremely large databases are involved in the reporting systems for large agencies and universities. Nevertheless, performance reporting may be a minuscule percentage of administrative costs in these institutions. On the other hand, very small agencies—such as the Board of

Table 9.4: Accuracy of Performance Measures in Audited Texas Agencies

Release Date	Report No.	Number of Agencies	Number of Measures	Percent Certified**	Number of Agencies at 100% Certified	Percent of Agencies at 100% Certified
2/28/96	96-052	22*	105	54	5	23%
7/31/96	96-071	20	105	68	8	40%
1/6/97	97-029	20*	94	82	7	35%
3/26/97	97-047	14	31	94	13	93%
8/8/97	97-077	26	165	69	9	35%
5/22/98	98-040	36	150	65	16	44%
5/31/00	00-030	11	43	54	4	36%
11/30/00	01-007	25*	105	65	7	28%
8/29/01	01-036	12*	119	53	2	17%
11/18/03	03-008	14*	68	47	1	7%

* Includes educational institutions.

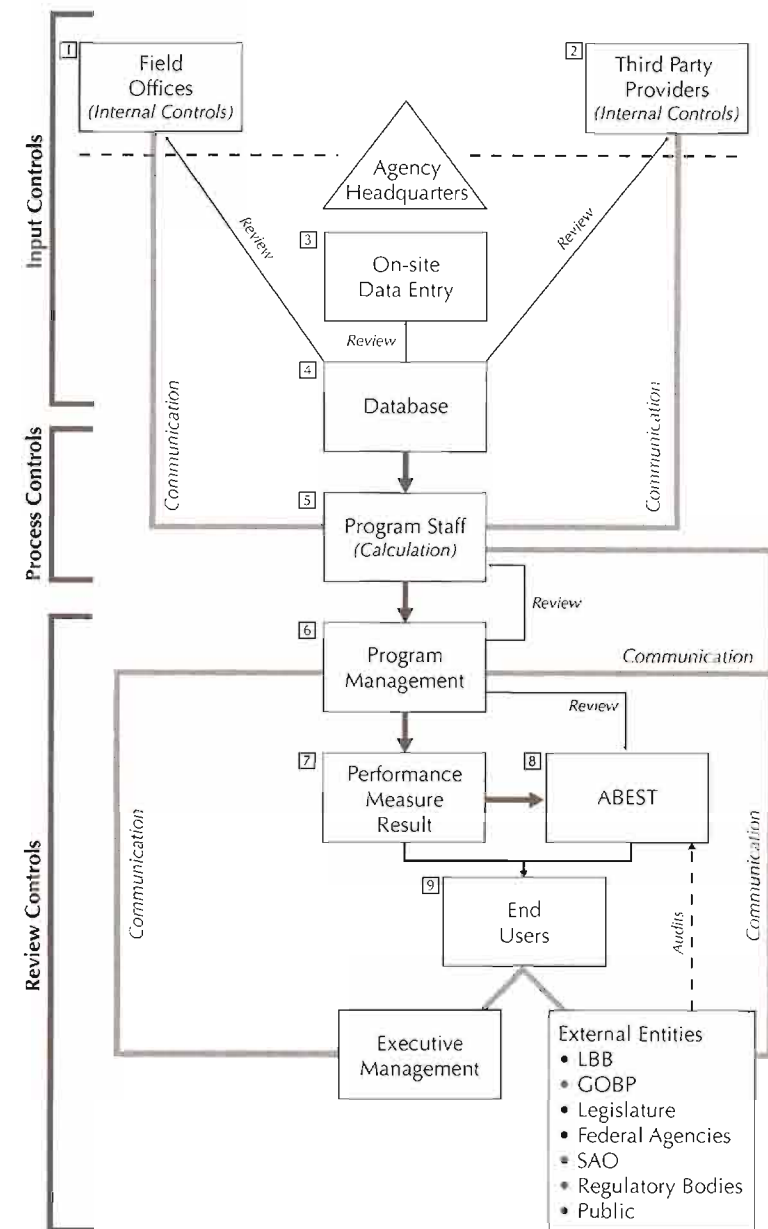
** Certified or "certified with qualification"

Source: Compiled from State Auditor's Office reports

Tax Professional Examiners, with a staff of only four people—may devote a higher percentage of administrative staff time working on their reports.

In a nearly perfect system, the percentage of measures being certified would hover around 100 percent. As Table 9.4 indicates, the percentage of agencies having that kind of record is low. The best results of the certification process were reported in the March 1997 report—of the agencies reviewed, 94 percent of their measures were certified as accurate, with 93 percent of the agencies having 100 percent of their measures fully certified. This particular report examined the courts' reporting accuracy. On the other hand, the low point in the certification results is the most recent report, where only 47 percent of the measures passed certification and only 7 percent of the agencies had 100 percent of their measures certified.

Considering the intensity and rigor with which the State Auditor's Office conducts its testing during the certification process, one might be more surprised by the high percentages of successful certification during the earlier years than the declines in certification in the more recent years (see Table 9.4). Still, there is evidence that the reliability of the performance measures reported to the LBB could be improved. Accuracy is only one of many caveats that one should entertain when reviewing agency performance.

Figure 9.1: Model of Performance Measurement System with Ideal Reporting Structure

Source: State Auditor's Office, Guide to Performance Measure Management: 2000 Edition, p. 50.

Explanation of Variance: Processes and Performance Drivers

According to a former state representative and Appropriations Committee member, studying internal processes is one of the fundamental steps in developing a performance-based budget system.²³ As fundamental as it might seem, documenting business processes is relatively uncommon among state agencies. With the notable exception of such legislative mandates as information technology projects, disaster recovery plans, rules reviews, and the agencies' Compact with Texans, legislators and agency administrators pay relatively little attention to documenting operational processes. The often repeated truism that continuing to do the same things should produce the same results rings especially true when reviewing agency explanations of variance, which commonly refer to the same causes—generally external factors.

Writing an acceptable explanation of variance is a rare art form. Patterns are sometimes hard to discern, but experienced LBB analysts can attest to the common features of the artfully crafted explanation of variance from agencies whose reported performance differs more than 5 percent from the established target for key measures. Agencies are required to provide an explanation of variance in quarterly and annual performance reports. In cases where the variance is negative, a plan of remedial action is also required. Given the general lack of detail available to reporting staff, and the urgent turnaround time required to request and obtain comments from agency managers, it is not surprising that most explanations of variance tend to be brief and plans of action are carefully non-committal.

During the two years that the author reviewed explanations as an LBB analyst, some patterns emerged. Typically, explanations of variance from Texas state agencies refer to external factors. Most often, agencies cite the state of the economy to explain fluctuations in client caseloads, the economic situation of regulated entities, or market prices for increased costs. While the economy is especially common in explaining variances, direct correlations are not typically presented. Associated plans of action are likewise very general and involve activities such as monitoring performance or, for a lucky few, the obvious benefits of filling staff vacancies in the affected program area. With the rare exception of staff vacancies, references to internal processes or procedures are largely absent from explanations of variance. Plans of action typically include such passive activities as monitoring and reviewing performance, with little description of specific proactive activities.

Another possible explanation for the brevity and vagueness of explanations of variance is that too much discussion might have the potential to

raise even more scrutiny. Agencies are not likely to say that performance declined because of poor management or bad decisions.²⁴

In earlier editions of the *Summary Assessment of Agency Performance*, LBB analysts would comment on the adequacy or inadequacy of an agency's explanations of variance as well as the timeliness of its performance report submission. In recent years, assessments have included the status of lawsuits as well as audit reports, but have dropped items concerning timeliness and explanation adequacy. Without more information about internal processes, analysts are not in a very good position to determine whether an explanation is adequate, but merely plausible.

If one starts with the premise that changing results generally means changing the process by which results are obtained, performance measures should help guide process redesign. Performance measures can be used as aids in monitoring the impact of changes in procedures. Without a linkage to processes and procedures, performance measures have limited operational value.

Although agencies are required to have disaster recovery plans, self-assessments associated with "sunset" reviews, and the required Compact with Texans, there is no general statutory requirement to which agencies must adhere regarding the documentation of internal processes. When one considers that the same three to six people at any particular agency are responsible for all the activities that make up the performance budgeting and monitoring system, it is not surprising that a rigorous analysis of internal processes takes a back seat to the immediate challenges of quarterly explanations of variance, preparation of the necessary reports to the Legislature, and responses to specific questions raised in hearings or by legislative and gubernatorial staff. Under these conditions, anecdotal comments about recent events, local office staffing, and the like have to suffice, but they are symptomatic of a lack of understanding about performance drivers.

Beyond Performance Measures: Understanding Processes

Despite its quantitative, highly empirical appearance, the implementation and management phases of the Texas performance-based budgeting system are difficult, filled most often with vagueness, anecdote, and perception. Information is often incomplete, irrelevant, or contradictory. The larger and more complex the agency's operations, the more challenging it is to make sense of what is driving performance. Performance drivers are sometimes counter-intuitive.

For example, one of the most controversial measures at the Texas Workforce Commission is the number of children on the waiting list. Budget ana-

lysts and policy makers have struggled with this measure for years, and it still plagues the members and agency staff to this day. The problem is that there is only enough funding to serve a relatively small fraction of eligible children, so the waiting list declines at times when it appears to be unlikely for those on the list to receive care. However, when additional funding is announced, and word reaches potential recipients, the waiting list actually increases as expectations of service rise. This continues until the funding is absorbed and the list stabilizes once again, frustrating policy makers, who had hoped to see a permanent decline.

This is why it is essential to step back and make time to outline and map out the internal processes, organizational and fiscal capacity, and the associated performance drivers. Without a deep understanding of performance drivers, explanations of variance do not contribute much to budgetary or policy discussions, or to performance management. State leaders are not oblivious to these concerns.

As early as 1995, the SAO began working on activity-based costing methods, and many agencies have worked on a pilot—among them, the Texas Workforce Commission.²⁵ These efforts bring structure to examining agency business processes and help fill the gap for entities that skipped the study of their internal processes. Such analyses are expensive and time consuming, but they can also be very enlightening for managers. Combined with performance measures, activity-based management techniques may prove to be a powerful combination. But those efforts will take time to mature.

Perhaps one of the best assessments of performance-based budgeting came from Jonathan Walters, reflecting on the biennial conference on performance measures in 2000 sponsored by the LBJ School of Public Affairs at the University of Texas at Austin:

Proponents of performance measurement in government need to reevaluate how they sell this stuff. In particular, they need to think hard about the message they want to send when they advocate for performance-based budgeting. Implicit in the discussions of the day is that data on performance is going to somehow make budget decisions easier: Because agency X is achieving certain results, then some budget decision or other will evince itself as eminently sensible—or flat-out obvious. In other words, information on performance will make budgeting a relatively cut-and-dried exercise. And so there was a good deal of earnest discussion about such classic performance-based budgeting questions as whether high-performing agencies should get more money and weak performers less, or vice versa.

The fact is, governments are never going to achieve anything like what performance measurement proponents view as textbook performance-based

budgeting. That doesn't mean state and local governments should stop trying to measure performance. If anything, they should escalate their efforts. But it's not so legislators can plug that data in some budgeting version of TurboTax to come up with a precise annual allocation based on last year's performance and next year's performance goals. It's so legislators can, if they choose (and that's still a big if), have better-informed debate and discussion come budgeting time.²⁶

There is much truth in these comments, but if performance measures are designed solely for the sake of discussions in budgeting, we have missed the point. Performance is what we should expect from government, and improvement is what we should require. To make this happen, we need to make sure that performance measures add value to operations and activities. Measures that do not add value should be discarded. Building solid linkages between performance measures and internal processes is essential. Only then is a performance measurement system worthwhile.

It is also the author's intention to make policy makers think about *improvement* as a standard for rewards and punishments instead of negotiated targets. Improvements are objective, easy to calculate, and fairly easy to comprehend. Most importantly, improved agency performance is what citizens deserve.

Lessons Learned

The Texas performance based budgeting system has evolved into a complex system for holding agencies accountable. While it still needs refinement to become a full-fledged performance management system, much of the infrastructure for such an effort has already been laid. Further refinement, in the area of process mapping and capacity analysis can contribute to better performance management.

Lesson 1: Define Success

Methods of rating agencies define success. Performance measurement systems should emphasize improving performance, not just meeting targets.

This problem is exacerbated by the use of targets that are bounded by natural limits. It is generally not possible to exceed 100 percent of a measure, though some agencies do exceed 100 percent when they catch up on backlogs of various kinds. Setting targets at 100 percent may seem like a

Lessons Learned

1. Define Success
2. Define Performance Drivers Before Adopting Measures
3. Focus on A Few Indicators
4. Measure Management's Effectiveness
5. Standardize Measures
6. Coordinate Measures Across Levels of Government Whenever Possible

good standard for accountability, but it is not the same thing as improvement. In the worst case, it can be a hindrance to improvement.

Texas' performance-based budgeting system is designed to align resources with performance. In that context, setting targets in the GAA is a practice that makes intuitive sense. The GAA becomes a kind of invoice for payments to agencies. Performance targets help keep track of those agencies that deliver and those that do not. For this purpose, the system delivers what its creators expect.

Scoring agency performance on target achievement alone shortchanges those agencies that make substantial improvements but are marginally under target. Only 25 to 30 percent of the performance targets set by the LBB changed from one year to the next within a biennium. While there is still much more to learn about the causes of variations in performance among agencies, the pattern of overall improvement across biennia is very suggestive. Perhaps better performance would be obtained if agencies were asked to make improvements more often.

Lesson 2: Define Performance Drivers Before Adopting Measures

It would be useful to define performance drivers for every measure before it is adopted. This practice would provide a basis for meaningful feedback to the Legislature and other oversight agencies about opportunities and obstacles to improvement. A well-developed logic model, of the kind Harry Hatry and others have recommended, would make measures far more useful for management and for policy makers. A concise flow chart detailing the causes and effects of the agency's processes would help discipline administrators and provide policy makers with knowledge that could help inform policy.

As it stands, too many agencies are caught off guard by changes in performance. Given the scramble to assemble quarterly reports, there is little time, if any, for careful analysis or reflection. While a month may seem like a long time to compile a report for a previous quarter, navigating a complex system of quality controls as outlined by the SAO takes time.

As Hatry has noted, "What gets measured gets attention."²⁷ Performance measures are visible, but they are only indicators, records of what has happened. Measures are not answers, but questions to be answered. Why? How? Those are the critical questions. A well-developed performance measure should be connected to a documented operational process that answers questions about what causes it to change. As in business, focusing on internal processes is more likely to lead to better performance than discussions about the state of the economy or the weather.

Lesson 3: Focus on A Few Indicators

On average, each agency has about 10 key measures; however, many of the larger agencies have more than 50 key measures. Together with non-key measures, some agency plans contain more than 100 performance measures. To the uninitiated, the idea that fewer performance measures are better than many may not be obvious. It is almost impossible to manage anything using 100 different measures. Cognitively and practically, that is unmanageable. A handful of good measures is more practical.

Almost all experts on performance measures agree that it is important to focus on a "vital few" measures. Moreover, such an approach engenders a kind of myopia where one cannot see the forest for the trees. Most importantly, it adversely affects time management. Some advocates of the balanced-scorecard approach use the term "dashboard" to describe the presentation of measures for busy executives. This idea has been incorporated in recent editions of the LBB's *Budget and Performance Assessments*, which show a graph or two of selected measures.

Of course, as a legislator noted in the SAO survey, one solution to the problem of being overburdened with attention items is to prioritize.²⁸ Unfortunately, in the LBB's performance assessment scheme, a measure for a \$500 million strategy has as much weight as that of a \$5 million strategy. Too much time may be devoted to answering questions about comparatively small programs, while much larger issues languish. In an over-measured environment, the attention given to small programs and issues robs executives of the time needed to manage large issues.

Another solution to the problem of too many measures is to aggregate measures, combining results from several areas into one or two larger aggregates. From a big-picture perspective, this can make a lot of sense.²⁹

Aggregate statistics are generally less susceptible to environmental disturbances than disaggregated statistics. And much of the variation in disaggregated statistics is meaningless noise, or what statisticians call “white noise.” Like static on the radio, over-measurement, by itself, generates its own noise. This makes it difficult to tell what is really happening.

As mentioned earlier, explanations are required for variations of more or less than 5 percent of the established target. Disaggregated measures produce more need for explanations than aggregated measures. Meaningless variations produce meaningless explanations, and the credibility of the system suffers under the weight of unnecessary verbiage and communication. In such a context, important variances and communications have to compete with less important messages.

For example, if a strategy has 10 equally distributed programs dealing with the same goal and five produce results that are 6 percent lower than the target and the other five have results 6 percent higher than the target, no explanation is needed for the aggregate result. But if the results are disaggregated, 10 program-specific explanations are required! Staff in performance reporting departments have a term for these unhappy excursions: “rabbit hunts.”

Rabbit hunts are costly in a number of ways. Given the time constraints associated with performance reporting, the more measures there are to explain, the less likely there will be quality explanations for any of them. As a result, audiences receive mere speculation about causes, barely more than plausible hypotheses. The economy may be cited as a cause for increases as well as decreases without any clear causal connection being demonstrated. Eventually, credibility suffers. Rigorous analysis and investigation takes time. If the number of measures artificially multiplies variances, the cost of analysis and investigation is only magnified.

In a well-tuned performance measurement system, static or noise is reduced. Performance is clearer and easier to identify with fewer, more focused measures. It is worth the time to map processes and analyze trends to identify the performance drivers that are the critical success factors management needs to monitor. Otherwise, performance measures and their explanations become work products, not management tools.

Lesson 4: Measure Management's Effectiveness

One of the biggest holes in the Texas performance measurement system is the lack of any measure for functions of indirect administration. Yet administrative activities are very similar across agencies.

In addition to compliance with a number of specific regulations concerning procurement, minority contracting, human resources, accounting,

and performance reporting, the executive functions can be measured using indicators similar to those sometimes employed to measure the performance of private sector chief executive officers. Of course, such measures would need to be tailored to a public sector environment, where stock prices or other private sector indicators are not relevant. Still, the agency's overall improvements across “direct” strategies can also be gauged, either in the LBB style or the ratios presented here. Regardless of the specific metrics to be used, in principle, it seems odd that anywhere from 5 to 30 percent of an agency's budget should go unaccounted for with respect to performance indicators.

As many in the private sector have noted, CEOs need to be held accountable for their institutions.³⁰ Summary assessments of agency performance are not enough. A standardized scorecard for administrative functions is needed. During the legislative session, agency executives answer to the Legislature regarding their agency's overall performance; however, lower-level program managers are generally asked to supply any specific information related to specific measures or to comment on a program's status. This is generally appropriate, but there are well-defined administrative functions that vary only marginally from one agency to another.

Some standardized performance measures can be taken right off the shelf. There are already some indicators of management performance, information systems, accounting systems, and so on. Most managers are aware of 360-degree appraisal and evaluation systems. Appropriations Committee members usually receive a briefing package on each agency that includes a kind of checklist of items about the health of the agency, including outstanding lawsuit liability, recent audit findings, compliance with audit recommendations, travel caps, hiring limitations, the use of historically underutilized businesses, and recent internal and state audit findings. It is a lot to digest.

One of the critical success factors in many organizations is executive management. Indeed, in many organizations, implementation of new procedures or changes in operations requires executive decisions, support, and follow-through. Seemingly mystical managerial buzzwords like *agility* and *leanness* can be operationalized in measures of administrative support functions. Decision turnaround time, delivery time for information requests, number and quality of policy decisions, reporting accuracy rates, adherence to planning timelines or progress on remedial action plans, procurement efficiencies, and the like can all be measured. These are just a few of the indirect administrative operations that lack measurement in the current system.

Most importantly, measuring management functions sends a message that everyone is on board and understands they are accountable, not just mid-level managers and line staff.

Lesson 5: Standardize Measures

Compiling and analyzing the massive volumes of data for this chapter has brought to light the difficulties facing anyone trying to assess the overall value of the system. Very few common denominators for comparing performance between agencies or the overall health of the agency exist.

One of the most difficult lessons for decision support staff to learn is to summarize information. The vast quantity of statistical information available to state policy makers is overwhelming. It is not realistic to expect part-time legislators to closely examine stacks of performance reports on more than 200 entities. This creates a dependence upon staff to sort through the deluge of information transmitted each reporting period.

What helps people understand performance measures is having a standard gauge of performance that is quickly understood and accessible, as well as valid and reliable. The Legislative Budget Board and the Governor's Office of Budget and Planning recognized this when they issued their memorandum on standard measures for licensing-related agencies.³¹ While every agency's mission is different, there are some common features among functional groups. For example, measures for universities are already more or less standardized.

The LBB's summary assessments of agency performance have striven to provide a consistent look and feel for legislators, much like the dashboard concept mentioned earlier.³² By comparison to the Legislative Budget Estimates, these assessments are easy to digest and consistent in the manner of scoring agency performance, though the LBB's target-based approach can mask improvements or declines in performance. Still, there is an awareness of the need for easily utilized information, as the state's recent emphasis on customer satisfaction indicates.

Outcome measures that use percentages are quickly understood, but most of the measures examined in this study are idiosyncratic, applying to the unique mission of a particular agency. With few exceptions, one cannot compare outcome measures from one agency to another agency, even in the case of customer satisfaction. In such a vacuum, objective measures lose objective meaning. Compared to what? Objective, quantifiable, and quantified benchmarks are essential.³³

Lesson 6: Coordinate Measures Across Levels of Government Whenever Possible

Finally, it is important to remember that the costs of maintaining competing systems of measurement are probably much higher than the cost of each system separately. Potential disagreements between systems generate problems and issues that would not exist had the systems been synchronized

and coordinated. This point goes well beyond mere turf issues. Competing figures must be reconciled.

Very few of the headaches associated with multi-layer measures occurred because of turf battles. In fact, the majority of issues raised were generally initiated by innocent third parties asking simple questions such as, "Why are these numbers different?"

If the questioner's predisposition or professional responsibility is to be skeptical, competing measures only heighten the skepticism. News reporters, legislative staff, auditors, and budget analysts are all trained to sniff out suspect information. Differences between state and federal reports are fair game. Questions about consistency will not go away. Such circumstances are a fact of life.

Good performance measures summarize events or encapsulate information in an easily digested manner. One indication of a less than perfect measure is the number of questions it raises about itself. Some measures are not worth the time they take to explain. All things equal, such measures should be discarded or reworked to be less mysterious. Decision times are not enhanced by superfluous questions. In general, the more questions raised in the decision-making process, the more likely the issue being considered will be lost in the fog of a busy legislative session. Policy makers tend to avoid making decisions about things that are unclear. Vague or obscure measures contribute to that inclination.

Likewise, competing measures—those that have the same name or a name similar enough to be mistaken for another measure—cause consternation and confusion among various audiences. This situation frequently occurs when an agency is reporting similar information to different audiences for different purposes, such as those measures reported to state and federal entities in different reporting cycles. To some audiences, the differences in reported figures require explanations.

During committee hearings, overly technical answers to such questions do little to stimulate understanding or inspire confidence. Answering that differences are the result of different definitions may not help, either. If the querulous party is agreeable to make the effort to walk through the calculations and timing processes, it is worth the time to educate them about the differences. The long-term payoff in trust and confidence is well worth the investment, but it is far better to avoid problems in the first place. Reconciling competing figures as a result of differing reporting cycles or similar naming conventions is generally not a productive use of time.

Endnotes

1. Maria Pilar Aristigueta and Joseph S. Wholey. *Managing for Results in State Government* (Westport, Conn.: Quorum Books, 1999), 109–110.
2. Proclamation by the Governor of the State of Texas, June 22, 2003, p. 1.
3. See General Appropriations Act, 75th Legislature, Article IX, Sec. 66; General Appropriations Act, 76th Legislature, Article IX, Sec. 6; General Appropriations Act, 77th Legislature, Article IX, Sec. 6.31.
4. State Auditor's Office, *Guide to Performance Measures* (Austin, Texas, 1995), p. A-4.
5. Christopher T. King, Robert E. McPherson, Daniel P. O'Shea, and Ying Tang. Improving Performance Measurement for Texas Workforce Development Boards: Phase One, Summary Report (Workforce Leadership of Texas, February 2002). A rider in the General Appropriations Act requires that any contract for services contain performance measures.
6. State Auditor's Office. Report 01-022, An Audit Report on the Local Workforce Boards (Austin, Texas, March 2001), p. 16.
7. General Appropriations Act, Article V-12, 76th Legislature.
8. General Appropriations Act, Article III-190, 77th Legislature.
9. Texas Legislative Budget Board, "Budget and Performance Assessments: State Agencies and Institutions, Fiscal Years 1999–2003."
10. Legislative Budget Board. Standard Measures for Licensing-Related Agencies, Memorandum from John Keel, Director, Legislative Budget Board, and Albert Hawkins, Director of the Governor's Office of Budget and Planning, to Licensing-Related Agencies (September 1, 1995).
11. State Auditor's Office. Member Perceptions of the Performance-Based Budgeting System (Austin, Texas, March 1998), p. 3.
12. Ibid, p. 15.
13. Ibid, p. 6.
14. Ibid, p. 3.
15. Office of the State Auditor, Performance-Based Budgeting Survey Results (Austin, Texas, October 1998), pp. 3, 25.
16. Ibid, p. 24.
17. Ibid, p. 29.
18. Enrique Roberto Cuellar. *A Comparative Analysis of Legislative Budget Oversight: Performance-Based Budgeting in the American States* (Doctoral dissertation. The University of Texas, 1998), p. 93.
19. State Auditor's Office. Member Perceptions of the Performance-Based Budgeting System (Austin, Texas, October 1998), p. 13.
20. Ibid.
21. Ibid, p. 8.
22. State Auditor's Office, *Guide to Performance Measure Management* (1999), p. 28.
23. Cuellar, p. 103.
24. This reflects the author's experience, which included four years of writing explanations for an agency.
25. State Auditor's Office, Report 95-139. *Guide to Cost-Based Decision Making* (August, 1995); General Appropriations Act, Article IX, Sec. 69, 75th Legislature.
26. Jonathan Walters. "Management and the Holy Grail," *Governing.com* (May 1, 2000).
27. Hatry, Harry P., *Performance Measurement: Getting Results*. Washington, D.C.: The Urban Institute Press, 1999. p. 55.
28. State Auditor's Office (October 1998), p. 19.

29. State Auditor's Office (March 2001), p.16.
30. Carroll Lachnit, "Big Rewards for Lackluster Execs," in *Workforce* (August 2002), Vol. 81 Issue 8, p. 20.
31. Standard Measures for Licensing-Related Agencies, Memorandum from John Keel, Director, Legislative Budget Board, and Albert Hawkins, Director of the Governor's Office of Budget and Planning, to Licensing-Related Agencies, September 1, 1995.
32. Legislative Budget Board, *Summary Assessment of Agency Performance: Fiscal Year 1996*, Submitted to the 75th Texas Legislature; *Budget and Performance Assessment, State Agencies and Institutions, Fiscal Years 1997–2001*, Submitted to the 77th Texas Legislature; *Budget and Performance Assessment, State Agencies and Institutions, Fiscal Years 1998–2002*, Submitted to the 78th Texas Legislature.
33. Despite much discussion about benchmarks in statewide planning, the Governor's Office never quantified the statewide benchmarks agencies are supposed to use in their planning process.

Bibliography

- Aristigueta, Maria Pilar, and Joseph S. Wholey. *Managing for Results in State Government*. Westport, Conn.: Quorum Books, 1999.
- Cuellar, Enrique Roberto. *A Comparative Analysis of Legislative Budget Oversight: Performance-Based Budgeting in the American States*. Doctoral dissertation. The University of Texas, 1998.
- Hatry, Harry P. *Performance Measurement: Getting Results*. Washington, D.C.: The Urban Institute Press, 1999.
- Juarez, Joseph C. Regional Administrator, U.S. Department of Labor, to Barbara Cigainero, Director of Workforce Development, Texas Workforce Commission, May 26, 2000, re: 20 CFR 663.320 and 20 CFR 663.600(d).
- King, Christopher T., Robert E. McPherson, Daniel P. O'Shea, and Ying Tang. *Improving Performance Measurement for Texas Workforce Development Boards: Phase One, Summary Report*, by Workforce Leadership of Texas, February 2002.
- Legislative Budget Board. *Budget and Performance Assessment, State Agencies and Institutions, Fiscal Years 1997–2001*, Submitted to the 77th Texas Legislature.
- . *Budget and Performance Assessment, State Agencies and Institutions, Fiscal Years 1998–2002*, Submitted to the 78th Texas Legislature.
- . *Legislative Budget Estimates for the 2004–2005 Biennium*, Submitted to the 78th Texas Legislature, January 2003.
- . *Legislative Budget Estimates for the 2002–2003 Biennium*, Submitted to the 77th Texas Legislature, January 2001.
- . *Legislative Budget Estimates for the 2000–2001 Biennium*, Submitted to the 76th Texas Legislature, January 1999.
- . *Legislative Budget Estimates for the 1998–1999 Biennium*, Submitted to the 75th Texas Legislature, January 1997.
- . *Legislative Budget Estimates for the 1996–1997 Biennium*, Submitted to the 74th Texas Legislature, January 1995.
- . *Legislative Budget Estimates for the 1994–1995 Biennium*, Submitted to the 73rd Texas Legislature, January 1993.
- . *Standard Measures for Licensing-Related Agencies*, Memorandum from John Keel, Director, Legislative Budget Board, and Albert Hawkins, Director of the Governor's Office of Budget and Planning, to Licensing-Related Agencies, September 1, 1995.
- . *Summary Assessment of Agency Performance: Fiscal Year 1996*, Submitted to the 75th Texas Legislature.
- State Auditor's Office. *Guide to Performance Measures*. Austin, Texas, 1995.

- . *Guide to Performance Measure Management: 2000 Edition*. Austin, Texas, December 1999.
- . *Member Perceptions of the Performance-Based Budgeting System*. Austin, Texas, October 1998.
- . *Performance-Based Budgeting Survey Results*. Austin, Texas, March 1998.
- . Report 03-008, *An Audit Report on Fiscal Year 2001 Performance Measures at 14 Entities* November 2002, November 18, 2002.
- . Report 01-036, *An Audit Report on Performance Measures at 12 State Entities for Fiscal Year 2001*, August 29, 2001.
- . Report 01-022, *An Audit Report on the Local Workforce Boards*, March 2001.
- . Report 01-007, *An Audit Report on Performance Measures at 25 State Agencies and Educational Institutions, Phase 14*, November 30, 2000.
- . Report 00-030, *An Audit Report on Performance Measures at 11 State Agencies, Phase 13 of the Performance Measures Reviews*, May 31, 2000.
- . Report 98-040, *An Audit Report on Performance Measures at 36 State Entities, Phase 12 of the Performance Measures Reviews*, May 20, 1998.
- . Report 97-077, *An Audit Report on Performance Measures at 26 State Agencies, Phase 11 of the Performance Measures Reviews*, August 8, 1997.
- . Report 97-047, *A Review of Financial Controls and Performance Measures at the Courts of Appeals and Other Judicial Branch Agencies*, March 26, 1997.
- . Report 97-029, *An Audit Report on the Performance Measures at 13 State Agencies and 7 Educational Institutions, Phase 10*, January 6, 1997.
- . Report 96-071, *An Audit Report on Performance Measures at 20 State Agencies and 1 Educational Institution, Phase 9*, July 31, 1996.
- . Report 96-052, *An Audit Report on Performance Measures at 21 State Agencies and 1 Educational Institution, Phase 8*, February 28, 1996.
- . Report 95-055, *An Audit Report on Performance Measures at 18 State Agencies, Phase 6*, February 15, 1995.
- . Report 95-139, *Guide to Cost-Based Decision Making*, August 1995.
- Walters, Jonathan. "Management and the Holy Grail," *Governing.com*: May 1, 2000. (<http://governing.com/view/vu50100.htm>). Accessed April 2, 2003.