

PN-ABJ-429

7384

**MEASURING PROGRAM PERFORMANCE
FOR FEDERAL AGENCIES**

ISSUES AND OPTIONS FOR PERFORMANCE INDICATORS

Gerald M. Britan, Ph.D.
Agency for International Development
April 1991

Prepared for the General Government Division,
U.S. General Accounting Office

BEST AVAILABLE COPY

PREFACE

This paper was prepared during the course of a four month "developmental assignment" with the Federal Management Issues Area of the GAO's General Government Division, under the auspices of the U.S. Office of Personnel Management's Executive Potential Program.

I very much appreciated the opportunity to work with and learn from my temporary GAO colleagues that the Executive Potential Program provided. In particular, I want to express my thanks to David Mathiasen and Benjamin Nelson for facilitating my entry into the GAO and to Assistant Comptroller General, Richard Fogel; Federal Management Issues Director, William Gadsby; and Federal Management Issues Assistant Director, Earl Walters for their continuing help and support. I also want to express my appreciation to my superiors at the Agency for International Development, particularly Annette Binnendijk, Janet Ballantyne, and John Eriksson, for making this opportunity possible and for their continuing support as I simultaneously juggled two sets of responsibilities. Above all, I hope that my suggestions prove useful both to GAO and A.I.D. in our mutual efforts to manage government programs more strategically for results.

Ultimately, of course, this paper does not reflect official positions of either the Agency for International Development or the U.S. General Accounting Office, but solely the views of the author.

SUMMARY

This paper reviews some of the key issues and options involved in measuring *program performance* in federal agencies and assesses their implications for *program performance indicators* efforts by the GAO and Congress.

After first defining "program performance," the paper considers different *levels of performance* in relation to organizational structure and hierarchies of objectives, articulating the use of *objective trees* as analytical tools. This provides a basis for identifying *types of indicators for different levels of performance*, including *direct indicators* of results (inputs, outputs, outcomes, impact, and significance) and *relational indicators* that measure "how well" results are being achieved (efficiency, effectiveness, relevance, and sustainability). Other key elements in a *performance indicators typology* are also discussed, including *performance indicators for different kinds of programs* (service delivery, regulatory/oversight, intergovernmental, grants, and defense) and *different ways of measuring program performance* (direct measures, indirect measures, intermediate indicators, leading indicators, quantitative and qualitative measures, and measurement scope).

Performance indicator options are then considered in relationship to *management needs at different organizational levels*. While lower level managers generally require information on operational performance (inputs, outputs, and outcomes), senior managers are primarily concerned with program impact and significance, relying on summary statistics and "management by exception" in monitoring routine operations. Effective management by top executives requires not only clear "vision" and "values," and but also a clear delineation of *strategic objectives* that both define core programs and suggest how the performance of these programs should be measured.

Few government agencies appear to have well defined performance objectives at this strategic level. Congress could play an important role by requiring such objectives and by working with top agency executives to clarify and confirm them. In so doing, Congress would, essentially, be defining *performance "contracts"* for which such agencies could then be held accountable. This would provide a better basis for *managing for results*, reducing the need for "micro-management," while also providing Congress with more relevant and targeted program and policy information.

Such a *performance management* focus also has important implications for the GAO. Nearly all of the GAO's General Management Reviews have identified substantial agency performance information deficiencies. By helping to delineate clearer program objectives and performance criteria, GAO could define a continuing *strategic framework* better linking its financial audit, evaluation, and GMR activities in an agency over time.

INTRODUCTION

Although there are major differences between managing public agencies and private firms, there are also fundamental similarities. To manage effectively, both public and private sector executives need some basis for assessing their organization's performance. For private firms, mobilizing private resources to earn private profits, the primary performance indicator--the financial bottom-line--is relatively straight-forward. For public agencies, utilizing public resources to produce a broad array of public goods and services, measuring program performance tends to be more complicated and multi-dimensional. However, it is just this complexity and ambiguity--combined with the public trust that public service embodies--that makes the development of appropriate program performance indicators so essential.

In recent testimony before the Senate Committee on Governmental Affairs, the Comptroller General strongly supported the concept of performance indicators for federal agencies. In subsequent meetings with GAO representatives, members of the Committee's minority staff requested GAO's assistance in developing a strategy to promote the use of such performance indicators, particularly as specified in S. 3154 (1990) and S. 1 (1991), bills introduced by Senator Roth. Based on these meetings, literature reviews, internal GAO discussions, and consultations with OMB and others, GAO proposed a four part plan of action to support the Committee's work and to provide a clearer framework for pursuing the GAO's own Federal Sector Management objective to develop better measures of agency performance.

This "issues" paper and accompanying "indicators typology" represent the first two steps in GAO's proposed strategy. The paper summarizes key issues in measuring program performance in federal agencies, develops a typology of performance measures, and assesses the utility of various performance indicator alternatives for major planning, management, and organizational functions. Further proposed steps, to be implemented later this spring or summer, include an "assessment of user (primarily Congressional) interests and needs" and the implementation of "pilot performance indicator demonstrations" with selected agencies.

WHAT IS "PROGRAM PERFORMANCE?"

Before we can delineate useful program performance indicators, we must first consider what kinds of performance we want to measure, for what programs, for what purposes, and for whom. Different programs, different objectives, different managers, and different audiences all require different kinds of performance information. Measuring program performance is, in other words, closely linked to processes of strategic planning, the clarification of organizational goals and objectives, the character of decision-making needs, and the needs of managers for information.

No single set of program performance indicators can satisfy every manager's information needs equally and simultaneously. Nearly every public agency pursues many different objectives, in different substantive areas, and at various organizational levels. Performance in achieving any of these objectives can be assessed in terms of a variety of not always fully compatible efficiency, effectiveness, outcome, and impact measures. And, the usefulness and significance of these measures varies greatly for different levels of management, in different

organizational settings, in coping with different organizational "environments," and in addressing different action opportunities.

Measuring program performance, in other words, involves much more than simply assuring public accountability for the legitimate use of public resources. Indeed, measures of program performance must usually go far beyond even the much more difficult questions of whether these resources are being managed efficiently and effectively. To be useful to top agency decision makers, program performance indicators must also provide a basis for assessing whether programs are achieving their intended results, whether these results are having broader impacts in alleviating the policy problems for which programs were created, and which among a variety of program options represent the best policy decisions and the best use of taxpayer dollars.

HIERARCHIES OF OBJECTIVES AND LEVELS OF PERFORMANCE

Just as organizations are structured hierarchically to manage people and resources, organizational objectives also tend to be structured hierarchically, reflecting the cause and effect logic necessary to achieve broader program goals. While organizational structure and hierarchies of objectives need not be entirely congruent, results achieved at lower organizational levels are often themselves inputs towards the achievement of higher level organizational purposes and goals. Not surprisingly, the kind of performance information needed, the type of performance being assessed, and the character of useful performance indicators also varies dramatically for different objectives at different organizational levels.

Sometimes the relationship between organizational structure and program objectives is primarily one of aggregation across space, with results in particular program locations simply being added to yield regional objectives and achievements. A state school milk program, for example, might seek to distribute a targetted number of milk servings statewide each week. At the same time, milk programs in individual school districts would each seek to fulfill their own smaller distribution targets. Overall state-wide performance would simply involve the summing of milk distribution figures across districts. Of course (as we will discuss in more detail later), different management levels would likely still have different performance information needs: State-level program executives, for example, would almost certainly require information about performance variations (and the reasons for them) across districts, but would be far less interested in performance variations across individual schools, which would be critically important to district program managers.

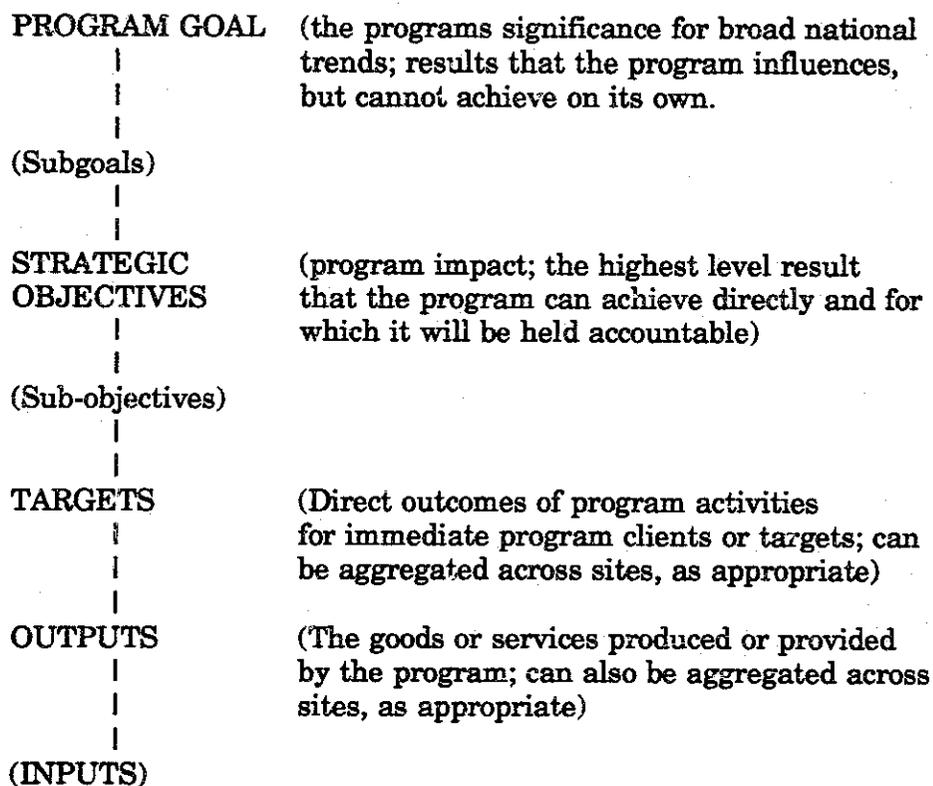
In addition to aggregation over space, most programs also embody a logical hierarchy of objectives that is reflected in organizational structure. The key objective for a local polio clinic, for example, may simply be to inoculate a certain number of pre-school children each year. The district polio immunization office, on the other hand, may be seeking to inoculate a certain percentage of a regional target population, the national polio prevention program may be trying to reduce the incidence of polio as a disease, while the parent "Child Health Agency" strives to reduce overall child mortality and contain the costs of curative child health services.

From this perspective, activities conducted and results achieved at lower organizational levels are necessary, but not necessarily sufficient inputs to achieving higher level goals. In

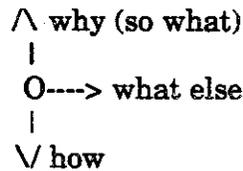
addition to clinic-based immunization, for example, the district polio office might also conduct a wide range of other activities--public education campaigns, community health outreach, school vaccination teams, and the like--all aimed at increasing the percentage of the target population that is immunized. Similarly, the National Polio program may encompass research, vaccine production and distribution, physician training, and a host of other initiatives to help achieve its broader disease reduction goal. At a still higher organizational level, the Child Health Agency might sponsor a variety of other immunization, emergency health, education, research, nutrition, and related activities, all aimed at contributing to reduced child mortality.

Spatially aggregated and hierarchically linked program objectives often exist side-by-side. Our national polio program may, for example, also have a more focused objective of innoculating so many children nation-wide--the sum of innoculations by individual clinics. Conversely, individual school district milk programs may be hierarchically structured components in more comprehensive statewide student health efforts encompassing a variety of health and nutrition activities.

Either way, the relationship among program objectives (and the need for performance information) can often be clarified by depicting performance objectives in a hierachical "objective tree." Such an "objective tree" begins with overarching program goals, lists the "strategic objectives" that contribute to achieving these goals, the lower level "outcomes" through which these objectives are achieved, down to specific "outputs" that result from discrete program activities. The objective tree, in other words, provides a graphic depiction of the overall program logic:



The logic of an objective tree is straight-forward: Reading down, we are asking the question, "how:" what did we need to do in order to achieve this result. Reading across we are asking "what else"--what additional things need to be done to achieve the next level objective. Reading up we are asking "so what"--what is the significance of our accomplishments; what will we achieve if we accomplish everything at this level. This logic can also be depicted graphically:



While all of the lower level results that are necessary and sufficient conditions for achieving higher level objectives should be included in the tree, not all need to fall within the purview of the program in question. Some may represent "assumptions" about conditions in the wider environment (such as the availability of transport for milk) or results that are expected from other programs (such as continued funding for general purpose school clinics). Care should be taken, however, to assure that assumptions are valid or that complementary results are in fact likely to occur.

A partial objective tree depicting the (somewhat simplified) logic of our hypothetical "Child Health Program" is depicted on the next page.

REDUCING CHILD MORTALITY

**REDUCING
INCIDENCE
OF PREVENT-
ABLE DISEASE**

**IMPROVING
NUTRITION**

**IMPROVING FAMILY
HEALTH PRACTICES**

**EXPANDING
AVAILABLE
CARE**

**INCREASING
POLIO
IMMUNIZATIONS**

**INREASING
OTHER
IMMUNIZATIONS**

**INCREASING
USE OF ORAL
REHYDRATION**

**EXPAND SCHOOL
IMMUNIZATIONS**

**INCREASE
AVAILABILITY
OF VACCINES**

**IMPROVE
PUBLIC
AWARENESS**

**INCREASE
SUPPLIES**

**IMPROVE
TRAINING**

**EXPAND
FACILITIES**

**PROVIDE
FUNDS**

**ENFORCE
REGULATIONS**

**PROVIDE
VACCINE**

total funding,
operating expenses,
program funding (for programs that provide goods,
services, or grants to third parties),
total person years,
administrative person years,
operational (and third party) person years,
materiel and supplies,
etc.

Output Indicators: Measure the quantity (and sometimes also quality) of goods and services that have been created or provided through the use of inputs. Depending on the kind of program this can include

clients vaccinated (by a health program),
farmers visited (by an extension program)
grants awarded (by a research program)
miles of roads built (by a highway program)
weapons procured (by a military acquisition program)
tax returns processed (by a tax program)
claims processed (by an insurance program),
etc.

Outcome Indicators: Measures the quantity (and sometimes quality) of direct results that have been achieved through the provision of program goods and services. Depending on the kind of program this can include:

illnesses prevented (through health vaccinations),
farm practices changed (through extension visits),
research publications or patents awarded (by grant
recipients),
vehicle usage (of roads built),
weapons available (through procurements),
taxes collected (through returns processed),
funds distributed (through processed claims),
etc.

Impact Indicators: Measures the degree to which wider program objectives are being achieved through the direct outcomes of program activities. Depending on the kind of program this could include:

reduced mortality/reduced curative health costs
(through illnesses prevented and other program outcomes),
increased agricultural productivity (through improved
farming practices and other program outcomes),
reduced traffic congestion (through increased usage of
new highways,
increased industrial productivity (through research
findings and other program outcomes),

increased military readiness (through more and better weapons availability),
increased taxpayer compliance/few taxpayer complaints (through improved tax collection),
improved client satisfaction/less financial hardship (through better insurance claims distribution),
etc.

Significance Indicators: Measures (where appropriate) trends with respect to the wider policy problems ("goals") which program impacts are expected to influence. Depending on the program, this could include:

improved national health statistics (through improved health system performance),
increased farm profits/reduced food costs (through farming improvements),
reduced transportation costs/expanded economic development (through highway improvements),
improved economic growth/enhanced consumer well-being (through the application of new technology),
improved deterrence/foreign policy success (through improved military readiness),
increased tax funds available (through fairer and more efficient tax processing procedures),
improved economic recovery (through more efficient insurance system),
etc.

Relational Indicators:

These measure *how well* results at one level of the objective tree have been translated into results at the next level. This includes:

Efficiency Indicators: In general, efficiency indicators measure the ratio of inputs needed per unit of output produced. This could include:

physical inputs/unit outputs
dollars/unit output
labor/unit output, etc.

From the perspective of performance assessment, accountability indicators (the central focus of much program and financial auditing) can be considered a subset of efficiency indicators. They measure the extent to which program resources are appropriately available for and applied to the activities (creation of outputs) for which they were targetted.

Effectiveness Indicators: In general, effectiveness indicators measure the ratio of program outputs (or the resources used to produce these outputs) per unit of program outcome. This could include:

number of vaccinations (or cost)/illness prevented,
number of farmers visited/farm practices changed,
number grants/research publications,
miles of road built/vehicle usage,
number (or cost) of procurements/weapons (or
firepower) obtained,
returns processed (number or time)/taxes collected,
claims processed (number or time)/funds distributed, etc.

Relevance Indicators: Where appropriate, measure the relationship of program outcome to program impact, either quantitatively (unit of program outcome/unit of impact) or qualitatively (degree to which program outcomes affect program impacts). This could include:

illnesses prevented/mortality rate changes,
farm practices changed/increased agricultural
productivity,
patents granted/increased industrial productivity,
new road usage/decreased traffic congestion,
improved weapons/increased military deterrence,
improved return processing/increased tax compliance,
improved disaster funds distribution/enhanced economic
recovery, etc.

Sustainability Indicators: Where appropriate, measures the persistence of program benefits over time, particularly after program funding ends. This could include, for example:

disease incidence trends after external funding for vaccination programs
ends,
persistence of changed farming practices after
extension visits are completed,
maintainance and use of roads after highway
construction ends,
or, more generally, the persistence of institutions
(programs, organizations, relationships, etc.) created to deliver program
benefits.

The Issue of Scope:

Inputs, outputs, outcomes, impact, and even significance can all be measured for discrete activities, for categories of activities, or for all program activities. Results can also be reported for individual sites or aggregated across sites. However, at higher management levels the emphasis is usually on program impact in achieving broader objectives, with senior managers relying primarily on aggregated and summarized data on program inputs, outputs, and outcomes across discrete activities and sites. More detailed or site specific information would usually be required only if summary data indicated problems that senior managers needed to address. A more detailed discussion of this "management by exception" and the "nested" performance information systems that it implies is provided later.

PERFORMANCE INDICATORS FOR DIFFERENT KINDS OF PROGRAMS

The kind of program being considered--particularly the kind of goods or services being created or delivered--has a major bearing on the choice of performance indicators. In this regard, most federal programs fall into one of the following categories:

Service Delivery Programs: Programs, such as those administered by the Veteran's Administration, the Postal Service, the Bureau of Indian Affairs, the National Park Service, etc. that provide direct services to individual clients. For such programs, *outputs* measure the actual delivery of services or products to clients; *outcomes* measure the use of services, the perception of services, or the effect of services on clients; *impacts* measure wider changes that programs were intended to produce in the characteristics of target populations, communities, or organizations.

Regulatory/Oversight Programs: Programs, such as those administered by the Environmental Protection Agency, the Nuclear Regulatory Agency, the Food and Drug Administration, or the Federal Trade Administration, that develop, promulgate, and enforce rules and procedures governing the behavior and operations of individuals, firms, organizations, and communities. For such programs, *outputs* measure the number or kinds of rules, regulations, or procedures developed or enforcement actions taken; *outcomes* measure changes in behavior or operations, the degree of compliance with regulations, or the number of successful enforcement actions; *impact* measures the achievement of broader strategic objectives that regulatory oversight is intended to yield, such as reduced environmental pollution, improved nuclear safety, faster availability of more effective drugs, increased economic competition, etc.

Intergovernmental Programs: Programs, such as those administered by the Food and Nutrition Service, the Agricultural Extension Service, the Community Development Administration, the Department of Labor (Job Training Partnership Act), or the Department of Education (State and School District Block Grants) that involve mixed oversight, review, management, and funding by different levels of government (and with federal resources usually allocated through formula funding). For such programs, *outputs* measure the extent to which resources have been delivered to targetted operating entities; *outcomes* measure the extent to which resources are being appropriately applied to targetted activities (and, at the local level, the quality and quantity of goods and services actually delivered to clients); *impact* encompasses improvements in the quality and quantity of goods and services being provided by targetted entities (and the extent to which wider changes that programs were intended to produce in the characteristics of target populations, communities, or organizations are being achieved).

Grant Programs: Programs, such as those administered by the National Science Foundation, the National Institutes of Health, the Office of Naval Research, the Department of Education, and the Department of Housing and Urban Development that provide funding for specific activities or projects in response to applications or proposals by individuals or institutions. For such programs, *outputs* measure the delivery of resources to appropriate grantees and the quality and quantity of activities undertaken by grantees using those resources; *outcomes* measure the direct results of grant-funded activities (publications and patents from research conducted,

education completed, housing units built, etc.); *impact* measures the significance of these results for wider program objectives (improved weapons technologies, improved health or economic competitiveness, expanded educational participation, or improved availability of urban housing, etc.)

Defense Programs, the single largest component of the federal budget, include grant, intergovernmental, and regulatory/oversight programs, but primarily involve the delivery of a particular kind of service: military preparedness. Depending on the specific defense program being considered, performance measures can vary widely. In general, *outputs* will focus on the quality and quantity of defense goods and services obtained (number of people trained, weapons systems produced, communications capabilities in place; *outcomes* will encompass how well these goods and services perform (power and reliability of weapons, capabilities of soldiers, performance of communications; *impacts* measure the achievement of wider strategic objectives (such as increased military readiness or an improved ability to project military power); while *significance* encompasses progress in terms of broader national objectives (deterrence achieved or foreign policy goals supported).

DIFFERENT WAYS OF MEASURING PERFORMANCE

In addition to measuring different "levels" of performance for different "kinds" of programs, indicators can also be differentiated in terms of how performance is measured. This includes:

Direct measures: As the name implies, direct measures precisely correspond to and directly embody expected program results at any particular performance level. Such measures might include quantities of goods delivered (output), counts of clients served (output), levels of weapons accuracy (outcome), changes in consumer behavior (outcome), decreases in infant mortality (impact), etc.

Indirect measures: If direct measures are too difficult, costly, or inconvenient to obtain, performance may be measured indirectly, through indicators that correspond less precisely to the performance we are seeking. This might involve, for example, using lower farmgate prices as an indirect indicator of increased agricultural productivity, using declining taxi-cab or freight tariffs as an indicator of decreased traffic congestion, or using reduced consumer complaints as an indicator of improved tax processing. Indirect measures can be relatively straight-forward (e.g., declining insurance claims as an indicator of safer car designs) or they can be quite ingenious and creative--one major museum, for example, measured "tile wear" to indirectly assess the relative popularity of its exhibits (cf: Campbell et al., 1966).

Intermediate indicators: Sometimes we need information on program results before final performance data is available. At other times, we may need to know whether our program is on track, even though final results have not yet been achieved. In either case, we could turn to "intermediate indicators" to provide an early assessment of performance. We might, for example, measure fertilizer purchases as a preliminary indicator of changed farming practices or increased nutritional knowledge as an indicator of changed eating practices.

As the term implies, "intermediate indicators" measure "intermediate" results, intervening steps towards achieving program purposes. Generally they represent changes that are associated with the performance we seek, but for which information can be obtained earlier. Sometimes, intermediate indicators represent preliminary links in a causal chain, and their utility depends greatly on the validity of the policy "theory" linking them to final results. Occasionally, intermediate indicators reflect results from initial or selected program sites, and their usefulness depends on the extent to which they in fact prove representative.

Leading indicators: Sometimes "leading indicators" (or indices of indicators) can be identified that have well established or statistically validated links with longer term results. While similar to "intermediate indicators," such "leading indicators" (for example, the index of leading economic indicators) generally have multiple applications and a track record of reliability.

Qualitative and Quantitative Measures: While often a consideration, the issue of quantification is, in itself, rarely a major problem in assessing program performance. There are always trade-offs to be made between the precision, cost, reliability, and validity of information, but more quantification is not necessarily better. The real issue is how precise our information needs to be in order to be useful. It is almost always better to have more valid and cheaper measures that are less precise than to have more precise measures that cost too much or are of dubious validity.

In principle, qualitative measurement refers to purely narrative descriptions of program activities, operations or results. Quantitative measures range from simple categorizations (nominal measures), to rank orderings of categories, to equal interval scales, to ratio scales (with a real zero point). For some performance assessment purposes ratio and interval measures, which provide the most scope for analysis, are easily available. In other cases, however, frequency counts in rank ordered or nominal categories are more than sufficient.

Measurement Scope: Performance indicators sometimes directly measure results for an entire target population of individuals or organizations through available administrative records, observations, or census surveys. Often, on the other hand, the scope of measurement is limited to a *sample* of targets or sites. When samples are used, an additional technical issue is introduced: how reliably can we statistically infer overall program performance based on this sample? Sometimes, performance is measured only in one or a small number of program settings. While such *case-studies* can provide quite useful information on how programs work, why they don't work as expected, and how they can be improved, one must be extremely careful not to presume that results from one site necessarily represent program performance overall. Indeed, such case-studies are usually conducted in the context of "linking studies" (discussed below) rather than as a replacement for broader performance data.

Linking studies: Often managers need to learn more about the causal linkages between program outputs, outcomes, and impacts. This is particularly true when indicators show that although planned program outputs are being delivered, broader program purposes are still not being achieved. Something clearly is missing: either our program logic is faulty, our assumptions are invalid, or some necessary "input" is

lacking. Linking studies--often including detailed case studies--are, in essence, "formative" evaluations, which examine how programs work and what is needed to make them work better. As such they often provide important feedbacks to program redesign and higher level policy debates.

PERFORMANCE INDICATORS AND PROGRAM MANAGEMENT

All managers--operational administrators, middle-level executives, and top policy-makers--need information on how their programs are performing as a basis for decision-making. However, more senior and more junior managers have rather different performance information needs. This is not only because of the wider scope of senior management responsibilities, but, more importantly, because senior managers are responsible for higher levels of program performance. While operational managers are primarily concerned with the tactics of program implementation, senior executives are primarily concerned with the strategic implications of program and policy alternatives.

Within the same organization, management information needs do, of course, overlap. All performance information, after all, ultimately derives from the same set of activities, reflecting the same core of inputs and outputs. And most performance information will be collected "on the ground," at operational levels. To meet the differing information needs of more senior and junior managers, performance indicators (like organizations and organizational objectives) usually need to be hierarchically structured. This can be accomplished through "nested networks" of partially overlapping performance information systems, in which lower-level performance indicators are summarized or selectively "skimmed" as a basis for higher level information.

Any particular performance information system will necessarily reflect the particular goals, structures, and activities of the organization it has been designed to serve. The following sections therefore provide only a general summary of the differing character of performance indicators and information needs at different organizational levels:

Operational managers:

Operational managers are concerned primarily with the tactics of implementing a particular program component. Within their relatively narrow areas of responsibility, they need to know that necessary inputs (human and financial resources) are in fact available and that planned outputs (goods and services) are in fact being produced. They need to know that resources are being used legitimately, that financial accountability is being assured, and that inputs are being efficiently transformed into outputs. Operational managers also need at least some information on program outcomes and impact to demonstrate that program outputs are in fact being delivered and achieving intended results, but this is primarily for the purpose of formative feedback, to fine-tune program implementation, and to motivate performance. Operational managers would not, however, normally be expected to assess the underlying logic of the program they are implementing or the policy theory on which it is based. Nor would they be expected to consider the relative merits of program or policy alternatives.

Middle managers:

Middle managers are usually responsible for implementing an entire program or a major program component--a range of activities, often across a number of sites, aimed at achieving some larger strategic objective. Middle managers generally play only a supporting role in defining these strategic objectives, choosing among them, or articulating the underlying policies from which these objectives derive. They are, however, very concerned with managing their program better, with selecting and fine-tuning program activities to improve program outcomes, and with enhancing their program's broader impact.

To achieve these aims, middle managers do need summary information on program inputs and outputs and, particularly, comparative information on the efficiency of implementation across sites so that they can identify emerging problems for more intensive "management by exception." Even more importantly, middle managers need comparative information on the achievement of program outcomes. This provides a basis for judging comparative effectiveness (across sites or activities), for deciding which program activities should be expanded and which contracted, and for identifying any gaps or inadequacies in the program strategy that need to be filled. Middle managers will also want at least some information on program impact to validate the significance of their program strategy and assure that its policy underpinnings remain valid.

Senior Executives:

Senior executives are primarily concerned with strategic management. They must articulate their Agency's mission (its vision and values), clarify the policy presumptions that underlie this mission, and identify the strategic objectives and program initiatives through which this mission will be achieved. Further, they must accomplish this in the context of a dynamic external environment--a continually shifting mix of Congressional oversight, interest group pressures, bureaucratic competition, and executive politics.

Senior executives do need summary information on program efficiency and effectiveness. They need to know that planned goods and services are being delivered and that desired outcomes are occurring. More pointedly, they need to be sure that any emerging implementation or outcome problems are being identified and addressed at appropriate management levels and they need to be able to reassure both clients and critics of this fact.

To manage strategically, however, senior executives also need comparative information on program outcomes and impact. This provides a basis for judging the comparative significance of program alternatives, for deciding which programs should be expanded and which contracted, and for identifying gaps and inadequacies in program strategies that need to be filled in order to achieve policy goals. When "significance" is measured in relation to the cost of program inputs, this information can also feed back directly into a performance based budgeting system.

Beyond this, senior executives also need information on the wider national trends to which their programs are expected to contribute, at least in part. If these trends are positive, that provides further evidence that their agency's mission is being fulfilled. If, however, relevant trends are negative or weak, that raises questions about possible gaps in Agency strategies or inadequacies in underlying policy presumptions.

MANAGING FOR RESULTS

Good program performance information is essential to good strategic management. Managers at all levels must know the results their programs are achieving--whether outputs, outcomes, or impact--if they are going to respond creatively to changing opportunities and constraints and achieve better results in the future. However, too many organizations, particularly federal agencies, seem to place their primary managerial emphasis on process and procedures, ensuring that activities are carried out rather than that results are achieved.

Program performance indicators will be at the core of any system for agencywide program performance *monitoring*. As such, they can tell us whether desired results are occurring and whether program outcomes are basically on track. They can also provide, at a fairly aggregate level, at least a rough basis for comparing the kinds of results that different programs are achieving (or that similar programs are achieving in different locations). Perhaps most importantly, they can provide a clear warning when something is wrong and when intended results are not being achieved.

By themselves, performance indicators cannot tell us why results have or have not occurred or which, among a range of program alternatives, is the most efficient and effective. However, in conjunction with the other evaluation, analysis, and reporting activities, program performance monitoring becomes a much more powerful tool. Program evaluations, for example, are specifically directed at answering questions about how programs are working, why results vary, and which program alternatives have the most impact and are the most cost-effective. Similarly, operational-level performance information provides a basis for summarizing service delivery, analyzing implementation efficiency, and assuring accountability in relation to Congressional earmarks.

Improving program performance--and the wider efficiency and effectiveness of federal agencies--will require a much stronger emphasis on managing for results. This kind of "performance management" will certainly require a much clearer formulation of program objectives and program performance indicators. Such indicators also provide the basis for other, related management improvements:

- o the development of explicit *performance contracts* between management levels, which clearly define program objectives and expected results, and for which managers can be held responsible;
- o the *decentralization* of program management based on such performance contracts, which avoids micro-management and frees senior executives for more strategic decision-making, while clarifying the responsibilities and decision-making authority of subordinates.
- o *more rational program decision-making* based on clearer program objectives, more comparative data on program performance, and a better understanding of program alternatives.
- o a better basis for *performance-based budgeting*, rewarding programs that achieve results and penalizing those that do not.

It should be noted, however, that the relationship between program performance indicators and budget decisions is neither simple nor straight-forward. Poor results may, for example, reflect the extent of the problem being addressed rather than inadequacies in the program, indicating a need for more rather than less resources. Alternatively, performance data may be more relevant to choosing among program alternatives than to establishing overall funding levels for a program area. Placing too much emphasis on narrowly-defined performance indicators in budget decisions could also distort the objectives being sought and the data being reported. More generally, *performance-based budgeting* raises difficult questions--not fully addressable here--about how various program objectives, performance levels, funding categories, and budget time frames should be inter-related.

IMPLICATIONS FOR CONGRESS

Good program performance information is just as essential to good Congressional oversight as it is to good agency management. Just as well articulated program objectives and performance indicators can help top agency executives focus on "managing for results," so too can they help Congress focus its oversight more strategically on agency policies, objectives, and achievements. Indeed, just as top executives can use performance indicators as a basis for clearer management contracts with subordinates, Congress could also use well defined strategic objectives and performance indicators as a basis for management "contracts" with agencies. Moreover, by clearly defining strategic performance expectations, and by reducing requirements for most operational performance reporting except in cases of program failure, Congress can greatly enhance its ability to assess major program and policy issues while reducing its need for operational "micromanagement."

Improved "performance management" by agencies is not, however, likely to happen on its own. One of the key findings of the GAO's General Management Reviews (GMRs) is that substantial performance information gaps exist in nearly every agency studied. Even more disturbing, the GMR's found that performance objectives in most agencies were themselves unclear, contradictory, ambiguous, and poorly related to program activities. Legislation requiring clearer program performance objectives and reporting would take an important first step towards filling this management need. Congressional involvement would likely be essential not only in helping to delineate program indicators, but also in helping to clarify the "strategic objectives" that these indicators are intended to reflect.

Program objectives and performance indicators are inextricably linked. Program objectives without corresponding performance indicators provide no basis for accountability, while indicators without accompanying program objectives remain empty promises. As the nation's highest policy and program overseer, Congress has a particularly critical role in working with top agency managers in defining key *strategic objectives* and corresponding performance indicators--the highest level results for which the agency will be held directly accountable. Congress also has a central role to play, particularly given resource limitations, in helping agency executives prioritize their objectives and make trade-offs among objectives that conflict.

It is strategic objectives, more than the historical peculiarities of agency structure, that define the core *programs* for which the agency would be held responsible. Once these strategic

objectives are clarified, lower-level performance indicators, while sometimes technically difficult to collect, are much easier to define.

By requiring agencies to develop, in consultation with relevant Congressional committees, clearer strategic objectives and performance indicators, Congress could provide the basis for more explicit performance "contracts" between itself and the Executive. Such "contracts" could also help distinguish information needed to meet internal management and external accountability concerns. Just as important, by reorienting agency executives towards "managing for results," such legislation would begin to revolutionize the meaning of "quality" in federal program administration.

IMPLICATIONS FOR THE GAO

An emphasis on "managing for results" also has a direct relevance to GMR's and other GAO activities. Certainly, helping agencies delineate clearer strategic objectives and performance indicators has already been an important part (though sometimes phrased in other terms) of most GMRs. However, a more explicit focus on such strategic "performance management" issues, could define a much clearer framework for future GMR activities. In addition to clarifying agency-wide goals, this might involve assessing the adequacy of "objective trees" and "performance indicators" for particular agency programs, assessing the correspondence of financial and program information, assessing the adequacy of performance information analysis and processing capabilities, etc.

Delineating clearer strategic objectives, performance indicators, and program logics could also define the key issues and criteria for GMR follow-on activities. Indeed, a clearer delineation of these strategic "performance management" concerns, might well define a continuing framework that would better link the entire range of the GAO's audit, evaluation, and GMR activities within an agency over time.

OBSTACLES TO STRATEGIC MANAGEMENT

We should not underestimate how difficult the transition to more strategic, performance-oriented management will be for many federal agencies. To be both relevant and practical, objectives and indicators must reflect a convergence between policy goals and operational program realities. This will require a collaborative "top down" and "bottom up" effort by managers at all organizational levels. And, a number of potentially serious roadblocks will have to be avoided before federal agencies will be able to manage more successfully for results.

Disagreements about Purposes and Goals:

Programs sometimes lack well defined strategic objectives because senior policy-makers are unable to agree on what their purposes and goals should be. Are foreign aid programs primarily expected to improve the well-being of people in developing countries or to advance U.S. national security interests? Are defense procurements intended to improve military readiness or to provide employment in key constituencies? Are school feeding programs created to avert hunger, improve school performance, or provide an income transfer to poor

families? Should customs inspections be designed to facilitate international trade or to police illegal imports and exports?

Performance indicators provide the clearest guidance for management action when they embody a linear program logic reflecting clear program priorities. The real world, however, is rarely so clear, and programs often encompass multiple and conflicting objectives. While it is difficult to portray multiple purposes within a single objective tree, it may sometimes be possible to define parallel objective trees reflecting these multiple purposes. Such performance data could then be useful in making decisions that reflect trade-offs among competing objectives, particularly when such objectives can be prioritized or weighted.

It is more difficult to deal with disagreements about goals when such goals are truly incompatible or mutually exclusive. Such disagreements may ultimately need to be resolved through political decisions by senior policy-makers or the Congress. Unfortunately, ambiguities about program objectives may themselves be politically desirable. In such cases, delineating useful performance indicators could prove quite difficult.

Lack of Correspondence Between Objectives and Programs:

According to the old adage, "if you don't know where you are going, any road will take you there." Without a clear understanding of what it is we are trying to achieve, it is impossible to define a program strategy, an appropriate set of program activities, or useful indicators of program performance. Indeed, in the absence of clear objectives, the arbitrary choice of "indicators" may implicitly define objectives that are inappropriate or unattainable.

But just as programs without corresponding strategic objectives provide no basis for accountability, strategic objectives that are not linked to real program activities remain empty promises. Strategic objectives need to be grounded in a logical structure of program activities that defines meaningful indicators at all management levels.

Working to the Measures:

Performance indicators are more useful for judging the performance of programs than for judging the performance of managers. Programs may fail because they lack sufficient resources, reflect inappropriate policies, face unanticipated external conditions, or for any of a variety of reasons. Managers should be held accountable for obtaining and using program performance data, for understanding why their programs are succeeding or failing; and for making appropriate changes to help their programs work better. Managers should, in other words, be responsible for *managing for results*, but not necessarily for results as such.

Linking personnel and budget decisions too closely to program performance can potentially corrupt the measures being reported or the objectives being sought. If managers know they will be judged on narrowly defined performance criteria, they will make sure, one way or another, that indicators come up right. This may involve outright dishonesty--keeping two sets of books--or merely too much emphasis on inappropriate program activities. One shouldn't forget the proverbial Russian tool factory that only made penny nails when its managers discovered that their performance ratings would be based on the number of items produced. Particularly if objectives are poorly defined, too much emphasis on performance

indicators may encourage risk-averse managers to implement only those activities they are sure will succeed.

CONCLUSIONS

The concern with measuring program performance and using performance information in management decisions is certainly not new. Programming, planning, and budgeting systems; management by objectives; and other performance management approaches have been tried by federal agencies before and have, for the most part, not lived up to their expectations. What, then, are the grounds for optimism that current efforts will be more successful?

For one thing, performance management appears to be an idea whose time has come. The budget deficit, funding caps, and the need to make hard program and policy trade-offs has placed a premium on good performance information and made it much more relevant to decisions that now *have* to be made. Secondly, there appears to be clear Congressional recognition that more micro-management will be counter-productive, but a continuing need for the kind of accountability that performance indicators can help fulfill. Finally, there is a growing awareness that *quality management* can only be based on well-defined objectives and measurable achievements.

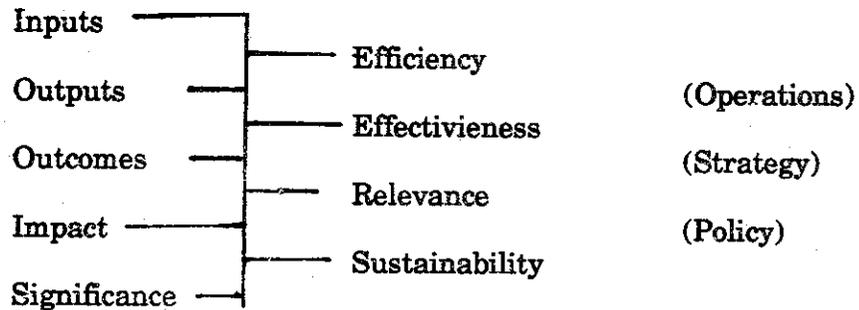
Even more important, performance information is now seen more clearly as part of a *strategic* management process, rather than as a mechanical programming and budgeting tool. Defining objectives and indicators is inexorable intertwined with the delineation of programs and strategies--a collaborative process involving all levels of program management.

There seems little doubt that federal agencies will be required to develop some kind of program performance indicators. The issue is doing it right, in a way that really improves the quality of federal services. The challenges are great, but the costs of not meeting these challenges are greater still.

Annex I

INDICATOR TYPOLOGY

Cause and Effect Hierarchy



Explanatory Approaches

- Analysis of Linkages
- Analysis of Assumptions
- Analysis of Environment

Program Types

- o Service
- o Police/Defense
- o Intergovernmental
- o Regulatory
- o Funding/Granting
- o Policy

Indicator Levels

Operational/Implementation	Activities	Benchmarks/Targets
Operational/Management	Programs	Targets
Senior Program Management	Strategies	Strategic Objectives (Indicators)
Top Agency Management/ Congress	Policies	Goals (Trends)

REFERENCES

AID

- 1989 *U.S. Assistance for Africa: The Development Fund for Africa--An Action Plan.* AID: Washington, D.C.

Bingman, C.F. & J.D. Young

- 1989 *An Assessment of the General Management Review Program of the U.S. General Accounting Office.* National Academy of Public Administration: Washington, D.C.

Britan, G.M.

- 1990 "Proposed Program Performance Indicators Strategy" (manuscript prepared for GAO/GGD)

Campbell, D.T.

- 1963 "From Description to Experimentation" In C. W. Harris (ed.), *Problems in Measuring Change.* Madison: University of Wisconsin Press.

Campbell, D.T.

- 1965 *Unobtrusive Measures.*

Davis, D.F.

- 1990 "Do You Want a Performance Audit or a Program Evaluation?" Public Administration Review, Jan/Feb 1990; pp. 35-41.

Fogel, R.

- 1990 "Approach to and Schedule for GMRs" (Internal Memo to Comptroller General Bowsher)

GAO General Management Reviews

Environmental Protection Agency: Protecting Human Health and the Environment Through Improved Management (GAO/RCED-88-101; 8/88)

General Services Administration: Sustained Attention Required to Improve Performance (GAO/GGD-90-14; 11/89)

Management of HHS: Using the Office of the Secretary to Enhance Departmental Effectiveness (GAO/HRD-90-54; 2/90)

Managing the Government: Revised Approach Could Improve OMB's Effectiveness (GAO/GGD-89-65; 5/89)

Social Security: Status and Evaluation of Agency Management Improvement Initiatives (GAO/HRD-89-42; 7/89)

GAO Reports

Evaluating A Performance Measurement System--A Guide for the Congress and Federal Agencies (FGMSD 80-57; 5/80)

Finding Out How Programs Are Working: Suggestions for Congressional Oversight (PAD 78-3; 11/77)

Office of Personnel Management: Better Performance Information Needed (GAO/GGD-90-44; 2/90)

Managing the Cost of Government: Building an Effective Financial Management Structure (GAO/AFMD-85-35-A; 2/85)

Quality Management Scoping Study (GAO/ACG-Ops-91-1; 12/90)

Gembacz, W.

1990 "VA-GMR: Performance Monitoring," Notes from a 10/31/90 Interview with Joseph Wholey.

n.d. "Outcome-Oriented Performance Monitoring: Key Step to Improving Management and Service Quality" (manuscript)

Hatry, H.P., et al.

1990 *Service Efforts and Accomplishments Reporting: Its Time has Come*. Government Accounting Standards Board: Norwalk, Connecticut.

Raaum, R

n.d. "Performance Results that Warrant Measurement and Reporting" (manuscript)

Wholey, J. S

1990 "Improving Government Management and Performance: Goal-setting, Performance Monitoring, and Annual Reporting," Testimony before the Senate Committee on Governmental Affairs. October 3, 1990.

Walter, E.F.

1989 "Future of the GMR Effort" (Internal Memo to ACG Richard Fogel)