

148250

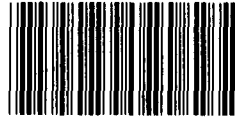
United States General Accounting Office

GAO

Transition Series

December 1992

Program Evaluation
Issues



148250

148250

GAO/OCG-93-6TR



**United States
General Accounting Office
Washington, D.C. 20548**

**Comptroller General
of the United States**

December 1992

The Speaker of the House of Representatives
The Majority Leader of the Senate

In response to your request, this transition series report discusses a topic that is critical to the effective oversight of government programs: the need for sound, evaluative information on how programs are operating and what they are actually accomplishing. This report, unlike our 1988 transition series report on this topic, cites some examples of good work being done within executive branch agencies. More generally, however, we feel that the attention being paid to evaluation issues is inadequate either for managing programs efficiently or for providing the Congress with the data necessary for informed program oversight.

The GAO products upon which this report is based are listed at the end of the report.

We are also sending copies of this report to the President-elect, the Republican leadership of the Congress, the appropriate congressional committees, and the designated heads of the appropriate agencies.

A handwritten signature in black ink that reads "Charles A. Bowsher". The signature is written in a cursive, flowing style.

Charles A. Bowsher

Contents

The Importance of Program Evaluation	4
Rebuilding Capacity	7
The Effects of Many Important Programs Are Unknown	10
Some Agencies Are Poorly Informed About Program Targeting and Outreach	14
Agencies Sometimes Rely Upon Flawed Studies and Ignore or Misuse Sound Analyses	18
Some Promising Initiatives Do Exist	23

Contents

Related GAO Products	25
Transition Series	28

The Importance of Program Evaluation

Over the next few years, the federal government will face powerful opposing pressures: the need, on the one hand, to reduce the federal deficit, and the demand, on the other, for a federal response to some potentially expensive domestic problems (expanding health insurance, restoring the economy, and the like). These pressures are likely to intensify concern with the effective management of federal programs and with the availability of objective information on the results of federal investments. In other words, are the federal officials who administer programs adequately informed about the implementation and the results of those investments? And can they, in turn, adequately inform the President, the Congress, and the nation about what has been accomplished?

Program evaluations contribute systematic information to federal decision-making that has been useful in a variety of ways, such as the following:

- Recent welfare-to-work evaluations funded by the Department of Health and Human Services showed a relationship between the increased employment of program participants and savings to the federal and

state governments achieved through reduced welfare payments.

- Evaluations of chemical weapons demonstrated gaps in the military's capacity to manage and use these weapons and played a major role in the termination of the Bigeye Bomb program and in the successful completion of ongoing arms control negotiations on chemical warfare.
- Evaluations conducted some years ago showed the effectiveness of the Job Corps program in preparing disadvantaged young men and women for employment and were a major contributor to the reauthorization of this expensive intervention.
- The Safe Medical Devices Act of 1990 included a number of provisions that were the direct results of findings in a series of medical device evaluations. Among other things, the act provides for increased recall powers for the Food and Drug Administration and for improved information reported to the Congress.

It is important to recognize that an objective and systematic evaluation function not only serves to protect an agency against wasted resources in the form of inefficient or ineffective programs. There may also be

elements of government programs that are in fact harmful to the well-being of some segments of society—unintended effects that a well-conducted program evaluation could prevent or detect. For example, an evaluation of the likely impacts of proposed immigration reform legislation—suggesting that the proposal would result in long waiting lists and delay the reunification of families—led to appropriate changes in the bill before its enactment in 1990.

If the nation is to have strong, well-managed federal programs that can deal efficiently and successfully with our domestic and international problems and if the President and the Congress are to be adequately informed of progress in meeting those challenges, the numbers and quality of the program evaluations conducted by executive branch agencies must be improved.

Rebuilding Capacity

In our 1988 transition series report, we found that there had been a 22-percent decline in the number of professional staff in agency program evaluation units between 1980 and 1984. A follow-up study of 15 units that had been active in 1980 showed an additional 12-percent decline in the number of professional staff between 1984 and 1988. Funds for program evaluation also dropped substantially between 1980 and 1984 (down by 37 percent in constant 1980 dollars). We have not repeated this survey, but discussions with the departments and the Office of Management and Budget offer no indication that the executive branch investment in program evaluation showed any meaningful overall increase from 1988 to 1992.

Apparently, the effort to rebuild the government's evaluation capacity that we called for in our 1988 transition series report has not been carried out. As in 1988, executive branch agencies have often failed to conduct the program evaluations that would provide officials with knowledge about the effectiveness of their programs. As in 1988, the Congress continues to turn to us and our sister agencies—the Congressional Budget Office, the Office of Technology Assessment, and the Congressional Research

Service—to do studies that might more appropriately be conducted by executive branch agencies. It is our mission to provide credible information to the Congress and to help ensure that the reports the Congress receives are not limited to those from special interest groups. However, we should not, and indeed cannot, do it all.

A first step in improving capacity is for agencies to review the adequacy of their current funding for evaluation. Some agencies, like the Department of Commerce and the Administration on Aging, devote few resources to evaluating their programs. Other agencies—like the Department of Education and the Public Health Service—dedicate major resources to evaluation. In these agencies, the task may be less one of rebuilding overall capacity than of strengthening areas of weakness. For example, the Department of Education conducts many evaluations of its elementary and secondary programs, but the investment at the postsecondary level is irregular in spite of some major ongoing problems. Similarly, some Public Health Service agencies do a great deal of evaluation, whereas others, like the National Institutes of Health, spend only a small percentage of available funds for program evaluation.

Our evaluations help fill the gaps in the information available to the Congress, but there is no substitute for a systematically planned, ongoing effort for an agency to evaluate its own programs. The next sections demonstrate that the limited capacity for program evaluation in the executive branch has some important consequences:

- agencies lack information on the effectiveness of their programs;
- agencies lack data on the targeting and outreach of their programs; and
- agencies need to improve their capacity to make sound decisions on the use of data for policy-making.

We are, however, able to conclude with some examples of promising agency initiatives.

The Effects of Many Important Programs Are Unknown

One of the most significant gaps in program evaluation information from the executive branch concerns program effectiveness. Program effectiveness evaluations show what, if anything, has changed as a result of implementing a program. Limited budget dollars can be more concentrated among programs that have demonstrated effectiveness, while programs with little evidence of effectiveness can be cut or reformed and restructured.

Do participants in federally funded elementary and secondary education programs for disadvantaged children show improvements in educational achievement similar to that of other children? Are federal housing vouchers shown to be effective in helping needy persons who would often not have adequate housing without the vouchers?

In short, is there evidence of some concrete benefit that results from a program that would not have occurred without the program? Program effectiveness evaluations estimate the effects of federal programs using statistical analysis of outcomes (such as educational achievement test scores or condition of housing) for groups of persons

receiving program services compared with similar groups of nonparticipants.

Our response to congressional requests over the last 4 years has yielded the following information on the effects of federal programs.

Education

Despite increased attention in recent years to removing the barriers that prevent the full involvement of persons with disabilities in work and other activities, the Department of Education has not evaluated the effectiveness of its \$1.8 billion-per-year program of vocational rehabilitation. Our evaluation using confidential income tax data showed only very modest overall long-term gains in earned income, in contrast to the dramatic short-term employment effects often cited by the program. The Congress has strengthened evaluation in the recent reauthorization of the program. The Department of Education should conduct such studies and help establish the overall impact of the legislation by identifying, for example, how people with certain disabilities have been helped by the program while those with other disabilities may need different assistance than the program has provided.

Health and
Human Services

The Department of Health and Human Services regularly makes public recognition awards to "promising" drug abuse prevention programs on the basis of reviews that required no hard evidence of program effectiveness. The problem here is that people in other communities could base new programs upon weak models while other, more effective programs go unrecognized. In response to our report, the agency agreed to begin seeking such evidence.

Health and
Human Services

To correct perceived widespread abuses of foster care, federal reforms were enacted in 1980 to ensure that the necessity and appropriateness of each foster care placement was periodically reviewed and that families received needed services. A 1989 evaluation found that these reforms had not been completely carried out and that no national evaluations had been performed. This evaluation gap means that children placed in foster care may still be unnecessarily at risk of abuses such as needless delays in returning them to their natural parents.

Defense

We evaluated the Department of Defense's (DOD) methods for selecting recruits for

training in technical occupational specialties and for assessing the effectiveness of the training. DOD's major selection instrument, the Armed Services Vocational Aptitude Battery, which has been extensively researched over the years, was moderately successful in predicting classroom performance in these more demanding training courses. But in most cases, testing of actual field performance—the end point of the training program—was either nonexistent or inadequate, making it impossible to evaluate the overall effectiveness of the services' training program. Thus, much of the investment that the Department had made in program evaluation was in this case inefficient because the final loop—showing which training programs produce the top performers and thus make the best use of DOD's human resources—had not been closed.

Some Agencies Are Poorly Informed About Program Targeting and Outreach

A second type of program evaluation that is useful—but often not available from federal agencies—concerns questions of program implementation. Agencies should evaluate different aspects of program implementation, such as the proper targeting of programs and their outreach—that is, whether they reach some or all of those eligible.

Analysis of program targeting demonstrates how well the program is reaching its intended recipients (such as determining the extent to which Chapter 1 aid, under the Elementary and Secondary Education Act of 1965, is targeted for and reaches educationally disadvantaged children). Studies of outreach examine participation in federal programs.

Such evaluations help agencies understand why their outreach may not be successful and what barriers may need to be overcome before participation can increase. There are important gaps in agencies' knowledge of the targeting and outreach of their programs.

The limited executive branch evaluations of targeting and program outreach led to congressional requests for us to do work such as the following.

**Health and
Human Services**

The National Cancer Institute is responsible for disseminating information on treatments proven to be effective in the treatment of cancer in experimental situations. An evaluation revealed various blockages in the processes that the Institute used to move breakthrough therapies from clinical trials to the patient. Recommended treatments had not been adopted for many patients in the samples we studied. Consequently, these patients had not received what the Institute considered state-of-the-art treatment.

Education

Federal student aid is especially intended to help persons of modest means gain access to postsecondary education, but sound information on school costs and the availability of aid has to reach potential recipients in time for them to make crucial decisions about their higher education plans. An evaluation documented both unawareness and incorrect understanding of the program that could have significant effects on the decisions students and parents make. The Congress acted on these findings in revising student aid laws in 1992 to require the Department to make improved information available and to evaluate its impact.

Agriculture

We found in reviewing the Department of Agriculture's Food Stamp program that less than half of the households eligible for food stamps participated in the program in the mid-1980s. Evaluators found evidence of a variety of outreach problems. About half of the eligible nonparticipants incorrectly thought that they were ineligible. Almost two-thirds of the eligible nonparticipants cited either a lack of information or program barriers, including administrative "hassles," as the reason for their nonparticipation. These findings suggest that the program should be changed to make food stamps more available to eligible persons who need them.

Aging

The Older Americans Act mandates that in the provision of services, the Administration on Aging should target older individuals with the greatest economic or social needs and give particular attention to low-income minority individuals. A program evaluation found that the data collection instrument and methodology used by the agency did not permit the generation of accurate counts of all participants, including targeted populations, in mandated programs and services. The data are thus inadequate to answer fundamental congressional questions

**Some Agencies Are Poorly Informed
About Program Targeting and Outreach**

about the degree to which agency programs target resources to persons with the greatest economic or social needs.

Agencies Sometimes Rely Upon Flawed Studies and Ignore or Misuse Sound Analyses

In 1992, we found some new consequences of the deterioration of the program evaluation capacity of the federal government in addition to the absence of information. In some cases, agencies have conducted evaluation studies, but the information produced is either flawed or improperly used for policy purposes. The studies may be based upon problematic data or analysis, or they may be properly conducted but ignored or misused in the formulation of policy. These problems suggest that management improvements are necessary. Agencies need to review the quality of their data and research more carefully and better integrate the findings from program evaluations and other analyses into agency decision-making.

Examples of inappropriate use—or neglect—of evaluation evidence by federal agencies in recent years include the following.

Environmental Protection Agency

An analysis of the shared responsibilities of the Environmental Protection Agency (EPA) and state environmental agencies for managing a national program of hazardous waste found that important information gaps remain. Problematic measurement and data

collection procedures limit the quality of some of the information produced, and the biennial reporting system still does not ensure that the states will collect or report to EPA all necessary data. The result is that it is not possible to determine whether the hazardous waste reduction goals are being met.

Multiple Agencies

We conducted a comprehensive evaluation of four executive branch agencies—EPA, the Food and Drug Administration, the Occupational Safety and Health Administration, and the Consumer Product Safety Commission—to determine how well they were protecting the public from exposure to reproductive and developmental toxicants. There were major gaps in the evaluative information available to these agencies. First, because no accepted federal list of these toxins has been developed, such as that mandated for carcinogens, these agencies have had no index of the most important hazards to reproduction and development. Second, risk assessment for these toxicants has been based upon a flawed threshold assumption: that is, that there is a specific dose level below which no problems occur. However, well-known hazards, such as lead and radiation, are

dangerous at any dose level. Therefore, it appears that current standards for exposure to lead and radiation could be resulting in more developmental problems in children and reproductive problems in adults than would occur under an alternative, nonthreshold approach.

Defense

Many unproven assumptions weaken DOD's decision-making in important national security areas. Our evaluation of the U.S. strategic triad found several major examples of assumptions that we found to be either inaccurate or unsupported by available data. In some areas, such as the threat posed by the former Soviet Union, the assumptions grossly overstated what the data actually support. In other areas—for example, the performance of weapon systems—available data instead show understated assumptions. In still other cases, such as specific assertions made by officials, no supporting data were available. Over the past 30 years, DOD has not conducted any comprehensive evaluation of the strategic nuclear triad. The lack of realistic assessments of the threat and the lack of rigorous analysis of the relative performance and merit of the weapon systems has resulted in the

Agencies Sometimes Rely Upon Flawed Studies and Ignore or Misuse Sound Analyses

questionable development and procurement of multiple costly modernization programs.

Agriculture

A series of studies examined the accuracy of various price, production, and supply forecasts made by the U.S. Department of Agriculture. Long-term commodity forecasts had large and systematic error rates over the period 1981-88. These errors contributed to a significant underestimate of the commodity program outlay estimates that were made in the President's 1990 budget submission to the Congress. Accurate forecasts are also important for administering such agency programs as acreage reductions and export enhancements. The Department agreed with recommendations to improve forecasting and set up a process to identify, report, and correct errors when they occur.

Transportation

We reviewed the quality and completeness of the analytical efforts supporting the testimony of the National Highway Traffic Safety Administration regarding the impact of continued requirements for automobile downsizing upon highway safety. The agency's finding that more than 1,300 fatalities each year can be attributed to the automobile weight reduction efforts that

**Agencies Sometimes Rely Upon Flawed
Studies and Ignore or Misuse Sound
Analyses**

began in the 1970s was not supported by available data. Instead, our analyses showed that the automobile weight reductions have had virtually no net effect on total highway fatalities. On the one hand, the very lightest cars have higher fatality rates than the very heaviest. On the other hand, the decreased number of heavy cars on the highways accounts for much of the total car weight reduction that occurred during the time period, diminishing the danger to the occupants of other cars.

Some Promising Initiatives Do Exist

This year, in contrast to 1988, we identified a number of agency evaluations that were well done or work in progress that seems promising.

Centers for Disease Control

The Congress requested an independent review of the investigation by the Centers for Disease Control (CDC) into the causes of the HIV infections found among patients of a Florida dentist. We concluded that CDC's research—especially that on genetic sequencing—was well done and our review supports CDC's conclusions that five patients became infected as a result of receiving care from the dentist with AIDS, although the mode of transmission remains uncertain. This review suggests that CDC made an exemplary effort to find and use the best information available at the time.

Agency for Health Care Policy and Research

We reported that despite considerable changes in the management of breast cancer treatment since the 1970s, there was no observable improvement in survival. That we were the only organization to examine nationally how well cancer patients fared underscored the fact that there was no federal agency charged specifically to examine health outcomes. That situation

changed in 1989 with the establishment of the Agency for Health Care Policy and Research. Currently, this agency has within its mandate the broad areas of determining the effectiveness of medical interventions, creating medical practice guidelines, and disseminating information on outcomes. This offers the promise of advances in the quality and quantity of information that will be available for making informed, future health care policy decisions.

**Office of
Personnel
Management and
Merit Systems
Protection Board**

In response to concerns in the Congress and among such groups as the National Commission on the Public Service that the quality of the federal professional and technical workforce was declining in the 1980s, we examined the available data and found no significant evaluation of workforce quality either in the major agencies or in the Office of Personnel Management (OPM). However, in response to our 1988 report, OPM and the Merit Systems Protection Board have in the last 4 years developed significant evaluation programs involving measurements of quality among those recruited and retained in a number of key occupations, and an expanded effort is being planned on the basis of a national advisory panel's work.

Related GAO Products

Importance of Program Evaluation

Bigeye Bomb: Unresolved Developmental
Issues (PEMD-89-27, Aug. 11, 1989).

Immigration Reform: Major Changes Likely
Under S.358 (PEMD-90-5, Nov. 9, 1989).

Medical Devices: FDA's Implementation of
the Medical Device Reporting Regulation
(PEMD-89-10, Feb. 17, 1989).

Program Evaluation Issues (OCG-89-STR, Nov.
1988).

Program Effects Often Unknown

Vocational Rehabilitation Program: Client
Characteristics, Services Received, and
Employment Outcomes (T-PEMD-92-3, Nov. 12,
1991).

Drug Abuse Prevention: Federal Efforts to
Identify Exemplary Programs Need Stronger
Design (PEMD-91-15, Aug. 22, 1991).

Military Training: Its Effectiveness for
Technical Specialties Is Unknown (PEMD-91-4,
Oct. 16, 1990).

Foster Care: Incomplete Implementation of
the Reforms and Unknown Effectiveness
(PEMD-89-17, Aug. 14, 1989).

Program
Targeting Often
Unknown

Higher Education: Gaps in Parents' and
Students' Knowledge of School Costs and
Federal Aid (PEMD-90-20BR, July 31, 1990).

Food Stamp Program: A Demographic
Analysis of Participation and
Nonparticipation (PEMD-90-8, Jan. 19, 1990).

Cancer Treatment: National Cancer
Institute's Role in Encouraging the Use of
Breakthroughs (PEMD-89-4BR, Oct. 20, 1988).

Failure to Use
Quality Studies

Summary of the Strategic Nuclear Triad
Evaluation (PEMD-92-36R, Sept. 28, 1992).

Reproductive and Developmental Toxicants:
Regulatory Actions Provide Uncertain
Protection (PEMD-92-3, Oct. 2, 1991).

Highway Safety: Have Automobile Weight
Reductions Increased Highway Fatalities?
(PEMD-92-1, Oct. 8, 1991).

USDA Commodity Forecasts: Inaccuracies
Found May Lead to Underestimates of
Budget Outlays (PEMD-91-24, Aug. 13, 1991).

Waste Minimization: EPA Data Are Severely
Flawed (PEMD-91-21, Aug. 5, 1991).

Promising
Evaluation
Efforts

AIDS: CDC's Investigation of HIV Transmissions
by a Dentist (PEMD-92-31, Sept. 29, 1992).

Federal Workforce: A Framework for
Studying Its Quality Over Time (PEMD-88-27,
Aug. 4, 1988).

Transition Series

Economics

Budget Issues (GAO/OCG-93-1TR).

Investment (GAO/OCG-93-2TR).

Management

Government Management Issues
(GAO/OCG-93-3TR).

Financial Management Issues
(GAO/OCG-93-4TR).

Information Management and Technology
Issues (GAO/OCG-93-5TR).

Program Evaluation Issues (GAO/OCG-93-6TR).

The Public Service (GAO/OCG-93-7TR).

Program Areas

Health Care Reform (GAO/OCG-93-8TR).

National Security Issues (GAO/OCG-93-9TR).

Financial Services Industry Issues
(GAO/OCG-93-10TR).

International Trade Issues (GAO/OCG-93-11TR).

Commerce Issues (GAO/OCG-93-12TR).

Energy Issues (GAO/OCG-93-13TR).

Transportation Issues (GAO/OCG-93-14TR).

Food and Agriculture Issues
(GAO/OCG-93-15TR).

Environmental Protection Issues
(GAO/OCG-93-16TR).

Natural Resources Management Issues
(GAO/OCG-93-17TR).

Education Issues (GAO/OCG-93-18TR).

Labor Issues (GAO/OCG-93-19TR).

Health and Human Services Issues
(GAO/OCG-93-20TR).

Veterans Affairs Issues (GAO/OCG-93-21TR).

Housing and Community Development
Issues (GAO/OCG-93-22TR).

Justice Issues (GAO/OCG-93-23TR).

Internal Revenue Service Issues
(GAO/OCG-93-24TR).

Foreign Economic Assistance Issues
(GAO/OCG-93-25TR).

Transition Series

Foreign Affairs Issues (GAO/OCG-93-26TR).

NASA Issues (GAO/OCG-93-27TR).

General Services Issues (GAO/OCG-93-28TR).

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are \$2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

**U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20877**

or visit:

**Room 1000
700 4th St., NW (Corner of 4th & G Sts., NW)
U.S. General Accounting Office
Washington, DC**

**Orders may also be placed by calling
(202) 512-6000 or by using FAX number
(301) 258-4066.**

**United States
General Accounting Office
Washington, D.C. 20548**

**Official Business
Penalty for Private Use \$300**

**First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100**