The GAO Review (ISSN 0016-3414) is published quarterly by the U S General Accounting Office, Washington, DC 20548 Second class postage paid at Washington, DC.
Postmaster: Send address changes to U S General Accounting Office, 441 G St N W , Room 4026, Washington, DC 20548

Contents

From Our Briefcase

Government Financial Management Topics

The bimonthly newsletter of the Association of Government Accountants (AGA), *Government Financial Management Topics*, includes a feature concerning developments, procedures, manuals, regulations, and other news in the financial management area. The feature, called "Technical Notes," generally has 10 to 15 such references. Excerpts from the May/June 1986 newsletter are highlighted here. For more information, contact Chuck Hamilton, Director of Publications, AGA, 727 S. 23rd Street, Arlington, Virginia 22202, (703) 684-6931.

Clearinghouse for Single Audits. The Office of Management and Budget (OMB) has designated the Bureau of the Census as the national clearinghouse for single audit reports from state and local governments. OMB explained that the law required it to report on specific governments that were not in compliance with the Single Audit Act. To do this, it was necessary to determine which governments were required to have an audit (those receiving federal assistance of $100,000 or more directly or indirectly through other state and local governments) and then to ascertain whether the audits had been made. The clearinghouse was the most efficient way of doing both, with state and local governments required to file copies of their audit reports.

The Census Bureau was selected, OMB said, because of its long-established public sector statistical programs. Call the clearinghouse coordinator, William Fanning, at (301) 763-4403 for more information.

Accountants' Supply and Demand. The American Institute of Certified Public Accountants' annual survey of the supply/demand for accounting graduates foresees a 3-percent rise in the number of graduates versus a 6-percent increase in public accounting firm hires. The firms anticipate that the greatest demand will be for recruits with graduate degrees in taxation. Female accounting graduates will constitute 48 percent of the total hired.

DioGuardi Introduces Financial Management Reform Bill. Congressman Joe DioGuardi has introduced the Federal Financial Management Improvement Act of 1986 (H.R. 4495), which would establish an independent Office of the Chief Financial Officer of the United States within the Executive Office of the President and an Office of the Assistant Secretary for Financial Management within each executive agency. A Federal Financial Management Council would be made up of these new positions.

Roth Proposes Management Reforms. Senator William V. Roth, Jr., has introduced the Federal Management Reorganization and Cost Control Act of 1986 (S. 2230). It would make changes in the federal management structure, accounting and internal control systems, agency financial statements, cash management, credit and debt management, and information policy. Among other features, the bill would

- create an Office of Federal Management in the Executive Office of the President, overseeing, among others, an Office of Financial Systems (OFS);
- require that the accounting and financial systems of all agencies be approved within 2 years (those systems not in compliance would come under the direct supervision of OFS after 2 years);
- require biennial independent audits of annual agency financial statements;
- require more timely disbursement of federal funds and state payments to the federal treasury;
- make federal loans unavailable to those in default on other federal loans; and
- establish debt collection targets for each agency, with sanctions if collections fell behind such targets.

Communicating for Audience Attention

Effective oral communication—especially public speaking—is a particularly important skill, yet one that causes great anxiety. Public speakers have used numerous techniques and tricks-of-the-trade with great success. A summary has been adapted from "How to Recapture Audience Attention," by Robert P. Levoy, in the November 1984 issue of the *The Toastmaster*, published by Toastmaster International. For more information on that organization, call Jim Strange, President of GAO's Generally Able Orators at (202) 275-4195.

According to Levoy, "During any speech, especially a long one in a warm, crowded room after a big lunch, an audience's attention is apt to drift, nay—likely to drift. The signs are unmistakable: glassy-eyed stares, clock-watching, stifled yawns, doodling."

What can a speaker do to regain the attention of an audience, make people perk up, and mentally rejoin the meeting?

Introduce a Change of Pace

The speaker who drones on and on has a soporific effect on the audience. To avoid this, vary the pace. Talk fast, talk slowly.

- Ask a rhetorical question—in a whisper.
- At an appropriate moment, slam the lectern.
- Stop talking altogether. Look at the audience. Say nothing for 10 seconds. One by one, you'll get everyone's attention. Continue by saying, "The reason I paused is. . . ."

- Show something—anything. Don't miss an opportunity to show the audience whatever it is you're talking about. It could be a machine part, a fountain pen, a dollar bill, a drawing, a graph.
- Pick up the pitcher of water placed at most podiums and s-l-o-w-l-y pour yourself a glass of water. Simple as it is, it will get attention.
- Drink the water, or put it down.
- On an easel facing the audience, display several brightly colored showcards with just one word printed on each (in contrasting colors). These can be "key" words in your speech, representing the three or four points you really want to put across to your audience. Flip the cards as their labeled points appear in your speech. Space them throughout your speech for maximum impact.

Involve the Audience

There are countless other verbal and visual "tricks" to recapture an audience's attention. One of the most effective is to get people involved and participating in the meeting itself. Ask them to do something. For example, ask for a show of hands. The question they'd be responding to is not as important as is the thought. Asking a question will bring an audience back from its daydreams.

Another way to get an audience involved and participating is to distribute a test or a puzzle, particularly one that people can score themselves. This is always challenging and a sure-fire attention-getter.

Give the group a test of their "powers of observation." The instructions are simply to read the following paragraph aloud once and ask the group to count the number of letter *f*'s.

To appreciate the point that was later made, here's the test paragraph—try it yourself:

"The necessity of training farm hands for first-class farms in the fatherly handling of farm livestock is foremost in the minds of farm owners. Since the forefathers of the farm owners trained the farm hands for first-class farms in the fatherly handling of farm livestock, the farm owners feel they should carry on with the family tradition of training farm hands of first-class farms in the fatherly handling of farm livestock because they believe it is the basis of good fundamental farm management."

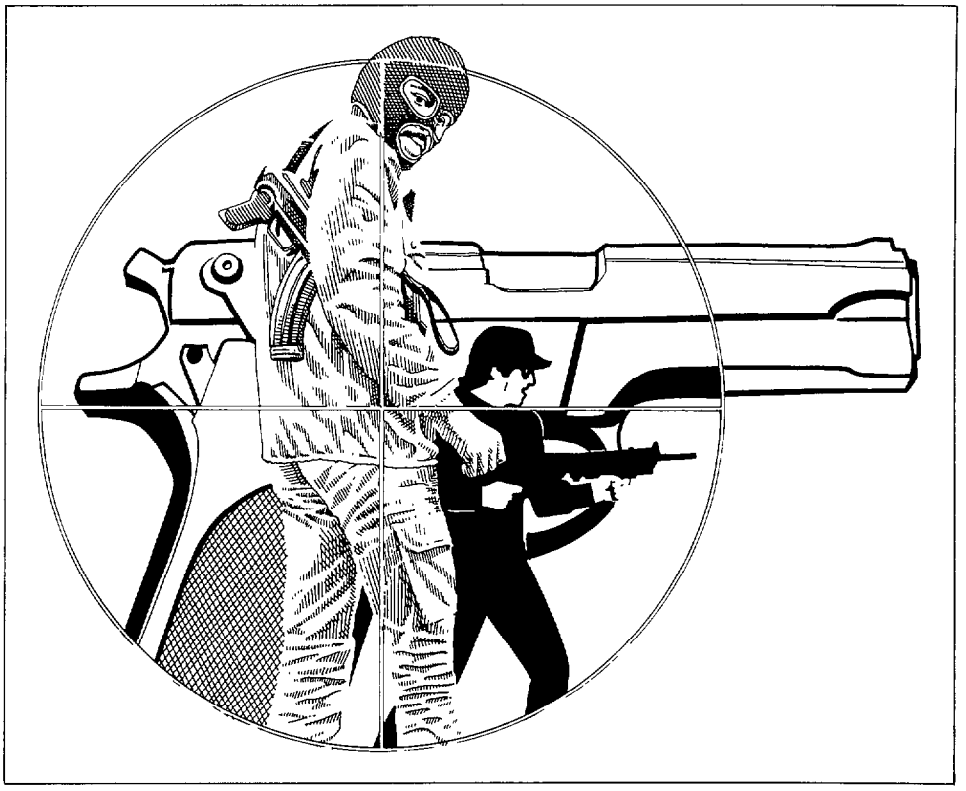
How many letter *f*'s did you find? In one group, the number of *f*'s varied from a low of 18 to a high of 37. Most people report in-between numbers. In fact, the amount of variation is always surprising. Mr. Levoy concludes that this test not only gets everyone's attention but also drives home a point that is most appropriate to this article: "Never underestimate the communication task—especially during a long speech in a warm, crowded room after a big lunch."

Brief Bibliography on Terrorism

GAO's Office of Library Services prepares its "Brief Bibliography" on numerous current topics. The bibliography entitled "Terrorism: 1985–86" includes citations of over 100 periodical articles, congressional hearings and reports, bibliographies, and Congressional Research Service (CRS) reports, covering the period 1985 through the first 4 months of 1986. It would be useful to researchers in security, psychology, or law and would alert GAO evaluators planning or conducting overseas audit assignments to issues they might face.

To introduce the bibliography, a CRS report defines terrorism: According to the Central Intelligence Agency, terrorism is the "threat of use of violence for political purposes by individuals or groups with the intent to shock or intimidate a target group rather than the immediate victims. Such acts may be directed against foreign nationals, institutions, or governments or against one's own nationals, insitutions, or government. Over 8,000 significant incidents of international terrorism have occurred since 1968, when the compilation of such statistics began. Over 50 percent of recorded terrorist incidents have been directed against U.S. interests, particularly against U.S. personnel and installations overseas."

The variety of subjects in the bibliography indicates the wide-ranging nature and influcence terrorism has on the world. Subjects include airport and air traveler security, vulnerability of chemical plants to terrorism, the media's role in international terrorism, psychological operations against terrorism, and legal aspects of terrorism. One citation highlights a speech on the need to fight terrorism through the law. In a 1985 address before the American Bar Association Convention, the Legal Advisor of the Department of State reviewed the hijacking of a TransWorld Airways flight from a legal perspective, pointing out the "inadequacies and obstacles to meaningful legal actions." He called for lawyers to fight and work for improved international laws to deal with lawlessness and enumerated specific actions to do this.



The bibliography entitled "Terrorism: 1985–86" is available from the Technical Library, order number OLS-86-2. Call (202) 275-5180 for more information. In addition, the State Department's Foreign Service Institute offers a 1-day seminar on dealing with terrorism abroad, which federal employees who expect to travel overseas on business may attend free of charge on a space-available basis. Call (202) 235-3417 for details. ■

On Location

Microcomputers in GAO

Microcomputers have become widely available in GAO and are increasingly affecting the way staff work. These powerful but relatively easy-to-use tools offer the possibility of significant increases in productivity if appropriately used. GAO has recently issued initial guidelines on micro use. *Using Micro Computers in GAO Audits: Improving Quality and Productivity* (Information Management and Technology Division Technical Guideline 1, Mar. 1986) addresses a range of issues that must be considered as computers become a more integral part of the audit process.

Information Technology Services

GAO has taken steps to ease the introduction and use of micros and to address the challenges that arise as use increases. The Office of Information Resources Management (OIRM) and the Office of Organization and Human Development (OOHD) have jointly established a program of user training, technical assistance, and information sharing. The purpose of Information Technology Services (ITS) is to encourage and assist all GAO personnel—managers, evaluators, and support staff—in learning to use information technology to improve personal productivity. ITS offers various microcomputer training and technical services. These services are supplemented in many divisions and regions by microcomputer services offered by unit technical assistance groups or micro information centers.

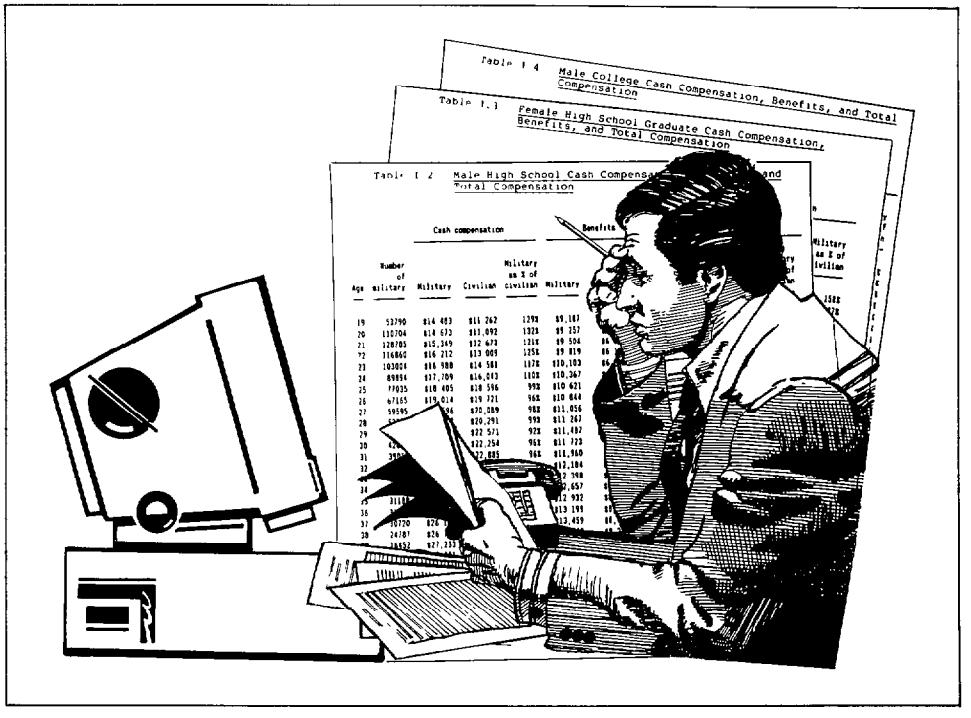
Hands-on Training. New microcomputer users are encouraged to enroll in formal training before using the technology in the workplace. OOHD offers courses introducing computer hardware and covering all of GAO' standard software, including WordPerfect, Lotus 1-2-3, dBase III, and Crosstalk. OIRM supplements formal classroom sessions by providing individual or small group training on specific topics. Hands-on training to individuals and small groups is also provided in many divisions and regions by staff within the unit.

Technical Assistance. Each division and office has a microcomputer "focal point" who serves as the first point of contact for assistance in using micros. Some divisions also have a network of group focal points assisting in providing technical assistance to operational groups. OIRM staff help to coordinate GAO-wide focal point activities. Additionally, they provide direct technical assistance through the OIRM Customer Service Desk (275-OIRM) and through electronic mail (GAO.MICRO.MAIL).

Information Sharing. Sharing of information and experiences among GAO micro users can be a major vehicle for ensuring the appropriate use of the new technology in the GAO work environment. Users can share the lessons they learn and seek help from colleagues through an electronic bulletin board maintained by OIRM staff. The bulletin board can be accessed through Crosstalk using an available micro. (Dial 275-1050.)

User groups have been established for Lotus 1-2-3 users and for dBase III users. Each group has had several meetings where users share information and expertise about these packages. Newsletters and meeting minutes are available upon request.

New Product Evaluation. ITS staff plan and coordinate reviews and evaluations of new hardware and software products that may be of general use in GAO. Reviews in progress include statistical software available for micros and graphics packages for microcomputers. Several pilot projects are examining ways of linking micros and other computers to provide a more integrated approach to automation in GAO.



Other Microcomputer Initiatives

The second annual technical conference (see next item) provided an opportunity for micro users throughout the agency to share experiences using microcomputers on the job. In virtually all divisions and regions, promising applications of micro technology are emerging. Sharing the benefits and the pitfalls of using micros on audits and evaluations will help enhance the potential of this technology while limiting potential threats to quality that may be inadvertently introduced by the computer.

Ed. note: Thanks to Elizabeth Powell, Information Management and Technology Division, for the preceding item. The article in this issue, "Controlling the Quality of Electronic Workpapers," documents some of the specific concerns that must be addressed as electronic spreadsheets are used in GAO audits and evaluations.

Second Annual Technical Conference: Tools for Future

GAO officials and staff members explored GAO's use of state-of-the-art audit and evaluation techniques and learned about new approaches and methodologies on the horizon at GAO's second annual technical conference, held April 9-10 at the University of Maryland Center of Adult Education.

Some 277 GAO staff members—division directors, regional managers, other senior officials, technical assistance staff, and evaluator and evaluator-related staff from GAO divisions and regions—attended the conference. The meeting, sponsored by GAO's Technical Services Committee, is a forum for GAOers responsible for technical development and technical assistance to share ideas with each other and with outside experts. Deputy Director Ray Rist of the Program Evaluation and Methodology Division and Deputy Director for Operations Dave Littleton of the National Security and International Affairs Division were conference cochairmen.

Keynote speaker Hale Champion, Executive Dean of Harvard University's John F. Kennedy School of Government, challenged GAO to stretch its thinking and its perspectives to deal with emerging issues that government decisionmakers will be required to address.

The rest of the first day was devoted to presentations by eight university, defense university, and GAO experts on new research and analytical techniques. "The purpose of these sessions was to stretch our perspectives—to learn of the wide range of new methodologies that can be applied to GAO's work," Rist explained. The second day of the conference comprised 17 concurrent sessions, featuring GAO staff members discussing their use of innovative techniques on GAO assignments. "These sessions were planned to give participants a sense of the breadth of what our organization is doing and to encourage staff members to learn from and with each other," Rist said. Selected topics included complex sampling methodologies, secondary data analysis, uses of computer modeling, improving questionnaire response rates, and microcomputer statistical analyses.

Assistant Comptroller General for Operations Frank Fee, who discussed opportunities and challenges for technical development in GAO, emphasized that the annual technical conference is an integral part of GAO's technical training and executive education programs. Observers from the Offices of the Canadian Auditor General, the Department of Health and Human Services Inspector General, and the Naval Audit Group also attended the conference.

Human Resource Management: A "Basic GAO Value"

Another second annual conference—on human resource management—was held at the Training and Career Development Center on March 27–28.

The conference, cosponsored by the Assistant Comptrollers General for Operations and Human Resources, Frank Fee and Gregory Ahart, attracted about 60 individuals. Deputy directors and assistant regional managers for operations, headquarters staff office directors, and others discussed issues related to managing GAO's human resources.

Human resource management, according to Fee and Ahart, should be as important a part of GAO's shared culture as its commitment to product quality, timeliness, professional standards, and service to the Congress. Human resource management is the foundation of all of GAO's institutional values.

Dr. Peter Vaill, professor of human systems at George Washington University, discussed his research on the qualities of "excellent organizations" in his keynote address.

Other panel highlights were as follows:

- Jim Brucia, San Francisco Assistant Regional Manager, discussed human resource management from the perspectives of his role at GAO and as Director of Audits and Personnel in the Department of Energy's San Francisco Office.
- Dexter Peach, Director of the Resources, Community, and Economic Development Division (RCED); Kansas City Regional Manager Dave Hanna; and Civil Rights Office Director Alex Silva participated in a candid assessment of GAO's efforts to improve performance management.
- Three evaluators-in-charge, Skip Jenkins and Jim Noel of RCED and Len Baptiste of the Washington Regional Office, discussed their experiences at managing work through people using the Behaviorally Anchored Rating Scales (BARS) performance appraisal system.

The conference also included sessions on managing in a low-growth environment, personnel implications of the budget reductions, and difficult-employee issues. It closed with a commitment at all levels to make human resource management work in GAO.

For more information, contact conference moderator Judy England-Joseph, Office of the Assistant Comptroller General for Operations, at (202) 275-5495.

Review Feature Writer Receives Award

Ms. Judith Hatter, Technical Information Specialist—Law, received a special award on May 23, 1986, for her "consistently superior contributions to *The GAO Review*, as author of 'Legislative Developments' in every quarterly issue for the last 15 years."

*Ms. Hatter receives her special award from Mr. Havens, **Review** editor.*

See Location, p. 33

Manager's Corner

Managing Organizational Change

This feature was coordinated by Ross Laguzza, Management Development Intern, Office of Organization and Human Development.

This issue's "Manager's Corner" focuses on managing organizational change. As GAO faces external and internal changes, managers are continually challenged to recognize, implement, and effectively manage change.

The reviewers for this edition participated in a recent Office of Organization and Human Development (OOHD) training program entitled "Managing Personal and Organizational Change." Their experience in the program gives them a perspective from which to evaluate literature on organizational change and to examine organizational change in GAO.

Linda Morra, Assistant to the Director in the Office of Program Planning, and David Kass, an economist in the Office of the Chief Economist, reviewed chapters from *The Change Resisters* by George Odiorne. Dr. Odiorne was the featured speaker at GAO's June 1986 Executive Speakers Program. These reviews evaluate the significance of "organizational resistance" and examine methods of managing change, drawing from examples within GAO.

In another review, Thomas Slomba, Group Director in the Resources, Community, and Economic Development Division, examines the merits of a model for overcoming resistance to organizational change proposed by Richard Hermon-Taylor in the book entitled *Organizational Strategy and Change.* The review is a skillful application of material written for a private sector audience to public sector organizations.

"Managing Change in the World of Change Resisters" and "Managing Change by Keeping Your Options Open," *The Change Resisters*

By George Odiorne
Englewood Cliffs, New Jersey: Prentice Hall, 1981

Reviewed by Linda Morra, Assistant to the Director, Office of Program Planning

Odiorne in Brief

Ours is a world populated by change resisters, according to George Odiorne. But even though the forces are arrayed against change, as the author points out, change has occurred, is occurring, and will occur. The trick is to manage it. Toward that end, Odiorne offers us strategic management.

Strategic management, as presented by Odiorne, is inventing a future (either for the individual or for the organization) and making that future happen. It is similar to defensive driving—the rules are to watch all around you, be on the alert, and always have alternative plans. Odiorne delineates six specific steps to strategic management:

- Keep your range of vision higher than you do now.
- Get a bigger picture of the world than you now have.
- Do not stare at the world; scan it.
- Make yourself visible to others and make your opinions known.
- Always have alternative plans.
- Practice your skill in managing your timing.

If there is one key, however, to strategic management, Odiorne would say that it is keeping your options open. The more options, the better. In an organization, the strategy for doing this would be (1) specifying a problem as a deviation from a standard, (2) involving a lot of people early on to generate solutions, and (3) where time permits, arriving at a consensus for action.

Odiorne sees different types of options, defensive and offensive, for solving problems. The only type he believes will not work is an authoritarian exhortation to produce a change in response to a problem. He believes that resistance is the primary action that this option will generate. Other types, however, may work well or poorly depending on the problem, and each type of option should be reviewed. One defensive option, for example, is to do nothing differently. Odiorne asks, if we didn't do anything other than what we are doing now, what would be the consequences? He believes that if the problem is not really severe, this may be the best option.

Another option is to find a scapegoat. While most of us might generally consider this a bad option, Odiorne points out that it is a common successful defensive action. It often removes pressure from the organization as a whole and unloads guilt and blame at the same time! Reorganization, according to Odiorne, is also a great defensive option, especially when one needs to buy time. More offensive types of options are redefining something noble about a change (e.g., making a crusade) or encouraging people to fulfill their highest potential. However, Odiorne believes that the best option for change is one created by the people who must implement it or one for which the implementers can claim ownership.

One Reader's Evaluation

Odiorne is instructive, stimulating, and sometimes even shocking, but he is also somewhat difficult to follow. He could have made his points, complete with illustrations, in a more organized, concise, and straightforward manner. For instance, a

section with the heading "Getting a Bigger Picture of the World" starts with a long discussion of the self-fulfilling prophecy. While it may be a good description of the phenomenon, its placement doesn't fit. There is also a ring of the Horatio Alger success story. Odiorne tells stories about people who looked at the big picture, aimed high, acted, and became huge successes. I kept thinking about the Hunt brothers and the silver market and wondered how many stories there were of "defensive drivers" who still couldn't anticipate the blind side that got them!

Applications to GAO

Overall, Odiorne addresses issues that are highly relevant to GAO. Having just completed GAO's course entitled "Managing Personal and Organizational Change," I was particularly interested in comparing Odiorne's perspective on managing change with the "Concepts for the Management of Organizational Change" presented by the course consultants, the Delta Consulting Group. Both share a focus on managing resistance to change. Both emphasize motivating change by building participation into the change. The Delta Consulting Group goes futher in identifying other steps needed to motivate change, such as identifying and surfacing dissatisfaction with the current state, building in rewards for desired behavior, and providing time to disengage from the present state.

The focus of both on resisters and resistance to change seemed to me, however, to treat "change" too generically. That is, resistance to change may be a function of the perceived difficulty/disadvantages of implementing the change minus its perceived benefits/utility. In GAO, we are in the midst of many changes. How do we personally manage organizational changes? As individual managers, Odiorne has much to say to us. His defensive driving analogy and steps to strategic management seem excellent advice in an environment where events will occur beyond an individual's control and where the objective is clearly to avoid "injury" and maximize opportunities by anticipating events but keeping options open.

In GAO, our issue area planning system provides an example of where we are, in effect, following much of Odiorne's advice. We seek the big picture, through development of planning documents, to ensure that GAO's resources are applied to the most important issues over a 2- to 4-year period. We use experts to help us scan the world and identify issues. We also make ourselves visible and discuss issues with congressional committees and their staffs to ensure that our plans reflect congressional interests and priorities. Our objective is a good match of individual requests with planned work. However, while we anticipate to the extent possible, we keep options open so that we can respond quickly to changing developments and congressional interests. For instance, we put aside resources for a realistic proportion of unanticipated congressional requests.

In relation to institutional management of change, however, it seems that we should first consider the perceived advantages versus the disadvantages of implementation to avoid overmanaging for change. For example, improving overall responsiveness to congressional requests has been a major GAO thrust since a study group reported its findings in the summer of 1985. One specific finding was that we needed to involve congressional members and staff more in determining the work we will do, and the policy was emphasized that responding to congressional requests was GAO's top priority. With no further formal directive, the congressional request assignment rate has jumped dramatically, and a large proportion of the requests reflect joint planning. While other factors may have contributed to this increase, GAO staff, it seems, responded strongly to the perceived need for change. In this case, the perceived benefits of the change and the ease of implementing it have apparently made many of the strategic management steps identified by Odiorne unnecessary. While GAO management is keeping close watch on the change and its effects, management did not need to "orchestrate" the change to make the desired future happen.

Concluding Comments

In brief, Odiorne has much to offer us with regard to managing change. But management of organizational change can be carried to an extreme, and the need to initiate change should be balanced against its perceived benefits and costs.

"How Bureaucracy Makes Cowards out of Heroes," *The Change Resisters*

By George Odiorne
Englewood Cliffs, New Jersey: Prentice Hall, 1981

Reviewed by David I. Kass, Economist, Office of the Chief Economist

As the titles of both this chapter and book imply, the author takes a dim view of bureaucracies. The larger an organization, the more lethargic will be its response to innovation. He believes that as a bureaucracy grows, so will managerial layers. Major decisions will then require additional approvals and, therefore, more time. Individual initiative is discouraged, and a "bureaucracy inevitably produces an organization of pipsqueaks."

Are smaller organizations necessarily or even usually better managed, more flexible, and more responsive to needs of their customers than larger ones? From my own experience, I would unequivocally say "no." I do not believe that being bigger necessarily implies a reduction in managerial or organizational efficiency.

GAO Links

How do the author's views apply to GAO? To provide some perspective, it might be useful to consider GAO's budget and the size of its staff. In fiscal year 1985, GAO had a budget of $290 million and emloyed about 5,000 people. In contrast, the Department of Agriculture employed 21 times that number, or 106,000, and had a budget of $55.5 billion. Similarly, the Veterans Administration employed 221,000 people (44 times the number employed by GAO) and had a budget of $26.3 billion. On the other hand, the Congressional Budget Office, with a staff of 200 and a budget of only $16 million, and the Office of Management and Budget, with a staff of 600 and a budget of $39 million, are considerably smaller than GAO. Thus, compared with other offices of the U.S. government, GAO might be considered a medium-sized organization.

Odiorne suggests that GAO would probably be less well managed than much smaller regulatory agencies but better managed than larger cabinet-level departments. Is this true? My response is that size alone is a poor indicator of managerial quality. Small organizations do not have a monopoly on innovative, open-minded, and responsive managers. Indeed, many small organizations stay small (or disappear) in the private and public sectors because they lack the managerial talent to properly motivate their staffs, satisfy their customers, and plan for the future. Finally, organizations that become large usually attain their size because they have achieved their goals (e.g., satisfying their customers). An example of what is universally accepted as a well-managed but very large organization is IBM, which had 400,000 employees and $50 billion in revenue in 1985.

Skill Matches

I certainly do believe, however, that there is more than a grain of truth in the author's hypothesis that large firms or agencies can and do stifle the individual. I do not doubt that this occurs and may even be valid for a majority of large organizations. Nevertheless, I do disagree with the impliction that it *must* happen.

Although Odiorne acknowledges that one of the major strengths of a bureaucracy is the efficiency to be gained through division of labor, he also attributes resistance to change to this organizational form. The author does not seem to appreciate, however, that division of labor (such as the way GAO is organized) represents an efficient allocation of resources by matching an individual's knowledge, skills, and abilities with the needs of a specific group (or division). I think that when an employee's needs and strengths are closely matched with those of the organization, the employee is likely to be more highly motivated and far more receptive to change.

Odiorne also states: "What bureaucracies don't realize is that when people are always treated as responsible, they will do better things. . . .Censorship and suppression are powerful forces against originality. . . ." Once again, the author assumes that *all* bureaucracies neccesarily stifle the individual. A responsive management that is sensitive to the needs of its staff, is interested in their professional growth, readily delegates significant responsibility, and encourages originality will not fall into the author's stereotype, regardless of the organization's size.

Concluding Comments

To sum up, I strongly disagree with Odiorne's negative characterization of large organizations. Each organization has its own strengths and weaknesses. But an enlightened management that appreciates that its most valuable asset is its people, that properly motivates its staff and is sensitive to matching their needs to those of the organization, and that encourages personal and professional growth, along with open communication, is well equipped to implement organizational change.

"Finding New Ways of Overcoming Resistance to Change," in *Organizational Strategy and Change*

By Richard Hermon-Taylor; J. Pennings, ed. San Francisco: Jossey-Bass, 1985

Reviewed by Thomas E. Slomba, Group Director, Resources, Community, and Economic Development Division

How can organizations successfully accomplish change, even when change threatens to violate the cultural constraints of the organization? This chapter proposes a possible answer to this complex question in the form of "a new model of change." The author is justifiably cautious in proposing the model and notes that it is based on "very limited experience with a handful of organizations."

Change Factors

Before presenting his model of organizational change, the author addresses a wide range of factors that can affect an organization's ability to successfully implement change. He does an excellent job of introducing factors such as communication, the nature and level of change, management style, management reasoning processes, and compromise. Presenting this wide perspective, however, forced the author to deal with very complex issues in simple terms. The reader should be careful not to oversimplify the concepts in this summary. In addressing level of change, for example, the chapter introduces the concept of "frame breaking" change. This level of change typically requires embedded behavior to be dramatically altered. Determining what is "frame breaking" as opposed to acceptable change for an organization is not easy, and the author points out that this determination depends on the size of the organization, its degree of bureaucracy, its leadership, and a host of other variables. What is a "frame breaking" change for one organization or unit may be acceptable change for another.

This chapter is written from a private sector perspective and addresses concepts such as competition, market growth, profitability, and return on investment. Nevertheless, I believe the ideas relating to organizational change apply to all organizations, including the public sector. Managers from the public sector will have to make adjustments for some of the terminology and examples presented.

Private Sector Links

This chapter's applicability to the public sector becomes clear when the proposed model of change is examined. It focuses on the change process in terms of diagnosis, formulation, and execution.

The model proposes that if change is to take place in an organization, the need for change must first be recognized or diagnosed. According to the author, this diagnosis should be made periodically. The diagnosis phase of the model is aimed at answering questions such as the following. Is the existing strategy working satisfactorily? Are there unperceived opportunities that suggest the need for change? What alternatives are available to the organization?

In the formulation phase of the model, management reviews the change identified in the diagnosis phase. This review includes analyzing alternatives and identifying organizational resistance.

During the execution phase, a plan for implementing the desired change is developed. This plan addresses ways to overcome any organizational blockages identified during the formulation phase.

In this model, these three phases are aimed at determining what an organization *should* do and what it *can* do.

The author suggests the need for an organizational "change agent" to perform the diagnosis and to help implement the change. The chapter outlines how to identify an appropriate change agent and the behaviors an agent must exhibit. One important characteristic of the change agent proposed is that he or she is an outsider. The author believes that for a member of the organization to effectively diagnose the need for change would probably prove impossible. I do not share the author's view on the need for an outside change agent.

GAO Program

My view is based, in part, on GAO's success in identifying the need for change through its Operational Improvement Program. This program largely parallels many of the diagnosis, formulation, and execution aspects of the Hermon-Taylor model. It provides a means for needed change to be diagnosed; a vehicle to present identified changes to management for review, analysis, and evaluation; and a strategy to implement changes. This program does not, however, rely on an outside change agent to perform the diagnosis and to help implement change. The change agent in the GAO program is the entire staff. The program's success in diagnosing areas needing change is evidenced by the 25 projects now being tested, for example, placing signature authority for reports with associate directors consistent with their in-

See Manager's, p. 33

Topics in Evaluation

Carl E. Wisler

Mr. Wisler is an Associate Director in GAO's Program Evaluation and Methodology Division

This issue's topic is qualitative data analysis

In 1981, the Polish government declared martial law and took action to suppress the Solidarity trade union. The United States responded by announcing a number of sanctions against Poland, including the end of U.S. Export-Import Bank credit insurance and the suspension of high technology equipment sales. For the next few years, the two nations played out a series of actions and counteractions on the international stage, each hoping to prevail over the other or at least to end the episode without serious loss.

Exchanges like those between Poland and the United States illustrate the use of economic sanctions, foreign policy instruments intended to affect national behavior. Debates rage among foreign policy analysts and historians about the effectiveness of such sanctions and with good reason, because it is often hard to precisely describe international events, let alone account for why they happen. This article explores the possibilities of qualitative data analysis and uses economic sanctions for illustration.

Qualitative Data Analysis

The phrase "qualitative data" does not mean the same thing to everyone. As frequently used by statisticians and some social scientists, "qualitative data" refers to information which can be classified or categorized. Examples include a person's marital status, another person's attitude toward South Africa, or a movie's rating. Though such information is called qualitative data (or better, categorical data), it can be analyzed statistically using techniques such as log-linear models and latent structure analysis.

Other researchers have something else in mind when they refer to qualitative data. They mean information in the form of words. In this sense, Margaret Mead's data base on Samoan youth was mostly qualitative. Traditionally, disciplines which have depended heavily on this form of qualitative data have produced huge amounts of information, analyzed the data in somewhat mysterious and nonreplicable ways, and reported the results in long narratives. Anthropology, enthnography, and some forms of sociology are examples. In the remainder of this article, qualitative data means information in words, not numbers.

Compared with the anthropologists and early sociologists, analysts in the more policy-oriented sciences, such as evaluation, social psychology, and political science, have developed more structured ways of analyzing qualitative data and more concise ways of reporting the results. Similarly, computer scientists have produced software for handling qualitative data. Our attention will be focused on some of these more recent developments.

Qualitative Data Bases

Suppose that we are in the first stages of analysis and that we just want to describe economic sanctions and to find out what factors are associated with one another. With quantitative data, we would start by organizing a data base, usually tables of numbers, and subsequently compute descriptive statistics, like means and ranges. To learn about relationships, we would probably compute correlation coefficients or other measures of association.

With qualitative data, we also need to organize our data base, but in general, we have to be much more flexible because we'll be dealing with words rather than numbers. Because we can't compute statistics, we need to use other means to summarize and display information. And finally, to draw conclusions, we must use some form of reasoning other than statistical inference.

The raw data in studying economic sanctions might consist of official government documents, public statements by government officials, accounts in the media, and personal interviews with participants in the actions. A data base, more succinct than the original documents but still in narrative form, can be built from the available information. Figure 1 shows a small segment about the conflict between the United States and Poland from such a data base.

Figure 1. A Portion of a Qualitative Data Base (unstructured)

In 1981, the Polish government declared martial law and took action to suppress the Solidarity trade union. The United States responded by announcing a number of sanctions against Poland: the end of Export-Import Bank credit insurance, the suspension of high technology equipment sales, the suspension of Polish fishing rights in U.S. waters, and the suspension of Polish airline landing rights. The goals of the United States were to persuade Poland to lift martial law, free union detainees, restore free speech. . . . Poland later outlawed Solidarity but freed Lech Walesa. Eventually the episode ended with decline of trade, martial law suspended, political prisoners granted amnesty. . . .

However, the United States-Polish confrontation is but an isolated case. In general, we will be able to understand economic sanctions better if we look at more than one instance. Therefore, we want our data base to consist of multiple cases, and because there have been well over 100 cases in the last 70 years, we can imagine a voluminous compilation. (Hufbauer, Schott, and Elliott, 1985)

A qualitative data base certainly has some similarities to a quantitative one in that it contains information about one or more cases that we want to understand in some sense. However, the differences are typically great: The data base consists of a lot of words instead of or in addition to numbers; the variables may not be sharply conceptualized much less defined; and in general, we just don't have the structure implied by a numerical data base (at least to begin with).

So suppose that we have thousands of lines of information like those in figure 1. We now face a central problem of qualitative data analysis. How can be bring order to an ill-structured word-oriented data base? How can we analyze the data? Our first choice is between staying in the qualitative mode or using a technique such as content analysis to transform the data into quantitative forms. Since our aim is to discuss qualitative analysis, we won't consider the feasibility of converting the data to numbers or the pros and cons of doing so.

So if we stick with qualitative data, we need some ways to structure the information and ways to analyze it. The suggestions to be offered must be regarded as tentative because not much attention has yet been given to methods for handling and analyzing qualitative data.

Manipulating the Data

To analyze data, we must be able to display it, move it about, compare it, and summarize it. When data are in the form of numbers or well-defined words, such as names and addresses, computerized data base management systems provide an efficient way to organize and manipulate the information. Systems for microcomputers, such as dBase III or PFS:File, were generally designed for business applications where the data tend to be well structured. Consequently, most microcomputer systems are cranky about requiring that data base elements like records, fields, and field types be set forth explicitly. Qualitative data of the type illustrated by the economic sanctions are less structured, however; the anthropologist's field notebook is the quintessential example.

Until recently, the qualitative analyst could either organize the data on paper (and do the analysis by hand) or force-fit the information into a structured data base management system. Now a third possibility exists. Some computerized data base management systems, such as askSam (Access Stored Knowledge via Symbolic Access Method) are more free-form and in general are better suited to data which are word based rather then number based. With askSam, data can be entered simply as a string of words without regard to what the variables might be. In the jargon of data base systems, this means that we don't have to specify fields, field widths, field types, and so forth. For example, the data in figure 1 can be entered into askSam just as they are displayed. Although we will eventually want to impose some structure on the data base, a flexible, free-form data base management system permits the user to construct the data base without planning every detail in advance (like choosing the variables and deciding how many characters to allow for each one) and permits easy alteration of the variables during the analysis stage. (AskSam also permits the usual preplanned structure of variables if that is desired.)

Once the data base is entered, askSam can perform the usual manipulations, such as retrieving cases which have specified characteristics or sorting the cases according to one or more criteria. For example, we might want to retrieve and print out information about all instances in which economic sanctions involved the freezing of foreign assets. AskSam searches through the cases (known as records in data base systems) to find those which refer to the "freezing of foreign assets." This step is analogous to a word or a phrase search on a word processor, except that with askSam, the result is the retrieval of a set of cases which include the key word or phrase. Also, the search can be done with various logical operations, such as finding all cases with word1 *or* word2, word1 *and not* word2, and so on.

As we begin to work with the data base, we will see how we want to structure it for better understanding. With a system like askSam, this structure can be imposed, using the data base editor, as we go along. That is, we don't have to completely

Figure 2. A Qualitative Data Base (partially structured in askSam format)

In DATE[1981], the Polish government declared martial law and took action to suppress the Solidarity trade union. The SENDER[United SENDER[States] responded by announcing a number of sanctions against TARGET[Poland] SANCTIONS[the end of Export-Import Bank SANCTIONS[credit insurance, the suspension of high SANCTIONS[technology equipment sales, the suspension of Polish fishing SANCTIONS[rights in U.S. waters, and the suspension of Polish SANCTIONS[airline landing rights] The goals of the United States were to persuade Poland to GOALS_OF_SENDER[lift martial law, free union GOALS_OF_SENDER[detainees, restore free speech. . . .] Poland later RESPONSES_OF_TARGET[outlawed Solidarity but freed Lech Walesa]. Eventually the episode ended with. END_STATE[decline of trade, martial END_STATE[law supended, political prisoners granted amnesty. . .

redefine and reconstitute the data base each time we gain some insights into useful ways to analyze the data. We might now decide, for example, to define the following variables: DATE the sanctions were first "sent" in an effort to influence behavior, the SENDER country, the TARGET country, the GOALS OF the SENDER, the SANCTIONS imposed, the RESPONSES OF the TARGET, and the END STATE of the episode involving economic sanctions. The appropriate variable labels are simply inserted into the data base, and it then looks like figure 2.

Now it is possible to perform new kinds of data manipulations and to develop displays which show all the cases in terms of the variables we have defined. A simple display is shown in table 1 where variables are displayed for two cases: United States-Poland and United Kingdom-Argentina.

Table 1. A Cross-Site Matrix

Variable	Case 1	Case 2
SENDER	United States	United Kingdom
TARGET	Poland	Argentina
GOALS OF SENDER	Lift martial law, free union detainees, restore free speech. . . .	Remove Argentines from Falkland Islands, restore United Kingdom administration. . . .
SANCTIONS	End Export-Import Bank credit insurance, suspend high technology equipment sales, and suspend Polish airline landing rights.	Freeze Argentine assets in United Kingdom, freeze official export credits, freeze United Kingdom bank loans to Argentina. . . .
RESPONSES OF TARGET	Outlawed Solidarity but freed Lech Walesa.	Froze United Kingdom assets in Argentina, suspended repayment of debt to United Kingdom. . . .
END STATE	Trade declined, martial law suspended, political prisoners granted amnesty. . . .	Trade declined, Falkland Islands regained by United Kingdom. . . .

Now suppose that we have the data base organized so that we can manipulate the information, either by an old-fashioned method, such as file cards, or by a free-form computerized system, such as askSam. What data analysis methods are available to summarize the data and to draw inferences from it?

Data Analysis Methods

If our data were numerical, there would be many standard ways of analyzing and displaying it. We would have descriptive statistics to summarize the data; scatterplots to display relationships; and more complex techniques, such as regression analysis, for looking at relationships and drawing statistical inferences. Methods for analyzing qualitative data are far less standardized and not well-known. However, a recent book by Miles and Huberman (1984) provides a rich array of 49 specific methods and many variations for analyzing nonnumerical data.

A principal problem in analyzing qualitative data derives from the sheer mass of information. Because usually much more data is available than the mind can keep track of or draw inferences about, we need to boil the information down without losing essential elements and to display the results so that patterns and relationships will be evident. Many methods described by Miles and Huberman are verbal analogues to the exploratory data analysis methods developed for statistical data by Tukey (1977), ways to tease meaning out of complex information.

Although most qualitative data analysis methods are best understood by looking at data on large sheets of paper or "paging through" a computer data base, one of the Miles and Huberman methods will be introduced here. The matrix is one of their key displays, and the authors offer many variations. One is called the case-ordered descriptive matrix, and in terms of structuring the data, it goes one step beyond the matrix in table 1. Suppose we have 100 cases involving the use of economic sanctions and we are interested in more than just describing the episodes; we are interested in discovering any relationships which may exist between the END STATE of the episodes and the types of sanctions used.

If we can in some way characterize the END STATE of the episode in terms of degree to which objectives were achieved, then the cases can be ordered. Hufbauer, Schott, and Elliott used judgment by experts to score each sanctions episode on a simple four-point success scale. (We are starting to cross the borderline between qualitative data analysis and categorical data analysis.) Using index cards or askSam, we could then reorder our 100 cases and look for patterns which may now appear showing the relationship between the degree of success and the types of sanctions used. Certain types of sanctions might tend to be more associated with success than others. If that is the case, we could go to use other methods, such as the case-ordered (also called site-ordered) effects matrix or causal networks that Miles and Huberman suggest, as ways to search for cause-and-effect relationships. In a case-ordered matrix, the cases are rank-ordered by degree of probable cause, and in a causal network, possible cause-and-effect relationships are depicted graphically.

For More Information

askSam: A Free-Form, Text-Oriented Data Base Management System. Perry, Fla.: Seaside Software, 1985. Version 2 software and manual available for the PC and other DOS machines. A more powerful Version 3 has been announced for late summer 1986.

Bogdon, R., and S. Taylor. *Introduction to Qualitative Research Methods.* New York, N.Y.: John Wiley, 1975. A widely used textbook on qualitative methods.

Case Study Evaluations: A Transfer Paper. Washington, D.C.: U.S. General Accounting Office, Program Evaluation and Methodology Division, forthcoming. The case study design frequently leads to qualitative data analysis.

Cook, T. D., and C. S. Reichardt, eds. *Qualitative and Quantitative Methods in Eval-*

See Topics, p. 33

Controlling the Quality of Electronic Workpapers

Stewart O. Seman

Mr. Seman, an evaluator in the Chicago Regional Office, has been with GAO since 1965. He received a B.S. in accounting from the University of Dayton in 1965 and an M.B.A. in systems from DePaul University in 1977. He is a certified public accountant and a member of the American Institute of Certified Public Accountants. He is also a member and past officer of the Association of Government Accountants. He has received several GAO awards and in 1985 received an Outstanding Achievement Award.

Ed. note: As discussed in "On Location," page 3 , GAO has taken a number of steps to train and assist staff interested in using microcomputers. Initial guidelines entitled *Using Micro Computers in GAO Audits: Improving Quality and Productivity* (Information Management and Technology Division Technical Guideline 1, Mar. 1986) have been issued. Sharing information and experiences among GAO users is likely to be a major vehicle for ensuring appropriate use of these tools. The following article takes a detailed look at some problems that can arise in using electronic spreadsheets and provides valuable suggestions on how to avoid problems and use the power of the computer programs to help ensure the quality of the results.

Warnings of high risks associated with microcomputer-generated data are becoming more frequent in the professional and microcomputer literature. Stories of major miscalculations and decisions gone awry have accompanied these warnings. One firm, for example, fired several of its executives who recommended, on the basis of erroneous microcomputer spreadsheet computations, what turned out to be a bad acquisition. It also dropped a major accounting firm that had approved these computations. Such stories illustrate the continuing importance of rigorously applied quality control for expert microcomputer users, as well as nontechnical neophytes. Because microcomputers can handle more data faster than can be processed manually, they hold the potential to make bigger mistakes faster. Microcomputer use is steadily increasing in GAO due to this ability to quickly process data. Thus, now is a good time to consider the measures needed to control the quality of electronic workpapers.

Here "electronic workpapers" refers to all the data recorded and analyzed in electronic form with a microcomputer. This includes word-processor-generated documents, as well as information generated with data base manager software, such as dBase III, and electronic spreadsheet software, such as Lotus 1-2-3.

The summer 1986 issue of *The GAO Review* carried an annotated "Microcomputer Quality Assurance Bibliography," which listed and briefly described a number of articles on detecting and preventing spreadsheet error. Most of them were drawn from microcomputer periodicals and thus reflected a nonevaluator's perspective. Still, they provide an excellent reference for those interested in applying the mechanics of control features.

This article will identify techniques that may be used to ensure the accuracy of electronic workpapers. Although the article will focus on electronic spreadsheets, it will also address the applicability of these control techniques to other software. Quality assurance is hardly a new subject in GAO; therefore, readers are already familiar with many of these techniques. In addition, the article will describe how the microcomputer's power can help ferret out errors.

Conceptually, electronic workpapers are not much different from the manual workpapers to which we are accustomed—only the medium is different. Therefore, we should be able to control the quality of electronic workpapers as we do for manual workpapers. In practice, however, it may not be quite as simple. Take several columns of numbers displayed on a video screen. Add a bunch of formulas—visible only individually—that refer to cells labeled "C81" and "X118." Toss in some macros (stored sequences of keystrokes that can be executed by pressing one key or a combination of keys) and you have the kind of spreadsheet that should give

any supervisor cause to ponder and might create some discomfort.

This uneasiness is well warranted. Errors can occur in data entry through alteration of the data (with no telltale erasure marks left behind), incorrect manipulation of the data, and software/hardware malfunction. Except for the last category, these same sorts of errors are common to manual workpapers. What is unclear is how to apply the manual workpaper standards with which we are familiar to the electronic environment.

The control features applicable to microcomputer-generated products fall into the following four categories:

- documentation,
- data entry controls,
- processing controls, and
- output controls.

Documentation

Documentation is a description of how the microcomputer analysis was performed. Electronic and manual workpapers require the same basic information (i.e., purpose, preparer, source of the data) as described in chapter 18 in GAO's *Project Manual.* But with electronic workpapers, the job doesn't end there. Appendix I to chapter 18 requires information sufficient for a referencer or a reviewer to duplicate the work done.

Documentation requirements are comprehensively discussed in *Using Micro Computers in GAO Audits: Improving Quality and Productivity,* Information Management and Technology Division Technical Guideline 1. Also, "Suggestions for Quality Assurance in Microcomputer Products," June 1985, prepared by a group made up of managers of GAO's Technical Assistance Groups, describes various required documentation for prepackaged and customized microcomputer software. These suggestions cover a wide range of automatic data processing (ADP) control techniques that are very helpful in working with data base managers and some of the other analytic software.

However, electronic spreadsheet documentation warrants extra attention. Problems with spreadsheet software have received the lion's share of attention in the media, not just because of the spreadsheet's popularity, but also because of its very design. Electronic spreadsheets tend to grow in an undisciplined fashion. The user may, for example, start by adding a column of numbers.

Once these numbers are entered, the user may find that a lot of time can be saved by adding formulas and other data. Either a spreadsheet Picasso or a nightmare can be created in the process, depending on the care taken in leaving "tracks" for others who will use the data. The purpose of documentation is not to inhibit this creative process, but rather to ensure that this creativity produces reliable results.

Two areas of spreadsheet documentation present concern: first, the information in the worksheet itself (internal documentation) and, second, the need for some form of hard copy to accompany the information stored on the microcomputer disk (external documentation).

Internal Documentation

Opinions differ on the best layout for a spreadsheet. They generally agree, however, that separate areas should be assigned for the worksheet description (documentation section), data input, formulas, computation, and report (results). Figure 1 illustrates such a layout. The amount of space devoted to each component will vary with the application, as will the number of components. For instance, in some cases, the information in the data input and processing areas will be printed as the report, thus eliminating the need for a separate reporting section.

The structure also lends itself to a modular, step-by-step approach to spreadsheet construction and verification. That is, data are input and checked for accuracy before they are manipulated, formulas are constructed and tested before they are used to manipulate the data, and so forth.

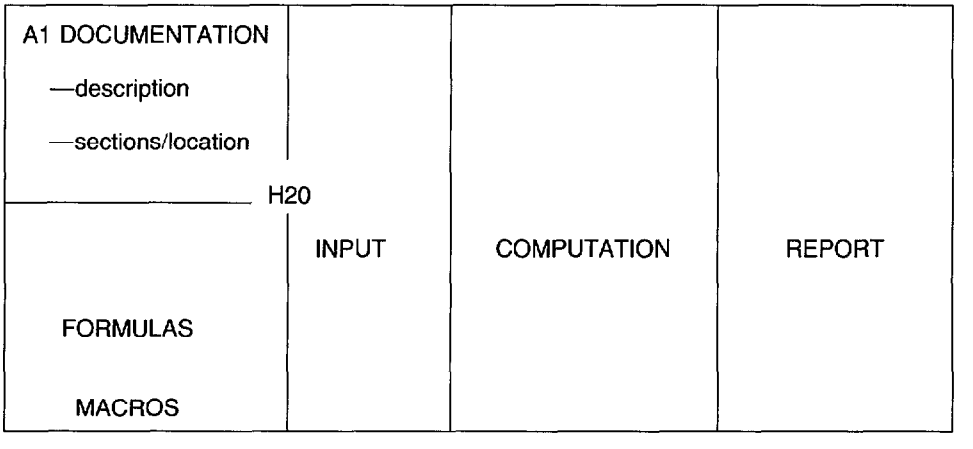
One hint on documentation is to place the key information in the top left-hand corner of the spreadsheet. One author suggests reserving the full "home screen" area for this purpose. In a spreadsheet, data are entered into cells, which are designated by the intersection of numbered rows and alphabetical columns. For example, the upper left corner of a spreadsheet would be cell A1 and the lower right corner would be cell H20. This area is analogous to both the usual worksheet heading and a table of contents. (See fig. 1.) The area identifies the spreadsheet and states its purpose, preparer, and date of preparation. In addition, it identifies where the various components can be found after the viewer leaves the home screen. Thus the statement "Case Sample Data = I1 to N155" indicates that the data are in the block starting with column I, row 1, and ending with column N, row 155. (Fig. 2 shows a section of a spreadsheet.) Anyone who has scrolled around a spreadsheet hoping to locate a specific segment will appreciate such a "road map."

External Documentation

In practice, the external documentation may be a printout of the spreadsheet documentation section or a printout of the entire spreadsheet. As spreadsheets grow in complexity, are used on multiple assignments, and are kept for permanent files, an audit trail is needed to describe their contents and function. While it is not necessary to print the entire spreadsheet to document it, the external documentation should include the record counts and key totals needed to authenticate the information in the electronic medium. It is also the place for the preparer and the supervisory reviewer to sign off on the work.

In some cases, other staff members may make subsequent changes to the file. Software or hardware malfunctions may also alter the data. Control figures are one mechanism for identifying such changes.

Figure 1. Electronic Spreadsheet Structure



Moreover, a detailed hard copy description is needed of each disk's contents. While we, as a matter of course, provide tables of contents that describe what is in our workpaper bundles, we frequently do not devote similar attention to the contents of disks. The meager information supplied by the operating system's directory command or the dedicated (Micom) word processor's index command may suffice for file management, but most folks agree that an 8- to 12-character file title is hardly enough to adequately identify a disk's contents. From a practical standpoint, a lot of time can be wasted trying to decipher what a file titled "DCIDAT.WK1" is supposed to contain. If this file title is accompanied by a one-paragraph description of the file's subject, purpose, etc., deciphering is not needed.

Data Entry Controls

Input controls generally consist of front-end edits, control totals, footing and cross-footing the results, and data verification either completely or on a spot-check basis. A brief description of each technique follows.

Front-end Edits

These are checks of the data as they are entered. These checks vary with the capability of the software. Most data base managers enable the user to define fields as either alphabetic or numeric. More-sophisticated software may allow the user to impose range restrictions on the data (e.g., values between 10 and 50).

Spreadsheet software generally does not have this sophisticated front-end edit capability. However, data base manager software can be used for data entry and the data then transferred to the spreadsheet. Data base manager software is especially useful for large-scale data entry tasks, since it has the added advantage of the ability to tailor "data entry screens." This gives the user a form to enter the data and in effect walks the user through the data entry process. It thus avoids cursor movement errors and permits use of personnel with minimal ADP skills for data entry.

Control Totals

These are established before the data that are entered into the spreadsheet or data base can be compared with the spreadsheet, or data base, totals. Record counts, data totals, and hash totals can be used for control purposes, as follows:

- Record counts are simply a count of the items entered. (See fig. 2.) On the Department of Transportation management review, for example, record counts for data transferred from a data base manager to a spreadsheet differed. We traced the error to a hardware problem we were having with the microcomputer we were using. The same spreadsheet on a different microcomputer gave the correct results. Using control totals enabled us to detect the error and thus saved us a lot of potential embarrassment.
- Data totals—the sum of numeric data, such as dollars or quantities—are a routine manual workpaper technique made easier by the microcomputer's number-crunching capabilities. They are usually calculated to check the correctness of the source data, as well as our own data.
- Hash totals are the sum of the numbers in a data field, such as social security numbers, which are not normally added. For instance, the "Grade" column in figure 2 was totaled by the electronic spreadsheet. This total (1,632) agrees with the total for the source records.

Figure 2. Electronic Spreadsheet Example—Input Section

	I	J	K	L	M	N
1	Cert.			Date	Number	Number
2	number	Series	Grade	requested	certified	appointed
3						
4	850097	GS-334[a]	12	10/16/84	4	2
5	850283	GS-1102	9	11/12/84	3	1
6	850921	GS-510	14	02/08/85	2	1
151	851839	GS-690	11	07/08/85	1	2
152	842246	GS-510	12	09/03/85	3	1
153	852367	WG-421[b]	9	09/24/85	3	1
154						
155	Total	150	1,632		390	203

Record count

Hash total

[a]General Schedule.
[b]Wage Grade.

Footing and Cross-footing

This procedure involves adding numbers in a spreadsheet both horizontally and vertically. It is a hallowed GAO tradition. A spreadsheet can be programmed to calculate and compare these totals and signal with a beep if they do not agree. This might seem like a case of overkill, but it can be quite effective in calling attention to a problem in a large spreadsheet where the result is out of view and may be overlooked.

Data Verification

This procedure is a comparison of what is on the computer with what is in the source document. It can be performed on the whole document or on a spot-check basis. Data can also be verified by reentering them and using the software to compare the results.

Processing Controls

Although the particular software may not permit range restrictions during data entry, these controls can be applied afterward. One of the most popular components of the Chicago Regional Office's workpaper training course has been "Can You Spot the Errors in This Spreadsheet?" The participants invariably found more errors than the author intended. What's really important, however, is the approaches—range checks and reasonableness checks—that they used to identify these problems. These techniques work well with spreadsheets, since they permit the review process to be automated.

Using range and reasonableness checks requires the user to determine the basic characteristics of and relationships between the data elements being examined. For example, in figure 2, our evaluators would know that the certificates are arranged by fiscal year, with the first two digits of the certificate number representing the fiscal year, and that "mid-level" refers to the GS-9 through GS-12 (characteristics). They would also know that the number appointed cannot exceed the number certified.

Depending on the software, the electronic workpaper can be examined for these characteristics using the search, sort, or logic analysis features. For example,

• Almost all the microcomputer software GAO uses for analysis has a search or query capability (e.g., the computer can be instructed to find the word "evaluator" and it will stop at the first occurrence of the word). Thus, if one of the characteristics of the data in figure 2 is that they should not include the "GS-718" series, this characteristic can be used as the criterion to search the workpaper.

• Entering the letter "l" for the number "1" is one of the most frequent mathematical errors on our dedicated and microcomputer-based word processors. This error is difficult to catch on the video display and a printout. The word processor's search function can be used to find and replace the errant figures in short order.

• The schedule in figure 2 should include only GS employees. While it is easy to spot the inappropriate WG employee in this illustration, the sort feature allows the user to quickly locate similar errors in a much larger selection.

• Spreadsheet and electronic data base software generally can identify data inside and outside a certain range. Figure 3 illustrates the formula that would flag with an "ERR" statement those certificates in our sample that are not mid-level. Figure 4 illustrates what occurs when this formula is applied to the data to be checked. In a spreadsheet, this process may be automated by using a macro. Figure 5 shows the formula devised earlier to flag certificates that were not for grades 9 through 12, converted to a simple macro.

Figure 3. Formula to Find Case Selection Errors in Column K, "Grade"

```
@IF(K4<9#OR#K4>12,@ERR,1)
```

Note: This is the formula as it would be written for the first row of values (row 4).

Reasonableness tests deal with the relationship between variables. In figure 2, column N, the number appointed should exceed the number certified shown in column M. On a small spreadsheet, the reasonableness test can be done manually. But on a more extensive spreadsheet, the evaluator may choose to use a macro. In either case, care must be taken to distinguish between data entry errors and errors in the source document.

Figure 4. The Result When the Formula in Figure 3 is Applied to the Values in Column K (See col. L)

	I	J	K	L
1	Cert.			Error
2	number	Series	Grade	check
3				
4	850097	GS-334	12	1
5	850283	GS-1102	9	1
6	850921	GS-510	14	ERR
////	////////////////////////////////			
151	851839	GS-690	11	1
152	842246	GS-510	12	1
153	852367	WG-421	9	1

Control totals verified at input should agree throughout the processing. Where records are dropped, new control totals should be established with a clearly documented link to the original totals.

the formulas can be entered using the range names instead of the cell references. Thus, if range names were used in the previous example, the formula could have been entered as "number certified/total." Another precaution is to pretest the formula with a range of values representing all possible situations to ensure it yields the correct results.

Also, a formula to be applied to each number in a column of numbers can be placed in a macro, tested with a range of values, and then applied to the whole column if it functioned properly. Figure 5 is an example of such a macro, along with the documenting comments that should accompany it.

Formulas and macros should always be tested on a copy of the spreadsheet. A macro running amok through live data can create more problems than it solves. Care must be taken, however, to ensure that the final spreadsheet has all the corrections. This is where a systematic methodology and good documentation techniques can help.

Figure 5. Keystroke Macro to Check the Grade Column

	A	B	C	D	E	F
40	E				RANGE CHECK MACRO	
41						
42	{go to} K1~				Go to column to be checked.	
43	{right}				Move right one column.	
44	/WIC~				Insert error check column.	
45	^Error {down}^Check{down}{down}				Title column "Error check."	
46	@IF(K4<9#OR#K4>12,@ERR,1)~				Enter error-checking formula.	
47	/C~L5..L153~				Copy formula for all items to be checked.	
48						

Formula Errors

Using formulas in spreadsheets probably accounts for the most problems. Errors may be made in formula construction because of incorrect cell references, incorrect constants, or mistakes in logic or command structure. A formula also may be copied incorrectly to other areas of the spreadsheet. For instance, the user who entered the formula +M155/J155 intended to compute the average number certified by dividing the number certified by the total in our case sample. This formula actually represents the number certified divided by the grade hash total. (See fig. 2.) When it is written out in words, it is obviously nonsense. In any event, formulas deserve special attention, and most authors agree that they warrant the most scrutiny in spreadsheet preparation and review.

Several precautions may be taken to decrease the potential for formula errors. One is the range name feature. This assigns a name to a cell or a rectangular group of cells. Once the cells are defined,

Output Controls

Output should also be reviewed carefully. In GAO, the results are reviewed by the supervisor and, often, by the referencer and the technical assistance staff. The documentation and design techniques described earlier can greatly expedite this process. If the microcomputer application has been well documented, it can be reviewed for logic and tested for correctness. If a modular approach has been used to develop the application, supervisory review can be concurrent with spreadsheet development. If subsequently the spreadsheet is significantly changed, it is important that the changes also receive supervisory review.

See Electronic, p. 33

An Inside Look at the Office of Personnel Management's Women's Executive Leadership Program

Aletha L. Brown

Ms. Brown is an evaluator in the National Security and International Affairs Division, as well as a 1986 graduate of the Office of Personnel Management's Women's Executive Leadership Program. She joined GAO in 1979. She received a B.S. in business management from the University of Maryland and an M.A. in management and supervision from Central Michigan University. Ms. Brown is a member of GAO's chapter of Blacks in Government and received the national organization's Meritorious Service Award in 1985. She has received a Certificate of Appreciation, an Outstanding Achievement Award, the Comptroller General's Equal Employment Opportunity Award, and a Special Commendation Award. Ms. Brown was nominated for selection to the Outstanding Young Women of America in 1985.

On January 10, 1986, 62 women graduated from the Office of Personnel Management's (OPM) Women's Executive Leadership (WEL) Program. This ceremony completed a 15-month pilot effort designed to (1) develop management skills of potential managers, supervisors, and executives and (2) increase the number of women executives in the upper levels of government. I am a WEL graduate, and this article describes some of my experiences in the program.

The number of women in the civil service professional ranks has increased significantly since 1972. That year, Executive Order 11246 called for affirmative action to recruit, employ, and promote qualified members of groups that had formerly been excluded from fuller participation in the workplace. However, women are still underrepresented at the GS-13 and above grade levels. Seeking to overcome this shortage, OPM targeted for recruitment women in grades GS-9 to GS-12 who showed managerial potential. OPM statistics revealed that the number of female employees available for managerial and supervisory positions decreased significantly at those grade levels. Thus, special incentives were required for women to develop their managerial capabilities.

The WEL program was intended to give participants a series of unique broad-based experiences (normally not provided before selection to middle management positions) that would increase their visibility and further enhance and develop their career potential for supervisory, managerial, and executive opportunities.

The WEL program is modeled after the Senior Executive Service Candidate Development Program. OPM's Office of Training and Development, Executive Programs Division, designed and managed the program; selected and assessed participants to develop individualized programs; coordinated, monitored, and approved individual development plans (IDPs); and assessed and provided program feedback. It cost each participating agency $1,300—less than one-third the actual cost for individual participants—to take part. The program's three phases—selection, orientation, and implementation—are described below.

Selection of Participants

OPM selected participants from nominations submitted by participating agencies. In all, more than 220 candidates representing 54 agencies applied.

At GAO, 17 women were nominated by their divisions offices for the program.

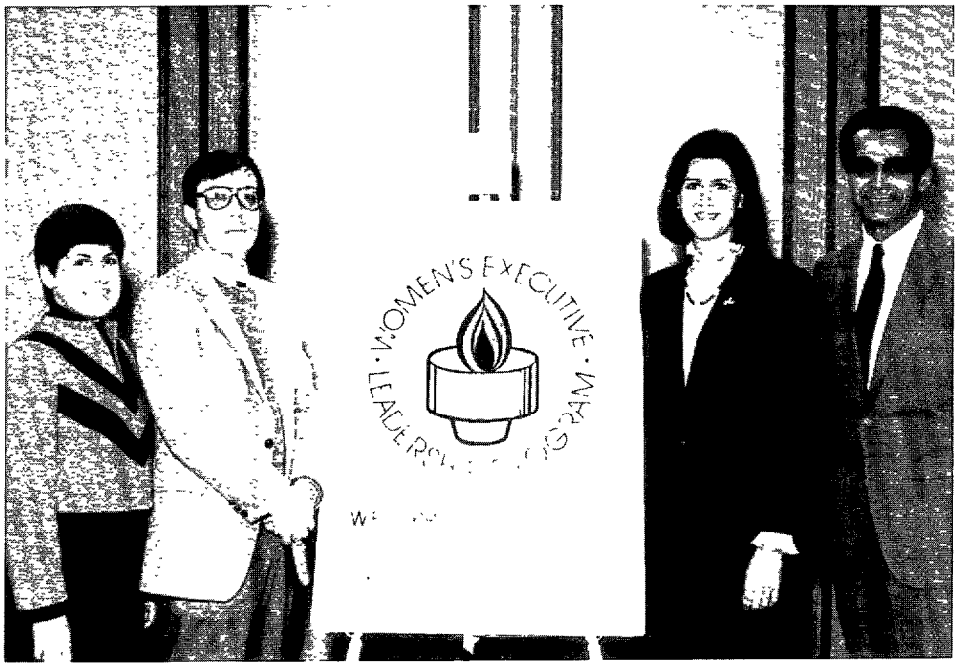
A panel representing the Assistant Comptroller General for Operations, Personnel, the Office of Organization and Human Development, and the Civil Rights Office evaluated nominees and submitted two candidates for OPM's consideration.

OPM selected 64 women, representing 52 agencies, on the basis of the candidates' job experience, performance, and potential for supervisory or managerial assignments.

Orientation

WEL's orientation phase began in August 1984. During the intense 4-month session, representatives from the government, academia, and private industry lectured participants on such topics as

- the constitutional and political basis of government;
- the interactions of the government's executive, legislative, and judicial branches in developing, enacting, and implementing policies, programs, and legislation;
- the manager's role in improving government operations; and



WEL program executives (L to R) Ann Brassier, Assistant Director for Training and Development, OPM; Terry Culler, Associate Director, Workforce Effectiveness and Development Group, OPM; Donna Alvarado, Deputy Assistant Secretary of Defense for Equal Opportunity and Safety Policy; and Gerald Hinch, Deputy Associate Director, Workforce Effectiveness and Development Group, OPM.

- the issues, concerns, and problems that women face as managers in the federal workplace.

These seminars were complemented with the administration of two assessment instruments, the Myers-Briggs Type Indicator and the Management Excellence Inventory. The Myers-Briggs Type Indicator was used to help participants identify individual personality characteristics and preferences that influence managerial style and determine how the participants relate to other people. The Management Excellence Inventory gave participants two perspectives (the participants' and their supervisors') on their skill levels, proficiency levels, and developmental needs. These instruments also identified the collective needs of the participants, thereby supplying OPM data with which to shape the training curriculum.

OPM began the program by briefing participants, their supervisors, and agency coordinators on program objectives, as well as the roles and responsibilities of agency coordinators and supervisors. OPM stressed that the program would not work without the coordinated efforts of OPM and the agencies. Agency coordinators helped prepare individual development plans, determined supplemental training, monitored the progress of the participants, and attended regular meetings at OPM.

The orientation phase ended with the public managers' workshop held in Lancaster, Pennsylvania, in December 1984. At this session, participants reviewed the elements of success outlined in OPM's validated management model, the national policy support role, and the role and responsibilities of the mid-level manager. At this juncture, we also began laying groundwork for our IDPs—identifying rotational assignments, formal training, and other activities on the basis of proficiencies and developmental needs.

The use of assessment instruments was helpful in preparing my IDP, but the advice provided by management of the National Security and International Affairs Division (NSIAD) was more beneficial. My former supervisor, Zeke Baras, along with Associate Director Henry W. Connor and Division Director Frank C. Conahan, provided suggestions for assignments and training that helped to build a solid IDP.

Implementation

The 11-month implementation period, which began in January 1985, was allotted to implementing those activities outlined in the IDP. In addition, OPM held monthly 3- to 4-day training sessions that were mandatory and included personnel management, executive communication skills, supervision and leadership, and interpersonal and organizational sensitivity. Monthly training was complemented by small group activities. These activities provided a means to practice those skills learned in the classroom. Activities included public speaking, personal financial planning, and simulated management activities.

Participants received on-the-job development training through rotational assign-

WEL pilot program participants at a spring workshop held at the Federal Aviation Administration.

ments identified in their IDPs. OPM recommended that each participant perform two rotational assignments that lasted at least 5 days. OPM also emphasized that these assignments should provide experience at a higher grade level or an opportunity to observe top-level managers in action.

I performed three rotational assignments during the period. For example, I worked 4 months with the House Armed Services Committee staff during the budget authorization cycle. I spent another 2 weeks learning and observing GAO regional office policy and operations and job planning and management while under the guidance of Atlanta Regional Office staff members Elkins Cox and Dave Gray. I also joined Atlanta's Suboffice staff in Huntsville, Alabama, for 8 weeks as a field evaluator on the survey phase of a NSIAD job. My last rotational assignment (also my present job) has been with my division's planning staff. Under the direction of George Egan, I have been involved in annual work planning, job management, scheduling and staffing, staff-year monitoring, and operation analyses.

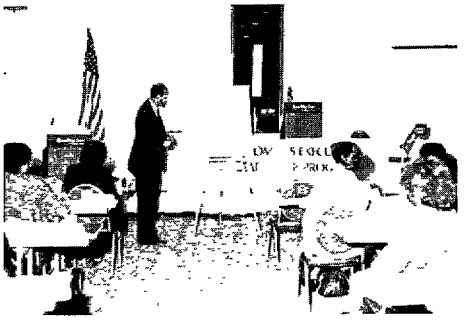
Although I was permitted to participate in the program full-time, many individuals participated while maintaining their normal work loads. For them, this circumstance limited rotational assignments to only a few days. Other participants worked on such rotational assignments as the U.S. Information Agency detail to the British Broadcasting Company and Radio Free Europe, the Department of Agriculture detail to the National Governors Association, and the Department of the Army detail to the 1986 Combined Federal Campaign Loaned Executive Program.

Program Assessment

The program was hectic and, at times, physically demanding. Fortunately, participants received time and stress management training early in the program and developed good networks and support systems with one another. In fact, program graduates are still communicating through the Executive Leadership I Chapter of Federally Employed Women. This chapter, chartered in February 1986, has 98 percent of the graduates enrolled. OPM Director Constance Horner is also a member.

OPM considers the pilot WEL program extremely successful and "attributes this success to strong support at the agency level and successful integration of some of OPM's executive development tools which identified developmental needs and enabled participants to accomplish career objectives in an exemplary manner." Agencies invested in each participant an average of $3,000 in additional training outside the program. So far, the Department of Health and Human Services and several defense agencies have started their own women's executive development programs.

Author leading workshop entitled "Differences in Socialization of Men and Women and How This Influences Management Styles" at the Public Managers Workshop, Lancaster, Pennsylvania.



A consultant making a presentation on leadership competencies at a WEL training session.

Thirty-three percent of the GS-12's who participated in WEL were promoted to GS-13 during the 15-month period. Overall, 32 percent of all participants were promoted.

OPM met its objective of increasing the pool of trained, talented women who were prepared to pursue careers at the executive level of government. Despite this success, OPM did not start a second program. Instead, it is conducting a new program, the Executive Potential Program for Mid-Level Employees. Through training and developmental experiences, this new program aims to prepare participants (GS-11's to GS-13's with high potential) for future opportunities as supervisors, managers, and executives. However, due to the cost of $4,900 per person and impending effects of the Balanced Budget and Emergency Deficit Control Act (P.L. 99–177), only 13 people (5 men and 8 women) from 6 agencies are participating. In light of budget restrictions and the controversy surrounding affirmative action programs, the future seems questionable, at best, for such special efforts as the WEL program.

I am delighted to have been a part of this unique experience. The diversity of my rotational assignments enhanced my technical skills and expanded my awareness of those attitudes and perspectives that an effective manager needs. Under normal circumstances, it might have taken me several years to accumulate the combined knowledge I gained in only 15 months through the WEL program. I am both proud of and grateful to GAO for its active interest and investment in my career development. ■

No Consensus on the Census

Jacob (Jack) Kaufman

Mr. Kaufman has been the Assignment Manager at GAO's audit site at the Census Bureau for the past several years. He has been employed by GAO since 1960. He originally worked for the New York Regional Office until 1973, when he transferred to the General Government Division. Mr. Kaufman received an Outstanding Achievement Award in 1985 and Certificates of Merit in 1980 and 1982. He was graduated with honors from Rutgers University with a B.S. in accounting and is a certified public accountant (New York).

The 1990 Decennial Census of Population and Housing will be the 21st in a chain of U.S. censuses that have been taken, as required by the Constitution, every decade since 1790. The census is used to (1) apportion seats in the House of Representatives, (2) determine congressional districts and potential redistricting of state legislatures, (3) distribute billions of dollars in federal and state funds annually, and (4) help government and businesses manage their operations.

GAO's Role in the Census

GAO has been requested by the Subcommittee on Energy, Nuclear Proliferation, and Government Processes, Senate Committee on Governmental Affairs, and the Subcommittee on Census and Population, House Committee on Post Office and Civil Service, to review planning for the 1990 census. They are most interested in the cost of the census and the quality of the results. In an era when cost cutting is stressed and the competition for federal and state aid is keen, the role of an independent evaluator is more important than ever.

The Bureau of the Census has begun a series of test censuses and special-purpose tests to evaluate possible procedures for use in the 1990 census. The quality of the potential census data, as evidenced by the test results and cost and benefits of procedures, will be closely evaluated by the auditors. GAO will monitor the ongoing tests so that their results will be useful to the above Committees and the Bureau for planning the 1990 census. This on-line audit activity differs somewhat from a more usual postaudit role. This can be difficult, as the auditors will not have the luxury of time. However, if GAO is going to make a significant contribution, it cannot wait to perform a postaudit. Census planning is a moving target adjusting to the results of tests. In addition, GAO has already testified several times concerning the 1990 census. The forecast is for an even busier hearing schedule.

Basic Questions

The primary purpose of the census is to count the population. This sounds simple and straightforward. But like many other things in life, the census is not as simple as it seems. Questions have been raised continually about the census, including: Who should be counted? Where should they be counted? How should they be counted? Can or should raw counts be adjusted using statistical methods? How should seats for the House of Representatives be apportioned? What should be asked? How much is the nation willing to spend on the census? The balance of this article explores these questions.

Who Should Be Counted?

Let's go back to the U.S. Constitution. Article 1, section 2, provided, at its initial ratification (1788), that

" . . . Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective Numbers, which shall be determined by adding to the whole Number of free Persons, including those bound to Service for a Term of Years, and excluding Indians not taxed, three-fifths of all other Persons. The actual Enumeration shall be made within three years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years in such Manner as they shall by Law direct. The Number of Representatives shall not exceed one for every thirty Thousand, but each State shall have at Least one Representative . . . . "

More recently, a debate has continued over whether illegal aliens should be counted and used as a basis for representation in the House. In the 96th Congress, a constitutional amendment (H.J.R. 581) was proposed to limit the count to citizens. In that session, a bill was introduced to accomplish the same objective by placing restrictions on the use of a fiscal year 1981 appropriation. In the 97th Congress, another constitutional amendment was proposed (H.J.R. 233) to exclude illegal aliens from future population counts for apportioning seats in the House of Representatives. None of these initiatives was successful.

Should Illegal Aliens Be Counted?

In the current Congress (99th Cong., 1st sess.), S. 1734 has been introduced, according to its sponsor, to prevent distortions in reapportioning seats in the House of Representatives caused by using census figures which include illegal aliens. The sponsor explained how including illegal aliens in the apportionment after the 1980 census had affected the allocation of House seats.

In connection with the 1980 census, the issue was reviewed in the federal courts *(Federation for American Immigration Reform v. Klutznik).* While dismissing the suit on legal grounds, the three-judge district court noted:

"It [the Constitution] required the counting of the whole number of persons for apportionment purposes, and while illegal aliens were not a component of the population at the time the Constitution was adopted, they are clearly 'persons' " (D.C.D.C. (1980) 486 F. Supp. 564).

The appellate court affirmed the opinion, and the Supreme Court denied the plaintiff's request to review the decision.

Aside from the legal issues, the practical question of including illegal aliens would present major problems. People who are undocumented aliens would probably not wish to identify themselves in the census, regardless of the spotless Census Bureau record of maintaining confidentiality over data collected.

In the 1980 census, the Bureau did include questions on its long-form questionnaire that could be used to approximate the number of illegal aliens included in the count. It asked: "If this person was born in a foreign country, — is this person a naturalized citizen of the United States?"

The respondents had to select from these possible answers: "a. Yes, a naturalized citizen," "b. No, not a citizen," or "c. Born abroad of American parents." The number of persons who designated that they were not citizens could be compared with Immigration and Naturalization Service (INS) records or alien registration to determine how many were in the United States under allowable circumstances.

Remember, this method would provide only an approximation because of the limitations in the methodology. The long form was provided only to a sample of the population—1 in 6 households, except in communities where the population was estimated at less than 2,500. In these locations, the long form was sent to one of every two households. As a result, there is a sampling error. But this limitation is not as critical to the process as the reliability of the responses and the accuracy of INS records. This methodology cannot be used for the next census because the program by which INS alien registration data was obtained was canceled after 1981.

What Groups Should Be Excluded?

Traditionally, the Bureau has excluded certain groups, for example, people living on the grounds of an embassy, a legation, a chancery, or a consulate, who are considered to be living on foreign soil and, therefore, not in the United States. Under current procedures, however, U.S. diplomatic personnel living outside the embassy are counted. Also, citizens of foreign countries temporarily visiting or traveling in this country are not enumerated because they have not established residences.

Americans who are overseas for an extended period, for example, people serving in the armed forces, working at civilian jobs, or studying at foreign universities, are excluded from the counts for apportionment purposes on the basis that their usual residences are outside the United States. On the other hand, Americans who are temporarily abroad on vacations and business trips are counted at their usual residences in the United States. Thus, an interesting condition exists. Members of the armed forces who are overseas and are entitled to vote are not counted, but illegal aliens who are not entitled to vote are counted.

Where Should People Be Counted?

In accordance with practice dating back to the first census in 1790, each person is counted as an inhabitant of his or her "usual place of residence," which has generally been construed to mean the place where the person lives and sleeps most of the time. This place is not necessarily the same as the person's legal residence or voting residence. People without usual places of residence are counted where they happen to be staying.

The Congress has delegated the authority to establish special residency rules to the Secretary of Commerce and has permitted the Secretary to delegate further to the Census Bureau. The rules established have been reviewed and upheld by the courts, most recently in the *Borough of Bethel Park, Pennsylvania, v. Stans* in 1971.

Members of the armed forces are counted as residents of the areas where the installations to which they are assigned are located. College students are counted where they are living while attending college. On the other hand, boarding school students below the college level are counted as residents of their parental homes on the assumption that they are not yet living independently and would return regularly to those homes.

The residence rules established by the Bureau can obviously affect state districting as well as House apportionment. The rules must consider legal, political equity, and practical considerations. Some states and locations can gain an advantage or be placed at a disadvantage because of the rules. For example, a state which has large military bases or ports, such as Virginia, could gain from the rules. Some states which have more college students than college berths from them could be at a disadvantage for apportionment purposes.

Important factors in developing the rules are the Bureau's ability to obtain accurate information at a reasonable cost considering the size of the population groups being counted. For example, there is no good source of data for the number of U.S. citizens overseas who are not affiliated with the federal government. Locating them would be difficult; thus, they are not counted.

The Count: How Accurate and Complete?

Counting the population is not like counting merchandise in a store. A store or a factory can be closed for inventory taking. This country cannot close so that a count can be taken. Moreover, the components of an inventory are not usually mobile. In a merchandise inventory, the total is the important result. In a census, the most important factor is not the total but the location of its components. People are missed in a census for several reasons: Housing units

are not identified and thus the people residing in them are missed; people may not permanently reside in fixed housing units; and people may intentionally wish not to be counted. In many situations, dwelling units may be difficult to find: Units are located in alleys, garages, basements, or other apartments and subdivided units. There is a whole class of unusual dwelling places, such as campers, cars, boats, and tents.

Also, many people are not permanent residents of fixed dwelling units. They may be drifters who sleep in such places as railway or bus stations, all-night movies, or streets. In addition, some people have temporary lodgings in shelters, boarding houses, motels, or institutions.

Most difficult to count are those deliberately omitted by the household respondent because they are undocumented aliens, are fugitives from justice, are behind in child support or alimony payments, and are violators of building occupancy requirements who fear identification. Although the Bureau has a spotless record of protecting confidentiality, to reassure these people may be difficult. The law does not require people to step forward and be counted. The only obligation is to respond truthfully when the Bureau finds them.

Do Census Methods Produce Quality Results?

Census-taking methods have changed over the years, but disputes about the accuracy of census results have persisted. The first census was conducted by U.S. marshals in 1790, who reported their results to the Secretary of State and filed a copy with the nearest U.S. district court for the scrutiny of the public. Confidentiality was not an issue then.

Even the first census results were questioned, as evidenced by this remark by George Washington:

"Returns of the Census have already been made . . . by which it appears that we shall hardly reach four millions; but one thing is certain, our real numbers will exceed, greatly, the official returns of them; because the religious scruples of some would not allow them to give in their lists; the fears of others that it was intended as a foundation of a tax induced them to conceal or diminish theirs, and thro' the indolence of the people, and the negligence of many of the Officers, numbers are omitted."

By 1840, it was evident that the methods used to gather census data did not produce quality results, and the American Statistical Association issued a report criticizing the accuracy of the results. Later, the office responsible for the census for a period of time established a Bureau of Revision to check the census reports and reconcile them before publication. Census results, including those of the 1980 census, have been questioned in the 19th and 20th centuries. In at least two cases in the 19th century, states were allowed additional repesentation when they claimed that the census results were invalid. After the 1920 census, because of political disputes and poor weather during the census, the Congress could not decide on an appropriate scheme for reapportionment.

For the 1970 census, a major change in the census taking was initiated. The Census Bureau changed from its traditional method of sending enumerators to canvass areas to the use of a mail-out/mail-back method. In the 1970 census, about 60 percent of the households received their questionnaires in the mail and were requested to mail back the completed forms. Enumerators contacted households only when the questionnaires were not returned or were not properly filled out. The new procedure was extended to about 90 percent of the households in the 1980 census.

Data Quality

A major reason for this change in procedures was to improve data quality. Studies conducted by the Bureau in 1950 and 1960 showed that significant problems existed with the completeness and accuracy of the census. Although the estimated overall percentages of people missed in the 1950 and 1960 censuses were not a significant problem, there were major differences in the miss rate among different racial groups and geographic areas. Thus, the Bureau believed that some areas had been much better counted than others. The studies also showed that enumerators had caused some errors. Consequently, the Bureau decided to rely on the households to complete their own questionnaires.

Despite the change, many complaints about undercounts were received about the 1970 census. By November 1970, the Bureau had received almost 1,900 complaints from various communities. As a result, the Bureau instituted several procedures. For example, in about 500 communities, corporate boundaries had to be verified. Also, field checks were made in 187 places where there were indications of improper counts. As a result of these activities, the counts for about 600 communities were revised before final figures were published. On the average, additions to the census constituted 5 people for every 10,000 originally enumerated in these communities.

Legal Suits

Several lawsuits were filed because of the 1970 census results. Generally, they were dropped by the plaintiffs or dismissed or the Bureau's procedures were upheld by the courts. The Borough of Bethel Park, Pennsylvania, unsuccessfully challenged the Bureau's practice of allocating college students and members of the armed forces away from home to the place where counted.

The increased number of grievances about the 1970 census may be attributed to the "one man one vote" concept stemming from court decisions. Also, starting with the 1970 census and to a greater extent for the 1980 census, more attention was placed on the census results because of their increasing importance as the basis for annual federal and state fund distributions. Although there is no precise calculation of the amount involved, the combined fund distribution for fiscal year 1980 could have been about $100 billion. Federal grants-in-aid to state and local governments alone totaled $82.9 billion.

With so much at stake, it is not difficult to understand why about 50 political jurisdictions or groups challenged the results of the 1980 census. The plaintiffs believed that the population counts should have been adjusted because (1) the census cannot count all people no matter how conscientious the effort, (2) the quality of the census was suspect, and (3) the census included illegal aliens. The communtities contended that because of an undercount or an improper count, they had been shortchanged in political representation and in receiving federal and state funds. Some of these cases have yet to be decided, but the ones that have been were all decided in favor of the federal government. Many lawsuits requested that the courts adjust the population counts to compensate for the undercounts. However, the legal and statistical issues about whether to adjust raw census counts are not easy to resolve. After considering ways to adjust for the undercount in the 1980 census, the Bureau decided against it. The Bureau believed that there was no statistically defensible way of accurately estimating the undercount at the subnational level and perhaps not even at the national level.

How to Apportion

The issue of how to apportion seats for the House of Representatives has been a

continuing source of debate among political scientists, elected officials, and mathematicians for almost as long as the history of the United States. George Washington's first veto concerned apportionment. For complete fairness, it would be desirable to assign each state the exact number of representatives proportional to its population. Unfortunately, this rarely works out to be a whole number. The accepted method of rounding to the nearest whole number cannot be used because the size of the House could vary from the required 435 and a small state's allotment could round to zero, which would violate the constitutional requirement that each state have at least 1 representative.

Historically, six methods have been used for apportionment. The Jefferson method was used five times between 1790 and 1830. The Webster method was used only once, in 1840. In 1850, Samuel Vinton, a Member of Congress, sponsored legislation based on a method first proposed by Alexander Hamilton. That method was used six times between 1850 and 1900. The modified Webster method, known as major fractions, was used in 1910 and 1930 (no apportionment was made after the 1920 census). The current method, called equal proportions, has been used since the 1940 census.

Some methods may favor large or small states. For example, the Jefferson method favored large states. Some contend that the equal proportions method shortchanges the larger states. For this reason, a Congressman from Indiana introduced H.R. 1990 in the 97th Congress, referred to as the Census Data Reform Act of 1981. One key provision of that bill was designed to eliminate the "imperfections" of the current equal proportions formula for allocating congressional seats by reinstituting the Hamilton-Vinton method. The bill did not pass, but the question of how the apportionment should be performed was reintroduced in the Congress. The issue of apportionment is a political one because no one formula will satisfy all conditions to provide fairness.

Questionnaire Content—How Much Is Needed?

The questionnaire content has varied during the 20 prior decennial censuses. The first census asked for the following information: number of persons, number of males or females, number of free people or slaves, number of males 16 years old or over, and names of heads of households. Between 1800 and 1830, questions were added relating to physical disability and people of foreign birth. By 1890, the census included 238 questions. Many of these were asked to obtain information about economic activity of the nation. Questions were included on such topics as mining and agriculture. Subsequently, this information was asked on special censuses. By 1930, the number of questions had dropped to 42. In 1970, the basic questionnaire consisted of 27 questions while 2 longer versions contained 81 and 94 questions and were directed to 15 percent and 5 percent of the population, respectively.

For the 1980 census, the Census Bureau distributed two types of questionnaires, a short form and a long form. About 81 percent of the nation's housing units received a short form that contained 19 population and housing questions. The other 19 percent received a long form that contained all the questions on the short form, as well as 20 more questions about the housing unit and 26 additional questions for each household member.

The Census Bureau is now determining the content and the design of the 1990 questionnaire. A 1986 form used in a pilot test for determining the content and the design has incorporated more questions than in 1980. The new questions being tested ask for more details on income, education, health, and even the number of smoke detectors in homes.

Although there are uses of these data, critics are concerned that including more and more detailed questions may adversely affect the quality of the population count and the cost of conducting the census. The census data, it is pointed out, help determine the need for a wide variety of social, educational, and economic programs. On the other hand, the primary purpose of the census is to ensure that seats in the House of Representatives are divided fairly among the states.

The expanded use of the mail-out/mail-back technique increases emphasis on the respondents' ability to read and understand the questionnaire. If respondents are turned off by the size and complexity of the form, they will not send back their questionnaires. As a result, the Census Bureau will have to send its enumerators to gather the data. This is not only costly but also defeats the purpose for having a mail-out/mail-back census.

How Much Should Be Spent on a Census?

For the 1980 census, the Bureau spent about $1.1 billion, much more than ever before. Even when inflation and increased work load are discounted, the 1980 census still cost about double the amount of the prior census. The Bureau's evaluation studies indicated that overall, the 1980 census had the lowest undercount. Nonetheless, there were still big differences between the estimated undercount rates between racial groups.

For the 1990 census, the Bureau has planned a two-pronged approach to obtain a complete and accurate count. First, a concerted effort will be made to obtain quality counts. This will incorporate procedures directed to ensure complete and accurate counts. Some planned procedures will be costly, including double-checking on the completeness of the address lists. Second, the Bureau has initiated procedures, which are now being tested, of comparing census results with other data collection results to detect missing people. This comparison has been tried in prior censuses. The possible innovation for the 1990 census is an automated matching of records between the surveys and the census. The automated procedures could expedite the matching procedure.

The improved census procedures and evaluation techniques will add to the census cost. This runs counter to one of the main objectives of the 1990 census, which is not to spend more per household, exclusive of inflation, on the 1990 census than was spent on the 1980 census. Cost estimates for the 1990 census have ranged from about $2 billion to $4 billion.

Conclusion

The importance of the census to governments at all levels and ultimately to the citizens has resulted in increased pressure to ensure an accurate count to provide for an equitable apportionment of seats in the House of Representatives and for other quality data. Various opinions have been expressed on how this should be done. Because of the importance of the census and the varied parties who have an interest in the results, debate on the census will likely continue.

If the trend continues, the next census, the bicentennial census, will be more closely watched than any census before. There will be the traditional concerns about data quality and equity. In addition, in that census, because of the budge-cutting environment, costs will become a major issue. On the basis of the experience to date, GAO

See Census, p. 33

Public Sector Truths, Private Sector Myths: A Contrast of Management Styles

Edward W. Gray, Jr.

Edward W. Gray, Jr., spent 1984-85 with GAO on an assignment through the President's Commission on Executive Exchange. He served as Special Assistant to the Director, General Government Division. Mr. Gray is Manager, Federal Systems Region, of R. R. Donnelley & Sons Company, a commercial printing firm in Falls Church, Virginia. In this position, he directs sales of Donnelley services to federal agencies, such as the Internal Revenue Service (printing of annual tax packages), the Federal Reserve System ***Federal Reserve Handbook****, and the Government Printing Office (production of federal directories, agency newsletters, and journals). Formerly, as general attorney for R. R. Donnelley, he proposed patent, confidentiality, and trade secret policy for the firm and negotiated printing agreements with Her Majesty's Stationery Office for producing classified and alphabetical telephone directories for most of Great Britain. He joined the company in 1973 as an attorney and subsequently served as a senior attorney from 1976 to 1981. Prior to joining R. R. Donnelley, Mr. Gray was an associate attorney with the Chicago law firm of Kirkland & Ellis. He has a B.A. in sociology from the University of Chicago and a J.D. from the Columbia University Law School. He belongs to the American, Chicago, Cook County, California State, District of Columbia, and National Bar Associations and was a vice president and director of the Chicago Council of Lawyers.*

Ed. note: The following article reflects the author's perceptions of and experiences in the Executive Exchange. He wishes to acknowledge the role of Comptroller General Charles A. Bowsher, Jim Howard, Terry Draver, Arnold Jones, Jerry Miller, Barbara Moran, GAO; Mike Nemeroff, Sidley & Austin; George Beckerman, Beckerman Associates; Cheryl B. Leggon, National Research Council; and William G. Norton, President's Commission on Executive Exchange, and to give special thanks to Sallee Garner, GAO, who provided the extensive reference sources.

Would you find appealing an opportunity to totally change your line of work and place of residence for 1 year? If so, you might enjoy the chance to participate in a public sector/private sector exchange experience designed to remove you from the familiar—the President's Commission on Executive Exchange. The Exchange is designed to give public and private sector participants a 1-year experience in the opposite sector to broaden perspectives on organizational and management issues that affect the entire society.

From September 1984 to August 1985, I was privileged to participate in Exchange XV of the President's Commission and to serve at GAO as Special Assistant to the Director of the General Government Division, William J. Anderson. Beyond relocating from a legal post at a Chicago printing company to Washington, the opportunity for close contact with various leading elected and appointed officials provided a richness of experience that my family and I cannot imagine being equaled. Such a complete change in environment stimulated me to research similarities in and differences between public and private sector management. My most significant conclusions are as follows:

- There is a lack of rigorous and systematic study of differences and similarities regarding what constitutes good management in both sectors. The largely unrecognized benefits from such study would flow from the public sector to the private and vice versa.
- Contrary to popular opinion, government managers *can* teach private sector managers useful things.

The relative lack of comparative literature on what constitutes good management in the public and private sectors (Lynn, 1984, and Ring and Perry, 1985) may result from an ideological stance that is rooted in American hostility to big government. Accordingly, what comparative literature there is seldom goes beyond simple similarities in and differences between business and government to discuss differences in effectiveness. (Baldwin, 1985)

Only recently, after painful losses to foreign public and private cooperative efforts, have Washington and business begun to question the American separation between public and private policy. For example, subsidies for agricultural products by the European economic community enabled European farmers to profit from sales to importers of agricultural products at the expense of the American farmer. However, even today, informed people tend to be emotional and ideological rather than systematic and objective in assessing where public, private, or combined solutions offer the most promise in addressing important public and individual needs. As a result, little literature is available on what management solutions can be transplanted to the other sector, which cannot, and why.

According to Lynn, "The development of theoretical and practical perspectives on the distinctive competencies needed for public sector management thus remains an

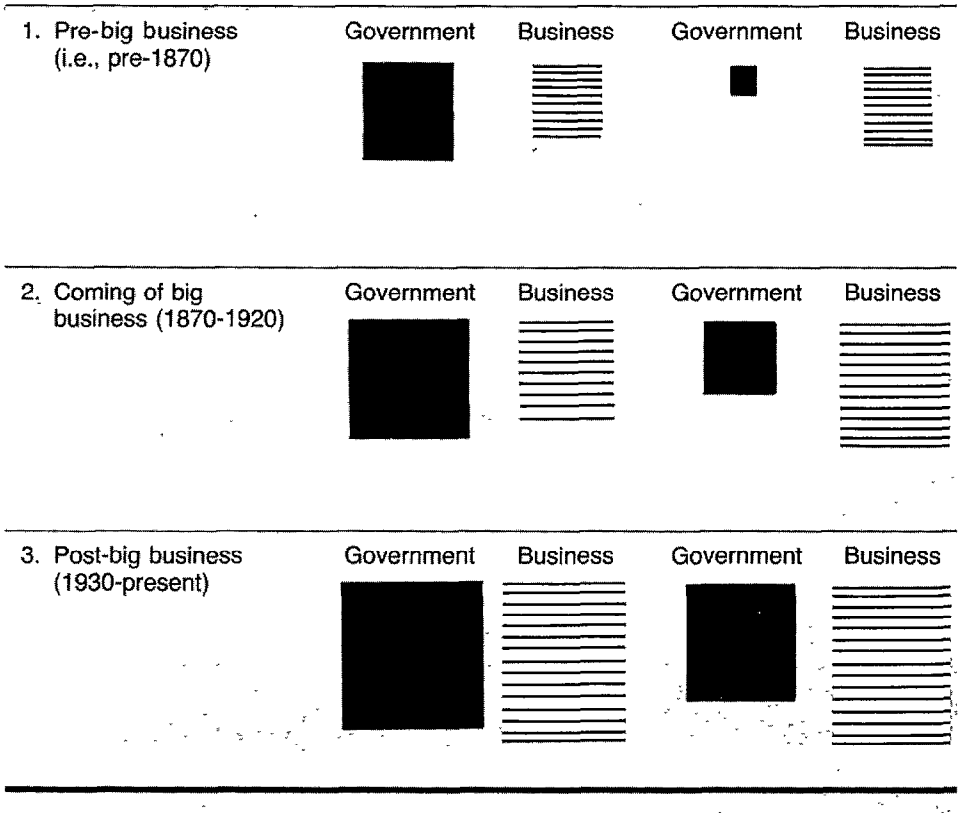
open, inviting yet sparsely settled intellectual and pedagogical frontier. Preoccupation with generic concepts relating to the general management of 'organizations' has precluded significant research into differences among organizations and the implication of those differences for management and management education. The few exceptional efforts along these lines—research concerning the distinctive character of not-for-profit organizations, public authorities, and government-owned corporations—prove the rule: there is no field of comparative organizations. It is an intellectual vacuum that should be filled. The emergence of a field of comparative research on the general management of organizations would deepen and sharpen understanding of general management of all kinds: public, private, and mixed." (Lynn, 1984)

To learn more about general management, I searched for contrasting themes, and I discovered several in the literature that I reviewed. While not meant to be all-encompassing, they provide a starting point for dialogue between public and private sector managers.

The American Conflict Between Public and Private

The history of the conflict between the public and private sectors in America begins with the enormous private growth offered by the frontier and 19th century entrepreneurship that anticipated the expansion of American government. The uniqueness of this history (McCraw, 1984) is shown by a comparison of the growth of the public and private sectors in Europe, Japan, and America. In America, the establishment of big business *preceded* the development of big government. Indeed, in the early years, government growth in the United States was often a reaction to the excesses of explosive private sector growth.[1] This circumstance provided a natural adversarial tone in 19th century America to the dialogue between big business and upstart big government; however, as shown in figure 1, big government in Europe and Japan preceded the growth of big business organizations by a considerable extent. Thus, unlike America, government in Europe and Japan could take credit for regulating growth in the private sector. One continues to this day to see evidence of more involvement by government with business decisions in Europe and Japan.[2] For example, trade issues underlie the creation of and are the most important activity regulated by the European Common Market structure. In Japan, government

Figure 1. Growth Rates of Government and Big Business



Source: McCraw, 1984.

more directly affects research and development and strategic planning by industry than in America.

While it may certainly be argued that the origins of the United States had a lot to do with economic considerations of the colonists, these considerations often paralleled concerns about too much government involvement in business that were not common elsewhere. McCraw says that few European or Japanese businessmen took it for granted that they could make important investment decisions without consulting the state but that American businessmen were outraged when the U.S. government first did claim such a role during the New Deal. (McCraw, 1984)

Attitudes Toward Public Employees

Perhaps nowhere is the conflict between the public and private sectors in America more painfully and senselessly played out than in public attitudes toward civil servants. Quite unlike the English example, it is currently fashionable in America to disdain and berate the merit-selected career public servant. This is witnessed by a real scarcity of professional managers with a strong background in both public and private sector management. As a consequence, regardless of which political group is in the White House, the executive branch is constantly being led by political appointees with little preparation for government while professional government managers are unsung, underused, and unappreciated, thus encouraging poor performance, which feeds the popular view of government. (Ingraham, Ban, 1986) Almost as a corollary, in the legislative branch, the professional nonelected Hill staffers have enormous impact on the legislative process and product but are overworked, underpaid, and virtually invisible to the public. (Heclo, 1977)

[1]Witness the Sherman Anti-Trust Act (Act of July 2, 1890, c. 647, 26 Stat. 209), the Pure Food and Drug Act (Act of July 1, 1902, c 1357, 32 Stat. 632), and the Interstate Commerce Commission Act (Act of Feb 4, 1887, c 104, 24 Stat. 379).

In Canada, on the other hand, the relationship seems more closely analogous to the American situation with a considerable anti-big-government tone to the public sector/private sector discussion (Auditor General of Canada, 1983).

Table 1. Strengths and Weaknesses of Public and Private Sector Management

Private Sector Management	
Strengths	**Weaknesses**
Freedom from statutory and constitutional constraints on rapid action. More flexible internal rules and procedures. More uncontrollable public scrutiny.	Lack of accountability to constituencies other than important shareholders: rough justice.
Bottom-line (profit) orientation, which facilitates goal setting.	Short-term nonstrategic perspective.
Recognition that pleasing the customer is paramount.	Vulnerability to economic dislocations.
Results orientation—getting the product out the door or the service performed.	
Greater ability to provide incentives and disincentives to workers and managers.	
Integration of revenue, treasury, and expenditure functions, which assists cost control and revenue enhancement efforts.	
Public Sector Management	
Ability to reconcile conflicting constituencies, i.e., govern.	Lack of consensus on policy issues or organizational goals and procedures.
Experience in managing political issues.	Conflict and lack of understanding between political appointees and merit-selected bureaucracy.
Ability to use media effectively.	Overemphasis on approval process as compared with implementation and results.
Size and resources sufficient to affect pace of change and to permit long-range planning.	Lack of efficient priority-setting mechanisms; arbitrary deadlines.
Complex decision-making prevents or reduces hasty, ill-considered actions.	Excessive statutory and legislative oversight constraints on management initiatives, which inhibit adjustment of programs to legitimately changed or extraordinary circumstances and cause poor accountability for results.
	Inadequate incentives for good management or managers.

Management Differences in Government and Business

There is wide agreement that the fundamental differences between management in the public and private sectors arise from the ever-present potential for political consequences from government policies and actions. Nevertheless, because political behavior covers such a broad range of conduct, politics alone provides an overly general explanation of the perceived and actual differences between the public and private sectors. Nor is the term "private sector" particularly helpful in comparing management issues in organizations. Is it really accurate to view all private sector organizations as constantly competitive, efficient, profit oriented, and unaffected by politics? Obviously, the answer is no. For example, public utility monopolies, such as electric and gas companies, large chemical companies, and major defense contractors, are private sector organizations clearly and frequently affected by public issues. Such organizations also often illustrate examples of bureaucracy and/or management failure as compelling as any such occurrences in the public sector. When, then, can valid distinctions between public and private sectors be drawn? What can each sector expect to learn from the other? Where are they both efficient and deficient?

It is as popular within the Washington metropolitan area to challenge the mythology of private sector superiority as it is to lament the failings of Washington in the rest of the nation. Such opinions are seldom based on a systematic or even an informed understanding of the other sector. Yet, even the most uncritical proponent for either sector must acknowledge shortcomings. As I reviewed the literature on management and environmental differences between the public and private sectors, I abstracted the scheme of strengths and weaknesses shown in table 1.

Do People Want Good Management in Government?

It is frequently said that in preference to tyranny, the framers of America's government self-consciously chose inefficient government. However, one may reasonably question whether they intended *as much* inefficiency as they have achieved. Despite obvious and long-standing problems in government and numerous private sector efforts to detail and attack such difficulties (e.g., Hoover, Gulick, Ash, and Grace Commissions), private sector managers generally have done little to build *political* support for good management in government. Such disinterest in building political support for management improvements in government explains why good management is such a low political priority in Washington. More people in America must understand and accept that good government cannot exist apart from good politics before good management becomes a political priority and thus a governmental reality. (Malek, 1978)

What Can Government Teach Business?

While to even suggest that government can teach business anything may be controversial to some, I have personally observed, and the literature confirms, that the public sector leads the private in the general ability of mid- and senior-level management to competently handle media relations in particular and public scrutiny in general. In addition, leadership skills and personality

traits, correlated with effective government management, might offer useful approaches to private sector managers seeking productivity improvements or experiencing work force conflict, particularly where the balance of power is held by labor. In addition, the examples I saw of the use of handicapped government employees illustrated a blend of opportunity, compassion, and hard-headed but fair performance requirements that I'd never witnessed in the private sector.

Conclusion

While the economic pressures of today may compel better public/private cooperation, we should not abandon this course as the economy improves. Indeed, since there will always be public and private sectors, the more they understand and cooperate with each other in promoting economy and efficiency in society, the less senseless conflict there will be and the greater will be the return on investment to citizens.

Sources on Public and Private Sector Management

Baldwin, J. Norman. *Differences and Similarities Between the Public and Private Sectors: A Bibliography.* Lubbock, Tex.: Texas Tech University, Department of Political Science, 1985.

Balutis, Alan P. "Public Excellence: Not an Oxymoron." *The Bureaucrat* (Winter 1985–86).

Blank, Rebecca M. "An Analysis of Workers' Choice Between Employment in the Public and Private Sectors." *Industrial and Labor Relations Review* (Jan. 1985).

Bowers, Joseph L. "Effective Public Management." *Harvard Business Review* (Mar.-Apr. 1977).

Constraints to Productive Management in the Public Service. Auditor General of Canada. Ottawa: 1983.

Farrell, Kevin. "Public Service in Private Hands." *Venture* (July 1984).

Gold, Kenneth A. "Managing for Success: A Comparison of the Public and Private Sectors." *Public Administration Review* (Nov.-Dec. 1982).

Helfand, Gary. "The Applicability of Japanese Management Techniques in the American Public Sector: Some Cultural Considerations." *Public Productivity Quarterly* (June 1983).

Heclo, H. *A Government of Strangers: Executive Politics in Washington.* Washington, D.C.: Brookings Institution, 1977.

Hodgetts, J. E. "Managing Money and People in the Public and Private Sectors: Are There More Similarities Than Differences?" Speech to the Institute of Public Administration of Canada (Sept. 1982).

Ingraham, Patricia W., and Carolyn R. Ban. "Models of Public Management: Are They Useful to Federal Managers in the 1980s?" *Public Administration Review* (Mar.-Apr. 1986).

Luce, Sally R. "Leadership and Change in the Public Sector." *The Canadian Business Review* (Autumn 1984).

Lynn, Laurence E., Jr. "Point of View: Improving Public Sector Management." *California Management Review* (Winter 1984).

McCraw, Thomas K. "Business and Government: The Origins of the Adversary Relationship." *California Management Review* (Winter 1984).

McManis Associates, Inc. *A Study of Private Sector/State and Local Government Personnel Operations: Prototype Specifications.* Washington, D.C.: 1985.

Malek, Frederic V. *Washington's Hidden Tragedy: The Failure to Make Government Work.* New York: The Free Press, Macmillan Co., 1978.

Miller, Gifford Willis. "The Ethical Coduct and Behavior of Public Executives Compared to That of Corporate Executives." Ph.D. Thesis. Ann Arbor, Mich.: University Microfilms International, 1981.

Nalbandian, John, and J. Terry Edwards. "The Values of Public Administrators: A Comparison With Lawyers, Social Workers, and Business Administrators." *Review of Public Personnel Administration* (Fall 1983).

Newland, Chester A. "Federal Government Management Trends." *The Bureaucrat* (Winter 1983-84).

Nigro, Felix A. "Public Administration: From Theordore Roosevelt to Ronald Reagan." *International Journal of Public Administration* (Mar. 1984).

Organizational Research and Consultation, Inc. *Key Leader Behaviors for the Successful Management of Organizational Changes.* New York: Dec. 1982.

Ouchi, William G. "The M-Form Society: Lessons From Business Management." *Human Resource Management* (Summer 1984).

Peters, Charles. "How Washington Really Works." Rev. ed. Reading, Mass.: Addison Wesley Publishing, 1985.

Posner, Barry Z., and Warren H. Schmidt. "What Kind of People Enter the Public and Private Sectors? An Updated Comparison of Perceptions, Stereotypes, and Values." *Human Resource Management* (Summer 1982).

Ring, Peter Smith, and James L. Perry. "Strategic Management in Public and Private Organizations: Implications of Distinctive Context and Constraints." *Academy of Management Review*, Vol. 10, No. 2 (1985).

Rudman, Warren B. "Putting Government out of Business." *INC.* (Nov. 1983).

Stein, Jay M. "Public/Private Sector Productivity: The U.S. Perspective." *Public Administration* (Winter 1983).

Ed. note: The above references furnished as guides only. Consult Sallee Garner in the Technical Library for further information. ■

Productivity Improvement in the Federal Government: An Emerging Reality?

Michael Dulworth
Dr. Brian L. Usilaner

Mr. Dulworth is an Evaluator in the General Government Division's National Productivity Group. Mr. Dulworth came to GAO in 1984 from the Department of Labor, where he was an industrial relations specialist with the Bureau of Labor-Management Relations and Cooperative Programs. Mr. Dulworth received a B.A. in history from the University of Michigan. Mr. Dulworth has also attended New College, Oxford University, England, and the University of Madrid, Spain. He is currently working on a master's degree in public administration and organizational behavior from the University of Southern California. Mr. Dulworth has received a Special Commendation Award and an Outstanding Achievement Award for his work at GAO.

Dr. Usilaner is an Associate Director of the General Government Division's National Productivity Group. Dr. Usilaner was associated with the Office of Management and Budget (OMB) for 7 years before coming to GAO. He held several positions at OMB, including Director of the Government-Wide Management Improvement Program, Director of the Cost Analysis and Reduction Unit, and Senior Analyst for Staffing Requirements. He was also Executive Secretary to the President's Advisory Council on Management Improvement. Dr. Usilaner received his B.S. and M.S. in industrial engineering from New York University and Columbia University and his D.Sc. in management science from George Washington University. Dr. Usilaner is an Adjunct Associate Professor at George Washington University and the University of Maryland. He is listed in **Who's Who in America** *and* **Who's Who in Government** *and has received several awards recognizing his federal service.*

Attention to the issue of improving productivity began in the private sector around the mid-1970's. The emphasis was directly related to the decline in our nation's productivity, which began in 1964 and culminated in consecutive years of negative productivity growth in 1979 and 1980. After almost a century of high productivity increases, U.S. relative productivity growth has fallen and our industrial base has lost its advantage to our competitors. From 1964 to 1984, our average productivity growth was just 1.2 percent, compared with 3.5 percent before this period. The 1.2-percent average was lower than the rates of all other industrialized nations. From 1977 to 1982, productivity grew 0.6 percent annually in the manufacturing sector: one-third Germany's growth rate, one-fifth France's rate, and one-sixth Italy's and Japan's rates. In 1983 and 1984, our productivity rose to 3.3 and 1.6 percent, respectively, but in 1985, productivity again showed a negative growth rate of 0.2 percent.

For many years after World War II, the United States relied on its productive capabilities to compete effectively against all other industrial countries and dominate world markets. However, with the rapid rise of Japan and the rebirth of some Western European countries as world industrial powers, the United States has had to reassess its basic approaches to producing competitive products.

The private sector's concern with productivity improvement has only recently spilled over into the federal government. There are many reasons for this heightened awareness, including the recent advent of $200 billion federal deficits, the Balanced Budget and Emergency Deficit Control Act of 1985, new policies and priorities under the Reagan administration, and the publicity associated with private sector productivity problems. Apparently, the public sector is not immune to the negative consequences of poor productivity and must set a course of action aimed at improving its productivity.

The federal government increased its productivity an average of 1.5 percent from 1967 to 1984. However, productivity increased only 0.4 percent from 1983 to 1984. The Department of Labor's Bureau of Labor Statistics (BLS) compiles annual productivity data on federal agencies covering 59 agencies and 67 percent of the federal work force. Organizational units are sorted by 28 government functions. These functions are categorized by similarity of tasks performed. For example, since 1967, the functions of financing and accounting loans and grants and information services have increased their productivity 4.5 percent and 0.9 percent, respectively. In 1984, productivity increased in 18 functions and decreased in 10 functions.

This article provides an overview of private sector productivity improvement efforts and relates what has been learned to recent federal efforts to improve productivity and decrease costs. The federal analysis will concentrate on the President's Productivity Improvement Program and its potential for providing mechanisms by which agencies can begin to improve their productivity and, consequently, their service to the public.

Before we go any further, it would be helpful to define productivity and explain its importance to improving the operations and effectiveness of organizations.

Productivity: What Is It?

Simply stated, productivity is the relationship between the output and input of resources—output divided by input. We define "productivity" as the efficiency with

which resources are used to produce and deliver services or products at specified levels of quality and timeliness. This criterion leads to a definition of "productivity" that has three elements: efficiency, timeliness, and quality. An organization can increase its productivity in three ways: (1) decrease inputs while keeping outputs constant, (2) increase outputs while keeping inputs constant, or (3) some combination of the two. Productivity can be measured many ways—on an individual, group, organizational, multiorganizational, or even national basis.

Why Is Productivity Improvement Important?

Productivity improvement is important because it is a way to increase concurrently an organization's efficiency, timeliness, and quality. The productivity improvement process is an attempt to orient organizations toward productivity and have their action use its elements as criteria in decision-making processes. The integration of these three elements into a productivity improvement process focuses an organization on critical areas for development while not allowing one element to improve at the expense of another.

The importance attached to productivity improvement today is enhanced by reports suggesting that many employees in the United States are not working to their potential. Daniel Yankelovich, in a report prepared for the Public Agenda Foundation in 1983, stated that the productive capabilities of the United States could be substantially improved if we tapped what he called workers' discretionary efforts. Yankelovich defines discretionary efforts as the differences between levels of work that have to be done to keep jobs and the maximum levels possible. The study reported that of all American workers

- 23 percent are not working to their potential,
- 44 percent do not put much more effort into their jobs beyond what is required to keep them, and
- 75 percent say that they could be significantly more productive.

Private Sector Approaches to Improving Productivity

In a November 10, 1983, report, *Increased Use of Productivity Management Can Help Control Government Costs* (GAO/AFMD-84-11), GAO found evidence that the private sector was systematically addressing its productivity problems and that many companies had initiated significant productivity improvement efforts. Twenty-five companies were visited, and many stated that for years they had ignored productivity or viewed productivity improvement as a staff, not core management, function. With changing markets, more worldwide competition, and high inflation levels in the 1970's, firms had to improve productivity or risk failure. They had to adopt strategies aimed at reducing operating costs through productivity improvements.

When first trying to develop productivity strategies, companies underscored the importance of avoiding quick fix or "program" approaches. We were told that the single greatest deterrent to sustained productivity improvement was the tendency to approach the effort as a "program." Programs carry negative connotations for most organizational members, implying a temporary add-on to regular activities. A suggested alternative to the program approach was to view productivity improvement as a management process, an ongoing and integral element of organizational functioning. If a productivity improvement process is integrated into an organization's procedures and practices, appraisal systems, goal-setting practices, management information systems, and human resource systems, it influences behavior in the organization and ultimately becomes part of the organization's culture.

On the basis of our analysis of productivity improvement efforts in private companies, we identified the conditions that need to be incorporated into the productivity improvement process, including the following:

- Top level support and commitment.
- Employee involvement.
- A dedicated high level organizational entity as the focal point.
- Written productivity objectives and goals and an organizationwide productivity plan.
- Productivity measures that are meaningful and useful to the organization.
- Use of the productivity plan and measurement system to hold managers accountable.
- Awareness and communication of productivity's importance to the organization.

While none of these elements is particularly innovative, their integration distinguishes systematic productivity improvement efforts from other approaches and makes them a powerful process for improving productivity and reducing costs. Studies on why productivity improvement efforts fail to achieve comprehensive and sustained improvements reveal certain pitfalls that seem to occur regularly. Most are due simply to the absence of the several key productivity elements mentioned above.

Why Is Productivity Improvement Important in the Federal Government?

The federal government now accounts for about 25 percent of the gross national product and has a total deficit estimated at $1.2 trillion. Budget restraints will be a way of life for federal agencies in the foreseeable future, placing severe demands on all managers to do more with less. This environment can, however, serve as a catalyst for emphasizing the potential positive results of productivity improvement.

Many recent studies indicate that substantial savings could be realized by small improvements in federal productivity. The Grace Commission estimated in 1984 that a federal productivity increase of 5 percent would reduce the government's approximately $90 billion annual personnel costs by about $10.5 billion in 3 years. GAO, in 1983, estimated that every percentage point increase in federal productivity would generate about $1 billion in savings. A 1979 Joint Economic Committee staff study estimated that a 10-percent increase in federal productivity could reduce federal costs by $8 billion while maintaining the present levels of service. Other equally important benefits of productivity improvement could include improved services to the public and higher morale in the federal work force.

Federal Productivity Improvement

GAO's 1983 report found a number of federal productivity improvement efforts, but most operated outside management's mainstream responsibilities, received little top management support, were narrow in scope, and were disjointed. The productivity efforts tended to be isolated from the decision-making process of the organizations and were, therefore, ignored or forgotten by changing top level leadership or administrations. Agencies did not integrate productivity into management processes, a critical element of successful private sector productivity efforts.

Productivity improvement received low priority in the federal government for two principal reasons.

First, government managers viewed productivity improvement from a short-term perspective, emphasizing budget reduc-

tions and short-term results, with minimum emphasis on long-term efforts to improve productivity. Accordingly, managers generally did not give productivity improvement sufficiently high priority. It is not surprising, therefore, that the savings generated by these efforts totaled less than 1 percent of payroll costs. This compares with our findings that successful private sector productivity improvement efforts saved 5 to 20 percent of payroll costs.

Second, federal managers lacked encouragement or assistance in addressing the numerous obstacles facing efforts to improve productivity. Neither the Office of Personnel Management (OPM) nor the Office of Management and Budget (OMB) provided sustained support for agencies' attempts to improve productivity. OPM's significant productivity leadership effort in the late 1970's to provide information clearinghouse services, technical assistance, and productivity training ended in 1980. OMB's management improvement projects under Reform '88 (a coordinated effort designed to modernize government through sound business practices) did not explicitly focus on productivity or address productivity improvement processes. Furthermore, OMB did not provide clear signals to agencies that productivity improvement was a priority issue.

Evolution of Federal Productivity Improvement

Since the early 1980's, the importance attached to federal productivity improvement has grown. Many studies indicate that substantial savings can be realized if federal productivity is improved by even small amounts. GAO's 1983 report, the Grace Commission report, a November 1983 report by the National Academy of Public Administration *(Revitalizing Federal Management: Managers and Their Overburdened Systems)*, and the 1979 Joint Economic Committee study all suggest the large potential for productivity improvements in the federal government and call for various approaches to a achieve these ends.

There is a growing consensus within the federal government that productivity improvement is a major issue and agencies need to develop strategies to increase their productivity. GAO, in 1983, recommended that the Director, OMB, encourage and support productivity improvements throughout the government by

- building on Circular A-11 to require federal agencies to specify in their budgets (1) short- and long-range productivity objectives and goals, (2) anticipated dollar savings from future or sustained efforts, and (3) prior-year dollar savings achieved through productivity improvements;
- requiring agencies to establish productivity management efforts that systematically identify opportunities for improvements;
- ensuring technical assistance was available to agencies developing productivity management efforts; and
- assuming responsibility for the Federal Productivity Measurement Program, administered by BLS, to monitor and encourage productivity improvements in agencies.

President's Productivity Improvement Program

In response to the growing consensus that productivity levels in the federal government are inadequate and substantial savings could be realized through productivity improvements, President Reagan recently signed Executive Order 12552, establishing a comprehensive Productivity Improvement Program for the federal government. The program's principal objective is to improve federal productivity 20 percent by 1992. To do this, the Order directs each federal agency to develop short- and long-term productivity plans to improve the quality, timeliness, and efficiency of their operations.

In a statement following the issuance of the Order, OMB Director James C. Miller stated that "The private sector has for many years been successfully applying productivity improvement techniques to remain competitive. The federal government must do likewise, especially in these times of financial austerity. Higher productivity will squeeze more value from every tax dollar spent on federal programs."

The oversight responsibility for the Productivity Improvement Program rests with OMB. OMB's Bulletin No. 86-8, issued February 28, 1986, outlines the program in detail. Executive agencies are directed to develop productivity improvement plans covering the remainder of fiscal year 1986 and all of fiscal year 1987 and to submit them to OMB for review as part of each agency's Management Improvement Plan. The Bulletin further directs each agency to

- designate a senior official to coordinate the effort;
- develop a productivity improvement plan for each fiscal year as part of its overall management improvement plan;
- specify four or five priority areas or functions offering major gains in quality, timeliness, and efficiency;
- develop a measurement approach that is appropriate for its productivity improvement efforts;
- encourage employee participation in developing and achieving productivity improvements through training and monetary and nonmonetary incentives and awards; and
- report annually by December 1 to OMB on progress made in productivity improvement during the prior fiscal year.

The authors believe that the President's program is a step in the right direction. If properly implemented, operated, and maintained, the program can benefit federal productivity and improve federal service to the public even in an environment of financial restraints.

Our major concern with the program stems, in a sense, from its name—the use of "program." We hope that this effort will not be perceived by agencies as just another program by or reporting requirement to OMB that will fade with time. If agencies do not perceive this effort as serious, they will be prone to go through the motions without expending the necessary time and resources to develop substantial and lasting productivity improvement efforts. Productivity improvement requires a long-term orientation.

OMB is responsible to ensure that agencies get the message that the productivity improvement "program" is serious. One possible way to hold agencies accountable for productivity improvements is to use the information collected through the program in their budget reviews. For example, those agencies that clearly exhibit serious efforts at productivity improvement that result in tangible productivity increases and cost savings might keep a portion of the savings resulting from their productivity improvements. Unless agencies are held accountable for productivity improvements and even penalized if they do not meet the targets set for them, few will approach this effort with the vigor and comprehensive treatment necessary to achieve the 20-percent productivity improvement target set by President Reagan.

The initial reaction from OMB is that the productivity plans for fiscal years 1986 and 1987 are a step in the right direction but need some work. OMB notes that the absence of productivity measurement data is the most serious problem with many productivity plans. Most agencies do not have or did not include productivity data to measure output at the function level that can be related to data on input costs. OMB

realized at the outset that productivity measurements would be one of the most difficult parts of the program and has detailed a measurement expert from BLS to help with this problem. Other OMB concerns with the program's early implementation efforts are as follows:

- Some agencies are not viewing the program as integral parts of their management processes.
- Implementation is being carried out by people in staff positions rather than by line managers.
- Improvement strategies are directed primarily to automation, while little attention is being given to human resource and process improvements.

To help overcome some of the initial problems with the program, OMB has established an information clearinghouse on productivity issues. OMB also has asked OPM to (1) review employee incentive, performance appraisal, and position classification practices and to suggest changes which would facilitate productivity improvements and (2) develop positive employee job placement and retraining programs and productivity training programs.

GAO's National Productivity Group (NPG) has been asked by the Senate Committee on Governmental Affairs to annually evaluate agency productivity improvement plans. NPG will assess each plan on the basis of the criteria previously mentioned, along with specific knowledge of agencies gathered through GAO's agency productivity management reviews. The key criterion will be whether agencies have integrated productivity improvement into their management processes.

Future of Federal Productivity Improvement

The ultimate questions each organization must ask itself when designing a productivity improvement plan are as follows: (1) Does it want a productivity "program" or a management process for productivity improvement? and (2) Does it want productivity improvement to be viewed as a short-term add-on responsibility that will cease to have emphasis after passage of time, or does it want to change the climate of the organization so that productivity becomes an integral element of organizational functioning?

Improving federal productivity has been a recurring goal; yet past efforts have been treated as "just another management program" that will pass as administrations change. However, with the pressures on agencies to maintain service to the public in a period of retrenchment, agencies must use productivity improvements to keep services constant.

Presidential involvement, an executive order making productivity improvement a national priority, and the explicit goal of seeking no less than total integration of productivity goals into the basic processes of government have already set this effort apart from others. GAO's work has gone a long way in convincing OMB and selected agencies that productivity management can be a powerful tool to contain costs and sustain quality service to the public. However, productivity improvement cannot be mandated. It is now up to individual agencies to take the initiative and develop comprehensive strategies for improving productivity. ■

National Productivity Group Efforts to Improve Federal Productivity

GAO's National Productivity Group has been focusing on productivity issues in government for 9 years. Its major focus has been on heightening awareness of federal productivity while identifying areas that offer the potential for substantial productivity improvement. Over the past 5 years, NPG has identified actual cost savings of $988.5 million. NPG reports have ranged from *Gainsharing: DOD Efforts Highlight an Effective Tool for Enhancing Federal Productivity* (GAO/GGD-86-143BR, Sept. 26, 1986) to *Improving Operating and Staffing Practices Can Increase Productivity and Reduce Costs in SSA's Atlanta Region* (GAO/GGD-85-85, Sept. 11, 1985).

Currently, NPG is working on assignments ranging from a survey of productivity management at service centers of the Internal Revenue Service to a review of the impact of human resource management practices on productivity in the federal government. NPG's strategy for increasing federal productivity is twofold: Of NPG's resources, about 50 percent is spent on productivity management reviews of federal agencies; the other 50 percent is used on more-specific reviews of productivity improvement techniques. This strategy allows NPG the latitude to approach productivity improvement from macro and micro perspectives. That is, NPG can evaluate different approaches to improving productivity, as well as agencies' efforts at integrating productivity management into their organizational processes.

NPG's productivity management reviews assess an organization's potential for improving productivity, focusing on the following four questions:

- How do the organization's processes work?
- How well has it performed in terms of efficiency, quality, and timeliness?
- What are the opportunities for performance improvements?
- What role has management played in organizational performance?

The assessment is conducted in four parts. Part one focuses on answering the question on how the organization has performed. Parts two and three identify process reasons why the organization may not be performing well and develop potential improvements through procedural, technological, and organizational changes. Part four identifies management reasons why the organization may not be performing well and assesses management's role in the organization's performance.

A Week's Worth

Lenora V. Brown

Ms. Brown joined GAO in 1981 as an evaluator with the Kansas City Regional Office. At present, she is an equal employment opportunity counselor for the region. During the last several years, Ms. Brown has worked primarily on assignments involving information management and technology. She is currently working toward a master's degree in management information systems. Ms. Brown began her career in 1976 at the Department of the Army while completing her master's degree in clinical psychology. This degree and her B.A. in psychology were earned at Southern Illinois University.

Monday

I arise before dawn on Monday morning to complete last-minute chores before driving to the airport. I must admit that I am looking forward to this trip to Washington, D.C. After spending a week each in Little Rock, Arkansas, and Grand Forks, North Dakota, I find that the fast-paced excitement of the "Capital City" is most appealing. I reach the gate in time for the 7 a.m. flight, and Laura Durland, my coworker, is relieved that I made it in time.

After landing, Laura and I proceed to a 1 p.m. entrance conference at the Defense Communications Agency (DCA) in Arlington, Virginia. DCA is an agency of the Department of Defense (DOD) under the direction of the Under Secretary of Defense (Research and Engineering). DCA performs system engineering and system architect functions for various military communication systems.

Another DCA function is to procure leased communication circuits, services, facilities, and equipment for DOD, where authorized, and for other government agenies, as directed by the Secretary of Defense. The purpose of our visit relates to this particular responsibility. We previously have done work at the Defense Commercial Communications Office, which is a field organization of DCA.

This audit was requested by the Subcommittee on Defense, House Committee on Appropriations. The audit objective is to survey DOD equipment-leasing practices at selected sites to determine the extent of lease-versus-purchase agreements and to determine whether a lease-versus-purchase analysis entered into the procurement decision. We are reviewing all leased equipment, including telephone systems, word processors, copiers, security alarm systems, medical equipment, and vehicles. Our goal is to define the current condition and report to the Committee.

This assignment is under the purview of the Information Management and Technology Division (IMTEC). The Committee's request is being handled by IMTEC because the issue involves telecommunications. IMTEC's role includes evaluator-in-charge responsibilities. Because the review concerns defense, IMTEC is also coordinating with the National Security and International Affairs Division on all work done. The Norfolk and Los Angeles regions are also involved.

At 2:30 p.m., we are escorted to a meeting with the Vice-Commander to discuss the purpose of our work. We also discuss previous GAO efforts at several other Air Force bases and other background data. Laura and I are eager to begin the actual audit work, which will entail interviewing agency officials and reviewing accountable property records, financial documents, and contract files.

We spend the rest of the afternoon with an official from DCA's Policy Office who will serve as our contact during our visit. We discuss in more detail the type of information we are seeking. We arrange for a full day of interviewing on Tuesday. With this much done, Laura and I leave to catch the DOD shuttle bus back to the Pentagon. From there we take the Metro, Washington's public transit system, to Crystal City, a nearby suburb and site of our hotel.

Tuesday

We get an early start to meet with our contact. He informs us that he is transferring responsibility for the audit, introduces us to our new contact from the Personnel and Administration Office, and departs. I repeat yesterday's briefing, which outlines the

Legislative Developments

Judith G. Hatter

Employment Opportunities for Disabled Americans Act

On March 19, Senator Robert Dole of Kansas introduced S. 2209 to make permanent and improve section 1619 of the Social Security Act, which authorizes the continued payment of Supplemental Security Income benefits to individuals who work despite severe medical impairment. The bill requires GAO to conduct a study of the operations of section 1619 to evaluate the program's effectiveness.

Federal Management Reorganization and Cost Control Act

On March 26, Senator William Roth of Delaware introduced S. 2230, the Federal Management Reorganization and Cost Control Act, to improve the management of the government by establishing an Office of Federal Management in the Executive Office of the President. Reports to the Congress by the Comptroller General are required with respect to a 5-year plan for the financial management activities of the federal government (to be prepared by the Office of Financial Systems of the Office of Federal Management) and an annual report on management of the executive branch by the President. (See this issue's "From Our Briefcase," p. 1 , for details.)

Anti-Kickback Enforcement

On March 26, Senator Carl Levin of Michigan introduced S. 2250, the Anti-Kickback Enforcement Act, to strengthen the prohibition of kickbacks relating to subcontracts under federal government contracts. The amendment authorizes GAO to conduct audits to investigate violations of the act.

Indian Gaming Regulatory Act

On April 21, the House of Representatives amended and passed, under suspension of the rules, H.R. 1920, to establish federal standards and regulations for the conduct of gaming activities on Indian reservations and lands. The bill establishes a National Indian Gaming Commission to implement and administer the act. Gaming is divided into three classes: class I, defined as social and traditional Indian gaming, which is left to the tribes exclusively; class II, defined as bingo and related games, which must be authorized by a tribal ordinance; and class III, defined as all other forms of gaming, including casinos and horse and dog racing. The bill provides for a study and report to the Congress by the Comptroller General on class II gaming on Indian lands.

Education of the Deaf Act

On May 6, the Senate passed S. 1874, Education of the Deaf Act, with a committee amendment in the nature of a substitute. The bill provides for audit by GAO of financial transactions in connection with the expenditure of appropriated funds for the benefit of Gallaudet University or for the National Technical Institute for the Deaf. The bill also provides for appointment by the Comptroller General of one member of the Commission on Education of the Deaf, which is established by the legislation. ■

purpose of our audit, for various department officials who should be able to provide the information we seek.

After this, we are directed to a partitioned area furnished with a desk, a table, and a telephone—our work area for the rest of the week. As I settle in to review notes for the upcoming interviews and to check in with my supervisor in St. Louis, a problem arises. The contact informs me that since security is very tight and Laura's security clearance is in process, the agency will not issue her a "no escort required" visitors badge. Laura will have to be escorted throughout all secured areas. Furthermore, unless someone sits with us at all times, we will have to find work space in a nonsecured area. "This," the contact says, "may take a little time to work out."

It is almost lunchtime; I ask that we be escorted to the cafeteria, which is a nonsecured area. Laura and I may have lunch unaccompanied but must call the contact from the guard's desk to be escorted to the next location. Neither of us is looking forward to working under these conditions.

After lunch, our escort meets us at the guard's desk and takes us to our meeting with an official from Finance. Since he was not present at either of our briefings, I again explain the purpose of our audit and the type of information we are seeking. The official explains that the information we need can be obtained from Finance's computer system by using a data retrieval program. This may prove difficult, though, because the agency is in the midst of year-end processing to be followed by a changeover from one computer mainframe to another. "This changeover," the official says, "will change the procedures for obtaining information from the system." He hopes he will be able to honor our request between the two events sometime during the following week. Since our visit is scheduled to end on Friday, I leave our mailing address for him to forward the information to us.

Next we meet with the head of Contract Management. The records are ready for our review, and a workplace has been prepared for us.

As we review the files, we find cases in which software was leased. We have not had cases like this at the other locations visited and are uncertain as to whether software should be considered equipment. I call St. Louis to obtain clarification from my supervisor as to whether software is leased property or not, and Laura and I continue through the files. The office assistant notes that the copies we want will be ready tomorrow and that we can pick them up at our convenience. With this work completed, our day is at an end. After dinner, at the Orleans House restaurant, we return to our hotel.

Wednesday

Today our contact has located workspace in a nonsecured area for us. This eliminates the need to have an agency official with us at all times. Welcome news indeed! To alleviate another problem, Laura's not being issued a "no escort required" badge, I have decided to do the "legwork" of data gathering (that is, collecting documents and conducting follow-up interviews) while Laura remains in our new workspace doing preliminary data analysis.

This morning, I meet with officials from Administrative Services. They are responsible for leasing word processors, copiers, and other office equipment for the agency. We discuss the policies and procedures they use in carrying out their leasing activities, and as usual, I request the necessary documentation. It should be ready for me by the end of the day.

This afternoon, I make a follow-up visit to the head of Contract Management to clarify some of yesterday's discussion points. Once the official purpose is accomplished, we chat about autumn's changing colors. My curiosity gets the best of me, and I venture a question about his British accent. He reflects on his British heritage and adds, "Ah, England is lovely in the spring."

Afterwards, I stop at the office assistant's desk to pick up the copies I requested yesterday. Tonight's agenda consists of a quick dinner at a cafeteria in Crystal City and a full night of sports excitement watching the National League Championship Playoff game between the St. Louis Cardinals and the Los Angeles Dodgers.

Thursday

I have set aside the entire day to discuss cost analysis procedures with the agency personnel. Initially, I am directed to one individual, but it soon becomes apparent that not all of my questions will be answered here. Several more interviews provide a fairly complete picture of the cost analysis process, but I still need written documentation (such as agency regulations or office memorandums) to support the testimonial evidence.

Tomorrow is our last day here, and as customary, I will conduct an exit conference with agency officials summarizing our findings. I jot down a brief outline and phone our contact to confirm the conference time. As I review my notes, I feel satisfied with what we have done. Then, I call it a day.

Friday

Smooth sailing this morning. The exit conference begins at 10 a.m. Laura and I present our information, respond to questions, and then leave for our flight with plenty of time to spare. Once we are on board, the pilot announces that our departure will be delayed due to weather conditions in St. Louis.

About an hour later, when our plane finally takes off, I mentally review our productive, but sometimes frustrating, week. In my 5 years as an evaluator, I have traveled to 22 states and have made numerous trips to Washington, D.C. Each trip has provided me with unique professional experiences that have contributed to growth in my personal life. As we land in St. Louis, I can appreciate the old adage "There's no place like home." ■

Location (continued from page 1)

The award, which included a plaque and cash, was presented by Assistant Comptroller General and *Review* Editor Harry Havens. Ms. Hatter was nominated for the award by *Review* editorial staff as a fitting complement to the *Review's* "Twentieth Anniversary Issue" (winter 1986). The regular feature, "Legislative Developments," is an important archive and summary of legislation that affects GAO's functions and responsibilities. The column was entitled "Hearings and Legislation" when Ms. Hatter's predecessor, Ms. Margaret MacFarlane, wrote it. It grew out of the work of the Law Library in developing legislative histories on all laws enacted by the Congress. During the informal awards presentation, it was noted that Ms. Hatter's column had certainly earned the accolade "longest running feature" in *The GAO Review.*

SFRO Employees Create Their Own Dental Insurance Program

San Francisco Regional Office (SFRO) staff took advantage of enabling legislation last spring to establish GAO's only employee-sponsored health insurance program. The new program was created over a 14-month period to offer dental insurance protection to SFRO staff at group rates. Nine persons joined the "San Francisco Dental Health Group" during the first annual open season held during 3 weeks in April and May.

The program grew out of SFRO staff interest in obtaining dental insurance for themselves and their families. Some staff reported that they could not afford to pay the premiums for federal health insurance plans having dental coverage. Others complained that the consolidated medical-dental plans did not provide adequate medical protection.

SFRO staff established a committee to launch the program, which began by screening and selecting dental insurance plans. The committee selected two plans out of the more than a dozen proposals received and distributed enrollment cards in SFRO and the Sacramento Suboffice during open season.

The two plans allow individuals and couples to reduce the amount they pay for dental services each year. Additional savings are also possible for staff who include dependent children in their plans.

Savings are achieved by not having to pay for basic services, such as check-ups, x-rays, and teeth cleaning, and by discount rates for other services. One of the two plans allows staff to select a dentist near their workplace while permitting the rest of the family to choose a dental office closer to home.

The committee initially encountered an obstacle posed by the GAO Personnel Act in the form of a provision prohibiting the agency from issuing insurance contracts. This provision raised questions about whether staff could issue contracts. Further investigation revealed that the act permits employee-sponsored activities benefiting the agency or staff and thus enabled SFRO staff to legally obtain insurance services.

The San Francisco Dental Health Group is managed by officers selected by the members of the group. GAO staff in other locations can contact Bruce McClellin, the Group Representative, at (415) 556-6200 for additional information. ■

Manager's (continued from page 7)

creased responsibilities, and establishing a divisional audit site in a regional office.

The author, in presenting his model, clearly states that it is too early to advance it as a general theory that would be helpful in all situations. Likewise, it is too early to evaluate GAO's program. Early signs are encouraging; however, more time will be needed to determine if desired changes result. Clearly, GAO's approach, like the model, is aimed at implementing change required to ensure the organization's long-term health.

Ed. note: A videotape of the June 20, 1986, Executive Speakers Program presentation by Dr. George Odiorne, entitled "The Strategic Management of Human Resources," is available for viewing. Odiorne's discussion before GAO's senior executives focused on how human resource management approaches can maximize organizational and individual effectiveness, especially in times of change. The videotape includes Odiorne's formal presentation and the ensuing question-and-answer period. Call Rusty Glazer, OOHD, (202) 272-3475, to arrange to see the tape. ■

Topics (continued from page 10)

uation Research. Beverly Hills, Calif.: Sage, 1979. A widely cited book which contrasts the qualitative and quantitative research traditions.

Hufbauer, G. C., J. J. Schott, and K. A. Elliott. *Economic Sanctions Reconsidered: History and Current Policy.* Washington D.C.: Institute for International Economics, 1985. A compendium of economic sanction episodes plus quantitative analyses.

Miles, M. B., and A. M. Huberman. *Qualitative Data Analysis: A Sourcebook of New Methods.* Beverly Hills, Calif.: Sage, 1984. Though many talk about doing qualitative data analysis, Miles and Huberman *show* how it is done.

Tukey, J. *Exploratory Data Analysis.* Reading, Mass.: Addison-Wesley, 1977. Presents an outlook on data analysis different from the classical statistical approach; includes a variety of specific techniques. ■

Electronic (continued from page 11)

Prompt technical review is also desirable. The Chicago Regional Office's Technical Assistance Group, as a rule, conducts its technical review as soon as possible after the application is completed.

Overall, it is actually the microcomputer user who makes the difference. Microcomputers can make bigger mistakes faster, but they also can be programmed to catch their own errors. Since we evaluators are already familiar with controls, we simply must (1) acquaint ourselves with the capabilities and idiosyncrasies of the hardware and software that we plan to use and (2) plan the application to capitalize on the microcomputer's ability to assist in error detection and correction. While much drudgery remains, it is a far cry from the days when this entire effort had to be done manually. ■

Census (continued from page 21)

will play a major part in helping the congressional oversight committees perform their role.

See Census, p. 49

Senior GAO Staff Changes

Ed. note: The staff changes listed in this and the following sections occurred during the approximate period January to March 1986.

Henry Eschwege

Mr. Henry Eschwege retired from GAO in March 1986, ending 30 years of GAO service. All but 2 of Mr. Eschwege's 32 years of federal service were spent with GAO.

Mr. Eschwege, who had served as Assistant Comptroller General for Planning and Reporting since 1982, had been responsible for determining the overall direction of GAO's work and for the quality of reports on specific assignments. He served in various supervisory positions in GAO before being named Director of the newly formed Resources and Economic Development (later the Community and Economic Development) Division in 1972.

Over the years, Mr. Eschwege was frequently recognized for his achievements and contributions to GAO. He received Meritorious Service Awards in 1965 and 1967, a Distinguished Service Award in 1968, and Comptroller General's Awards in 1977 and 1978. As a member of GAO's Senior Executive Service, he was named Distinguished Executive in 1981 and Meritorious Executive in 1985.

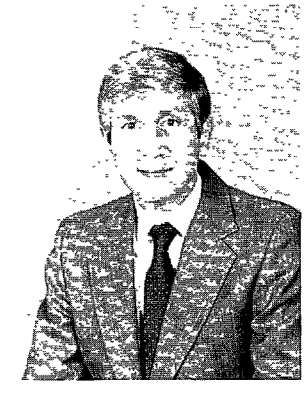
Donald J. Horan

Mr. Donald J. Horan was appointed Assistant Comptroller General for Planning and Reporting in March 1986. Mr. Horan had been serving as Director, Office of Policy, since May 1983.

After joining GAO, Mr. Horan worked in the New York Regional Office before transferring to the Audit Policy Staff of the Office of Policy and Special Studies in 1965. In 1968, he was named Assistant Director for Auditing Policy. He then served as an Assistant Director in the Procurement and Systems Acquisition Division from 1972 to 1974, when he was named Director, Office of Policy. In 1978, he was named Deputy Director, Logistics and Communications Division, and in 1981, he became Director, Procurement, Logistics, and Readiness Division.

Mr. Horan received a B.S. in accounting from King's College in Pennsylvania in 1955. In 1982, he completed the Executive Program in National and International Security at Harvard University's John F. Kennedy School of Government.

He received the Meritorious Service Award in 1968 and the Comptroller General's Award in 1978. In 1981, he received the rank of Meritorious Executive in the Senior Executive Service. Mr. Horan is a member of the Washington Chapter of the Association of Government Accountants.



Ralph V. Carlone

Mr. Ralph V. Carlone was named Deputy Director for Operations, Information Management and Technology Division, in March 1986.

After joining GAO in 1964, Mr. Carlone's responsibilities included audits at the Veterans Administration and the former Atomic Energy Commission. From 1976 to 1978, he served as Associate Director in the Energy and Minerals Division, where he was responsible for the activities of the Nuclear Regulatory Commission and for audit analysis of the Department of Energy's research and development programs. From 1978 to 1983, Mr. Carlone was the Regional Manager of the Philadelphia Regional Office. From 1983 to 1986, he served as Deputy Director for Planning and Reporting in the Resources, Community, and Economic Development Division.

Mr. Carlone has received the William A. Jump Memorial Award for Outstanding Public Service, the Distinguished Service Award, the Energy and Minerals Division Director's Award, and a Meritorious Service Award. He also held the rank of Meritorious Executive in the Senior Executive Service.

A graduate of Bloomsburg State College in Pennsylvania, Mr. Carlone served in the U.S. Marine Corps from 1956 to 1960. In addition, Mr. Carlone has attended executive development seminars at the Dartmouth Institute.

Neal P. Curtin

Mr. Neal P. Curtin was appointed Deputy Director for Planning and Reporting in the Resources, Community, and Economic Development Division in March 1986.

He joined GAO in 1970 in the Chicago Regional Office. He was assigned to the European Branch in Frankfurt, West Germany, from 1974 to 1978. Since returning to Washington in 1978, Mr. Curtin has worked in the Human Resources Division and the former International Division, served as a report reviewer in the Office of Quality Assurance when it was first established, and managed the report review function in the National Security and International Affairs Division (NSIAD) from its beginning in 1983 until November 1984. He was selected for the third Executive Candidate Development Program in 1984 and was appointed Director of the Office of Quality Assurance in April 1985.

Mr. Curtin graduated from Bradley University in Peoria, Illinois, with a B.S. in economics. He has received a Meritorious Service Award (1981) and the NSIAD Director's Award (1984), among other GAO awards. He is a member of the American Society for Public Administration.

Lowell Dodge

Mr. Lowell Dodge was appointed Director, Office of Affirmative Action Plans, in January 1986.

Mr. Dodge joined GAO in 1981 and served as Associate Director for Economic and Area Development in the former Community and Economic Development Division and the Resources, Community, and Economic Development Division until joining the Information Management and Technology Division in 1983.

A graduate of Yale University, Mr. Dodge earned a law degree from Harvard, then came to Washington in 1969 as Executive Director of the Center for Auto Safety, a citizen advocacy organization. From 1974 to 1975, he was Editor-in-Chief of the *Environmental Law Reporter.* Thereafter, he served as Staff Counsel for the Subcommittee on Oversight and Investigations, House Committee on Energy and Commerce. In 1978, Mr. Dodge joined the Consumer Product Safety Commission as Executive/Legal Assistant to the Chairman.

Mr. Dodge was active in the civil rights movement in the 1960's. He has worked as a community organizer in Washington, D.C.; Los Angeles; and North Carolina, concentrating on educational reform and urban tutorial programs.

Dennis J. Duquette

Mr. Dennis J. Duquette was named Associate Director of the Financial Audit Group within the Accounting and Financial Management Division (AFMD) in January 1986.

Mr. Duquette joined the Washington Regional Office in 1974. Following an assignment on GAO's 1976-77 banking task force, he joined the General Government Division and continued his work in the area of financial institution regulation. In 1980, he joined AFMD, where he was responsible for consolidating GAO's financial statement audit work. Mr. Duquette was selected for GAO's Executive Candidate Development Program in 1984.

Mr. Duquette graduated from Fordham University in New York in 1968 with a B.S. in accounting. He is a certified public accountant in New York and a member of the American Institute of Certified Public Accountants and the New York and Virginia State CPA Societies. He is also a member of the American Accounting Association, where he is active in the government and nonprofit section, and the Association of Government Accountants. Mr. Duquette has received a GAO Meritorious Service Award and Certificates of Appreciation for his work.

Frances Garcia

Ms. Frances Garcia was appointed as Assistant to the Assistant Comptroller General for Operations in January 1986.

Ms. Garcia came to GAO after working for 3 years as a partner for marketing and practice development with Quezada Navarro and Company, the largest Hispanic-owned accounting firm in the United States. Ms. Garcia spearheaded the firm's practice development at the federal government level and was responsible for opening the Washington, D.C., office.

Before joining Quezada Navarro and Company, Ms. Garcia was appointed by President Jimmy Carter to serve as Commissioner and Chair for the U.S. Copyright Royalty Tribunal. From 1968 to 1977, she worked at Arthur Andersen & Co., the last 2 years as a senior manager for the firm's Austin, Texas, office.

She holds a B.S. in business administration from Midwestern University, Wichita Falls, Texas, and is a certified public accountant in Texas and the District of Columbia. She is a member of the American Institute of Certified Public Accountants' Government Executive Committee.

Edward P. Henderson

Mr. Edward P. Henderson was appointed Special Assistant to the Director, Financial Audits and Accounting Policy, Accounting and Financial Management Division, in February 1986.

Mr. Henderson came to GAO from the New York State Comptroller's Office, where he had served for 3 years as Deputy Comptroller. In this position, he was responsible for the division that handles the accounting and auditing of financial transactions of New York State, the largest nonfederal fiscal entity in the United States.

From 1980 to 1983, Mr. Henderson served as the Director of Financial Audits for the Oklahoma State Auditor and Inspector. During a 1973-83 career with Lester Witte & Company, a national accounting firm, Mr. Henderson progressed from western regional audit partner to managing partner of the Houston office.

Mr. Henderson received his B.B.A. in accounting from Woodbury University in Los Angeles in 1958. He is a certified public accountant in California and Texas and a member of the Virginia CPA Society. He is an active member of the National Association of State Auditors, Comptrollers, and Treasurers, having served as Chairman of the State Auditors Committee on Single Audit Implementation.

James D. Martin

Mr. James D. Martin was appointed Regional Manager of the Atlanta Regional Office in January 1986.

Since joining GAO in 1958, Mr. Martin has served in the Office of Program Planning, the European Branch of the former International Division, and the Human Resources Division. In November 1980, Mr. Martin was named Regional Manager of the Dallas Regional Office.

Mr. Martin received his B.S. in accounting from Central Missouri State College in 1958 and attended the Program for Management Development at the Harvard University Business School in 1967. He is a certified public accountant in Virginia and a member of the American Institute of Certified Public Accountants and the Association of Government Accountants.

Mr. Martin received the GAO Career Development Award in 1967 and the Association of Government Accountants (Washington Chapter) Outstanding Achievement Award and Achievement of the Year Award in 1973. He headed the task force on health facilities construction costs, which received the Comptroller General's Award in 1973.

Michael E. Motley

Mr. Michael E. Motley was named Associate Director for the Research, Development, Acquisition, and Procurement Subdivision in the National Security and International Affairs Division in March 1986.

A graduate of Belmont Abbey College in North Carolina, Mr. Motley joined GAO in 1969. During his career, he has had two assignments to Capitol Hill and audited several civil agencies, including the National Aeronautics and Space Administration, the National Institutes of Health, and the Department of Health and Human Services. In addition, he worked in the Human Resources Division audit group that reviewed the Department of Defense's health care facilities and programs. From 1980 to 1984, Mr. Motley worked in the Office of Congressional Relations. In 1984, he was selected for the fourth Executive Candidate Development Program.

Mr. Motley has been an active member of the National Association of Accountants (NAA), where he served as president of the Northern Virginia Chapter in 1983, among other positions. He is currently Principal of the Potomac and Chesapeake Regional Council representing several NAA chapters. He has received several GAO awards and the Jaycees' Outstanding Young Men of America Award in 1983. Mr. Motley attended the Program for Senior Executive Fellows at the John F. Kennedy School of Government at Harvard University in the fall of 1985.

Joe E. Totten

Mr. Joe E. Totten assumed the position of Regional Manager of the Chicago Regional Office in January 1986.

Since joining GAO in 1962, Mr. Totten has served in the Civil Division, the Resources and Economic Development Division, and the Human Resources Division while concentrating on housing, forestry, education, occupational safety, and health programs. His most recent assignments have included working with the 1984, Post Assignment Quality Review System (PAQRS) team and serving as an Acting Associate Director in the National Security and International Affairs Division. In 1984, Mr. Totten was selected for the third Executive Candidate Development Program.

Mr. Totten received his B.S. in accounting in 1962 from the West Virginia Institute of Technology. He is a certified public accountant in West Virginia. Mr. Totten received the Comptroller General's Career Development Award in 1973, the Human

See Senior Staff, p. 49

Other GAO Staff Changes

SES Promotions

Name	Position	From	To
Boland, F. Kevin	Deputy Director for Operations in the Resources, Community, and Economic Development Division	ES-4	ES-5
Curtin, Neal P.	Deputy Director for Planning and Reporting in the Resources, Community, and Economic Development Division	ES-1	ES-3
Datta, Lois-ellin	Associate Director, Program Evaluation and Methodology Division	ES-3	ES-4
Dodaro, Gene L.	Associate Director, General Government Division	ES-1	ES-2
Gainer, William J.	Associate Director, Human Resources Division	ES-2	ES-3
Grant, George E.	Regional Manager, Los Angeles Regional Office	ES-3	ES-4
Grosshans, Werner	Director, Office of Program Planning	ES-4	ES-5
Hamilton, Mary R.	Regional Manager, New York Regional Office	ES-2	ES-3
Hanna, David A.	Regional Manager, Kansas City Regional Office	ES-4	ES-5
Herrmann, Walter C., Jr.	Regional Manager, Detroit Regional Office	ES-4	ES-5
Kepplinger, Gary L.	Assistant General Counsel, Office of the General Counsel	ES-1	ES-2
Kleeman, Rosslyn S.	Associate Director, General Government Division	ES-3	ES-4
Lauve, Ronald P.	Regional Manager, Washington Regional Office	ES-4	ES-5
Martin, James D.	Regional Manager, Atlanta Regional Office	ES-4	ES-5
McCormick, Thomas P.	Regional Manager, San Francisco Regional Office	ES-4	ES-5
McLure, Herbert R.	Associate Director, Resources, Community, and Economic Development Division	ES-3	ES-4
Mendelowitz, Allan I.	Associate Director, National Security and International Affairs Division	ES-3	ES-4

SES Promotions (cont.)

Name	Position	From	To
Rist, Ray C.	Deputy Director, Program Evaluation and Methodology Division	ES-3	ES-4
Simmons, Craig A.	Associate Director, General Government Division	ES-3	ES-4
Stillman, Rona	Chief Scientist, Information Management and Technology Division	ES-3	ES-4
Thompson, Lawrence H.	Chief Economist, Office of the Chief Economist	ES-5	ES-6
White, Daniel C.	Deputy Director for Planning and Reporting, Information Management and Technology Division	ES-3	ES-4

Additional Staff Changes

Name	Division	Title
Broderick, Thomas R.	Accounting and Financial Management	Group Director
Clark, David L., Jr.	Accounting and Financial Management	Group Director
Culkin, Charles W., Jr.	Accounting and Financial Management	Group Director
Gannon, Arthur G.	National Security and International Affairs	Assistant to the Director for Human Resources

New Staff Members

Name	Division/Office	From
Anderson, William	Accounting and Financial Management	University of Maryland
Boyer, Danette	Accounting and Financial Management	Western Illinois University
Carmichael, Mary	Accounting and Financial Management	Florida State University
Driscoll, Cheryl	Accounting and Financial Management	University of Maryland
Hench, Sharon	Accounting and Financial Management	Penn State University
Hynes, Richard	Accounting and Financial Management	University of Maryland
Mitchell, Beverly	Accounting and Financial Management	Honeywell, Inc.
Morse, George	Accounting and Financial Management	Christopher Newport College
O'Connor, John	Accounting and Financial Management	University of Lowell
Simms, Ross	Accounting and Financial Management	Army Audit Agency
Getz, Sandra K.	General Government	Not specified
Whitford, Cynthia L.	General Government	Northern Virginia Community College
Hellman, Steven	Human Resources	American University
Jeszek, Charles	Human Resources	University of California
Shaw, Lois	Human Resources	Ohio State University

New Staff Members (Cont.)

Name	Division/Office	From
Buckey, Kristin L.	National Security and International Affairs	Columbia University
Dyer, Sharon L.	National Security and International Affairs	Department of Agriculture
Guilliams-Tapia, Judy L.	National Security and International Affairs	Wayne State University
Hoagland, Daniel C.	National Security and International Affairs	U.S. Army
Knepper, Judith K.	National Security and International Affairs	University of Maryland
Peterson, Linda F.	National Security and International Affairs	Mt. Sinai Day School
Demlo, Linda K.	Program Evaluation and Methodology	Health Care Financing Administration
Benson, Audrey	Resources, Community, and Economic Development	George Washington University
Burros, Michael	Resources, Community, and Economic Development	Congressional Staff Office
Cates, Leah	Resources, Community, and Economic Development	Department of Education
Gillespie, George	Resources, Community, and Economic Development	George Washington University
Goldenkoff, Robert	Resources, Community, and Economic Development	George Washington University
Hampton, Matthew	Resources, Community, and Economic Development	American University
Knight, Gregory	Resources, Community, and Economic Development	Congressional Research Service
Shulman, Carol	Resources, Community, and Economic Development	Self-employed
Perruso, Richard R.	General Counsel	Boston University School of Law
Hutchinson, Angela	Financial Management	Private industry
Newman, Anita	Library Services	Pension Benefit Guaranty Corporation
Daugherty, Loraine L.	Information Resources Management	Department of Transportation
Morris, Charles B.	Information Resources Management	U.S. Army
Richardson, Glen M.	Information Resources Management	Republic Bank Corporation
Rivers, Charles T.	Information Resources Management	General Services Administration
Cox, Howard A.	Personnel	U.S. Marine Corps

Name	Regional Office	From
Brannon, Christopher T.	Atlanta	University of Georgia

New Staff Members (Cont.)

Name	Regional Office	From
Chastain, Marion S.	Atlanta	University of South Carolina
Darden, Patricia A.	Atlanta	Department of the Air Force
Garrett, George E.	Atlanta	University of Georgia
Lacy, James J.	Atlanta	University of Georgia
Landers, James H., III	Atlanta	Georgia State University
Ballard, Robert M.	Boston	Bryant College
Chin, Toy H.	Boston	Northeastern University
Johnson, Elmer	Boston	Suffolk University
Krinsky, Lisa E.	Boston	New England Medical Center
Starkey, Beth E.	Boston	Mt. Wachusett Community College
Taylor, Steven C.	Boston	Boston University
Nieto, Sandra	Dallas	Pan American University
Watson, Christine	Dallas	North Texas State University
Rael, Gloria J.	Denver	New Mexico Highlands University
Hubbard, Susan L.	Detroit	Wayne State University
O'Callaghan, Richard E.	Detroit	Kent State University
Westfall, Lynette A.	Detroit	Kent State University
Dorlac, Rose M.	Kansas City	University of Missouri, Kansas City
Welch, Nancy L.	Kansas City	University of Missouri, Kansas City
Dinapoli, Timothy	New York	Maxwell School, Syracuse University
Edwards, Stacy	New York	University of California
LaVelle, Dorothy	New York	American Field Service International
Pineault, Melissa	New York	Maxwell School, Syracuse University
Reilly, Eileen	New York	Citizens' Committee for New York City
Roach, Joann	New York	Veterans Administration
Sherwood, Cynthia	New York	Maxwell School, Syracuse University
Valentin, Sarita	New York	University of Puerto Rico
Harris, Darlene	Philadelphia	University of Pittsburgh
Maradeo, Janette	Philadelphia	Villanova University
Marshall, Andrew	Philadelphia	Bethune-Cookman College
Rullo, Marianne	Philadelphia	St. Joseph's University

New Staff Members (Cont.)

Name	Regional Office	From
Vera, William	Philadelphia	Aviation Supply Office
Waters, Irvenia	San Francisco	Department of Justice
Chaijaroen, Vallapha	Seattle	Private industry
Davis, Valerie	Seattle	Private industry
Carlise, Douglas H., Jr.	Washington	University of Virginia
Kollins, Susan	Washington	American University
Tapia-Videla, Juan F.	Washington	Wayne State University
Toda, Mark	Washington	Towson State University

Attritions

Name	Division/Office	Title
Berrios, Jorge L.	Accounting and Financial Management	Accountant
Cook, Gectris D.	Accounting and Financial Management	Secretary
Gelfand, Dennis	Accounting and Financial Management	Accountant
Giordano, Gina	Accounting and Financial Management	Accountant
McAndrew, Charles	Accounting and Financial Management	Systems Accountant
Carpenter, Rosemary	General Government	Evaluator
French, Richard E.	General Government	Evaluator
Johnson, Gail	General Government	Social Science Analyst
Katcher, Robert A.	General Government	Group Director
Murph, Jacqueline S.	General Government	Secretary
Philip, Constance	General Government	Clerk-Typist
Ketter, Ronald	Human Resources	Evaluator
Slattery, Jane	National Security and International Affairs	Evaluator
Sullivan, Arthur E., Jr.	National Security and International Affairs	Evaluator
Washington, Stephanie L.	National Security and International Affairs	Administrative Operations Assistant
Astor, Richard	Resources, Community, and Economic Development	Evaluator
Baldwin, Barbara	Resources, Community, and Economic Development	Secretary
Crowner, Evelyn	Resources, Community, and Economic Development	Secretary

Attritions (cont.)

Name	Division/Office	Title
Daniel, Beverly	Resources, Community, and Economic Development	Evaluator
Edmondson, Carlton	Resources, Community, and Economic Development	Evaluator
Kuchinsky, Catherine	Resources, Community, and Economic Development	Evaluator
Procter, Robert	Resources, Community, and Economic Development	Economist
Psoras, Cynthia	Resources, Community, and Economic Development	Evaluator
Reiger, Arthur	Resources, Community, and Economic Development	Housing Finance Specialist
Rodrigues, Lillian	Resources, Community, and Economic Development	Secretary
Schaefer, Edward	Resources, Community, and Economic Development	Agricultural Economist
Shepherd, Jacqueline	Resources, Community, and Economic Development	Secretary
Souther, Laura	Resources, Community, and Economic Development	Clerk
West, Doretha	Resources, Community, and Economic Development	Secretary
Brown, Charles L., III	General Counsel	Attorney-Adviser
Chalpin, Mark G.	General Counsel	Attorney-Adviser
Fitzmaurice, Edward L.	General Counsel	Attorney-Adviser
Ikwild, Jeannette M.	General Counsel	Attorney-Adviser
Sevigny, Robert J.	General Counsel	Attorney-Adviser
Jackson, Juanita	Director, General Services and Controller	Clerk-Typist
Magee, Christine	Director, General Services and Controller	Secretary
Industrious, Glanville	Acquisition Management	Purchasing Agent
Hudak, Elizabeth	Administrative Management	Management Analyst
Hartinger, Robert	Financial Management	Travel Accounts Assistant
Hill, Trinita	Financial Management	Teller
Moore, Marie	Financial Management	Travel Clerk
Herrell, Daniel B.	Information Resources Management	Computer Specialist
Ahmad, Rasheedah	Personnel	Clerical Assistant

Attritions (cont.)

Name	Regional Office	Title
Bivins, Gail	Atlanta	Clerk-Typist
McGee, Rebecca H.	Atlanta	Evaluator
Murley, David W.	Atlanta	Evaluator
Smith, Jessee J.	Atlanta	Evaluator
Watkins, Norris W.	Atlanta	Evaluator
West, Thomas J.	Atlanta	Evaluator
Benson, Frank M., Jr.	Boston	Evaluator
Morris, Regina L.	Boston	Evaluator
Hardin, George	Chicago	Evaluator
Krause, Alan	Chicago	Evaluator
Link, Kristine	Chicago	Evaluator
Roque, Phillip	Chicago	Evaluator
Hammond, Gerald G.	Denver	Evaluator
Rizzi, Richard W.	Denver	Evaluator
Long, Alfred C.	Detroit	Evaluator
Hammond, Timothy J.	Kansas City	Evaluator
Lincoln, Robin L.	Kansas City	Receptionist
Whitworth, Keith	Kansas City	Evaluator
Alexander, Sharon	Los Angeles	Clerk-Typist
Boudreau, Roger	Los Angeles	Evaluator
Crayton, Sondra	Philadelphia	Evaluator
Conti, Rodney R.	Seattle	Senior Evaluator
Bagley, Michael	Washington	Evaluator
Bartel, Margaret A.	Washington	Evaluator
Bartha, James D.	Washington	Evaluator
Benton, Gregory M.	Washington	Evaluator
Bolz, Steven D.	Washington	Evaluator
Conlon, Tracy S.	Washington	Evaluator
Efford, Richard E.	Washington	Evaluator
Lee, Susan F.	Washington	Evaluator
Leefman, Barbara J.	Washington	Evaluator
Mullins, James S.	Washington	Evaluator

Attritions (cont.)

Name	Regional Office	Title
Scott, Dean	Washington	Evaluator
Shafer, Gary	Washington	Evaluator
Stukes, Phillip E.	Washington	Evaluator

Retirements

Name	Division/Office	Title
Dacey, John J.	General Government	Evaluator
Heilmeir, Ludwig O.	General Government	Evaluator
Lucas, George B.	General Government	Evaluator
Iffert, Robert	Human Resources	Evaluator
Derricotte, Gladys N.	National Security and International Affairs	Secretary
Gorman, Richard L.	National Security and International Affairs	Evaluator
Hall, Robert B.	National Security and International Affairs	Group Director
Parsons, Robert E.	National Security and International Affairs	Evaluator
Fink, Lou	Program Evaluation and Methodology	Statistician
Lisle, Liselott	Program Evaluation and Methodology	Actuary
Byers, Ronald	Resources, Community, and Economic Development	Evaluator
Jenkins, Barbara	Library Services	Library Technician
Johnson, Daniel L.	Program Planning	Assistant to the Director

Name	Regional Office	Title
Barnhill, James R.	Atlanta	Evaluator
Rucker, Colonel S., Jr.	Dallas	Evaluator
Smith, Royce D.	Denver	Evaluator
Henry, Egbert C.	Detroit	Evaluator
Schmidt, Donald R.	Detroit	Evaluator
Newton, Lee W.	Kansas City	Senior Evaluator
Deibel, Karl	Los Angeles	Supervisory Evaluator
Hedrick, Harry	Los Angeles	Evaluator
McRae, Joel	New York	Evaluator
Antelman, Victor	Philadelphia	Evaluator
McMonagle, Cornelius	Philadelphia	Evaluator

Retirements (cont.)

Name	Regional Office	Title
Rodgers, James	Philadelphia	Evaluator
Zampino, Samuel	Philadelphia	Evaluator
Niedzwiecki, Genevieve	San Francisco	Secretary

Death

Name	Regional Office	Title
Johnson, Albert F. (Sandy)	Washington	Evaluator

Professional Activities

Ed. note: GAO staff participated in the following professional activities during the approximate period January to March 1986.

Office of the Comptroller General

Charles A. Bowsher, Comptroller General, addressed the following groups:
Institute of Internal Auditors conference, Washington, Feb. 12.

National Association of Accountants, Washington, Mar. 4.

Air Transport Association, Washington, Mar. 13.

Joint Financial Management Improvement Program conference, Washington, Mar. 18.

International Economic Policy Association, Washington, Mar. 19.

National Association of State Comptrollers, Crystal City, VA, Mar. 20.

With **Richard L. Fogel, Director, Human Resources Division,** and Dennis Whitfield, Under Secretary, Department of Labor (DOL), met with the faculty of Harvard University's John F. Kennedy School of Government to discuss GAO's management reviews. Mr. Whitfield discussed DOL's response to GAO recommendations, Feb. 19.

Harry S. Havens, Assistant Comptroller General:
Spoke on "Gramm-Rudman-Hollings and the Battle of the Budget" during a roundtable discussion at American University, Washington, Feb. 24.

Addressed a seminar, "The Gramm-Rudman-Hollings Emergency Deficit Reduction Act of 1985," sponsored by George Washington University, Washington, Feb. 27.

Spoke on "The Functions of the General Accounting Office" during a week-long conference on understanding federal government operations held at the Brookings Institution, Washington, Mar. 3.

Spoke on "Managing the Cost of Government" at the first national budgeting seminar presented by the International Law Institute in cooperation with Georgetown University, Washington, Mar. 5

Participated in "The Future of Financial Management Under Gramm-Rudman," a Joint Financial Management Improvement Program workshop, Washington, Mar. 18.

Addressed the topic "Implementing Gramm-Rudman and Implications for Future Deficit Reduction" during a monthly luncheon of the National Capital Area Chapter of the American Society for Public Administration, Washington, Mar. 20.

Delivered the keynote speech on "Developments and Implementation of the Gramm-Rudman-Hollings Legislation" during a combined meeting of the Institute of Internal Auditors (Springfield Chapter), the Association of Government Accountants, and the State Internal Audit Managers, Springfield, IL, Mar. 27.

Participated in a panel discussion on "Reform '88 in the Midst of Gramm-Rudman" during an executive forum cosponsored by the Federal Executive Institute Alumni Association and the National Academy of Public Administration, Washington, Mar. 28.

Accounting and Financial Management Division

Frederick D. Wolf, Director:
Addressed the National State Auditors Executive Committee, Washington, Feb. 10.

Addressed the Institute of Internal Auditors at their Internal Auditing in Government Conference, Washington, Feb. 12.

Participated in a panel discussion at a National Academy of Public Administration workshop on streamlining the internal control review process, Washington, Mar. 5.

Spoke to participants of the Interagency Federal Training Systems Program at the Executive Seminar Center, Kings Point, NY, Mar. 11.

Addressed participants at the Joint Financial Management Improvement Program Annual Conference, Washington, Mar. 18.

Virginia B. Robinson, Associate Director:
Moderated a workshop on reexamining the role of central agencies sponsored by the Federal Executive Institute Alumni Association, Arlington, VA, Mar. 27.

Was elected for a 3-year term, 1986-88, to the Board of Governors, Washington Chapter, the Institute of Internal Auditors.

William A. Broadus, Group Director, spoke on the proposed revisions to the "yellow book" before the Southeastern Intergovernmental Audit Forum, Atlanta, Feb. 28, and before the Southwestern Intergovernmental Audit Forum, Dallas, Mar. 4.

Bruce Michelson, Group Director, spoke on the federal government reporting study (a joint GAO-Canadian Office of the Auditor General project) to a group at the International Monetary Fund, Washington, Mar. 19.

General Government Division

William J. Anderson, Director, spoke at the Federal Executive Institute on the "Role of the Federal Management Agencies," Arlington, VA, Mar. 27.

Johnny Finch, Senior Associate Director, discussed GAO's tax policy and administration work before the American Bar Association Section on Taxation, Washington, Jan. 3.

Rosslyn S. Kleeman, Senior Associate Director:
Spoke at an Office of Personnel Management (OPM) executive seminar, Kings Point, NY, Jan. 8.

Discussed "What's Ahead for Federal Employees" before the D.C. Chapter of the International Personnel Management Association, Feb. 5.

Discussed "Federal Management Issues in the 80s" at OPM's Executive Seminar Center, Kings Point, NY, Feb. 11.

Discussed GAO's role and responsibilities at a career workshop sponsored by Indiana University's Washington Semester Leadership Program, Mar. 17.

Participated in a panel discussion on "Alternative Pay, Benefits, and Retirement Plan: The Search for Balance and Flexibility in the Compensation Package" at the Federal Executive Institute Alumni Association's "Executive Days," Washington, Mar. 27.

Brian Usilaner, Associate Director:
Spoke on "Productivity Measurement Systems in Executive Agencies: The GAO Perspective" at the Office of Management and Budget, Feb. 27.

Discussed productivity audits and participated in a panel discussion on "Productivity Basics in a Fast-Changing World" before the American Productivity Management Association, Los Angeles, Feb. 11.

Natwar Gandhi, Group Director, discussed taxation of insurance companies before the plenary session of the Conning Investment Conference, New York, Jan. 21.

Larry Herrmann, Group Director, discussed executive agencies' audiovisual policies and practices before representatives of the motion picture and videotape industry associations, New York, Mar. 13.

Human Resources Division

Richard L. Fogel, Director, spoke on evaluating policy outcomes and the mangement of federal agenices at the Executive Development Seminar, Executive Seminar Center, Kings Point, NY, Mar. 13.

Bill Gadsby, Associate Director, spoke on GAO's reviews of the community services block grant at the 1986 annual conference of the National Community Action Foundation, Washington, Mar. 14.

Barry Tice, Group Director; Bob Wychulis, Senior Evaluator; and Cam Zola, Senior Evaluator, discussed GAO's reviews of social security disability programs before the board of the National Association of Disability Examiners, Arlington, VA, Mar. 7.

Paul Roberts, Senior Economist, spoke on abandoned coal mine lands before the National Research Council's Commission on Physical Sciences, Mathematics, and Resources, Washington, Feb. 14.

Chris Crissman, Evaluator:
Participated in an expert panel on a Department of Education study on guaranteed student load default prevention and collection techniques, Washington, Mar. 17.

Discussed GAO's review of guarantee agencies' efforts to collect defaulted student loans before the National Council of Higher Education Loan Programs' Annual Default Conference, San Antonio, TX, Mar. 26.

Susan Kladiva, Evaluator, spoke on GAO's review of medical malpractice issues before the Hospital Insurance Forum, Scottsdale, AZ, Mar. 10.

National Security and International Affairs Division

Frank C. Conahan, Director:
Participated in a panel discussion on "Trends Toward Nationalization of the Defense Industry" at the National Security Industrial Association, Washington, Feb. 17-19.

Participated in a teleconference on "National Security Policy" from Washington before students at the Executive Seminar Center, Oak Ridge, TN, Feb. 27.

Paul Math, Associate Director, discussed Department of Defense procurement issues before the local chapter of the Association of Government Accountants, Chicago, Mar. 17.

Phil Thomas, Group Director, discussed GAO's work on agricultural trade before the National Commission on Agricultural Trade and Export Policy, Washington, Feb. 14.

Julia Denman, Senior Evaluator, gave a luncheon presentation entitled "Optimizing Limited Defense Dollars: A Challenge for Logisticians," before the Andrews Air Force Base Chapter of the Society of Logistics Engineers, Mar. 5.

Eileen Larence, Evaluator, was elected to the National Council Panel, American Society for Public Administration (ASPA). She was also selected to chair ASPA's National Planning and Evaluation Committee for 1986.

Program Evaluation and Methodology Division

Jill Bernstein, Project Manager:
Coauthored a paper, "Data on Home Health Care Services in the United States: What We Know and What We Need to Find Out," published in the *International Journal of Technology Assessment in Health Care* (Winter 1986).

Coauthored an article, "Home Health Care in the Era of Hospital Perspective Payment," with **Susan Van Gelder, Assignment Manager, Human Resources Division.** The article was published in the *Pride Institute Journal of Long-term Home Health Care* (Winter 1986).

Carolyn Boyce, Evaluator, spoke on "Continued Long-term Gains Through Early Intervention With Technology: A 13-Year Study" at the Southern Education Research Association meeting, Houston, Jan. 30.

Patrick Grasso, Evaluator, coauthored "The Performance of States and Localities: An Overview," the lead article in *Public Policy Across States and Communities*, edited by Dennis R. Judd, published by JAI Press (Winter 1986).

Resources, Community, and Economic Development Division

Dexter Peach, Director; Keith Fultz, Associate Director; and Dwayne Weigel, Group Director, discussed GAO's reviews of nuclear energy programs before the Atomic Energy Industrial Forum, Washington, Mar. 24.

Mark Nadel, Group Director, discussed "The Role and Politics of the Congressional Support Agencies" before graduate students and faculty of the Johns Hopkins University Political Science Department, Baltimore, Mar. 13.

Office of the General Counsel

Harry R. Van Cleve, General Counsel:
Participated in a panel discussion on bid-protest procedures at a symposium on "Competing for Federal Contracts: Changes in Pre-Award Forums, Procedures, and Remedies" sponsored by the American Bar Association, Washington, Jan. 17.

Discussed GAO's role, actions, and bid-protest case examples at the 1986 conference on Comptroller General procurement decisions sponsored by Federal Publications, Inc., Washington, Jan. 29.

James F. Hinchman, Deputy General Counsel, discussed GAO's role in implementing the Balanced Budget and Emergency Deficit Control Act of 1985 (Gramm-Rudman-Hollings) at a 2-day training program on the legislation for Office of Management and Budget staff, Washington, Feb. 20-21.

Rollee H. Efros, Associate General Counsel, spoke before the Legal Education Institute, Department of Justice, on "Fiscal Control and the General Accounting Office," San Diego, Mar. 26.

Seymour Efros, Associate General Counsel, spoke before the Legal Education Institute, Department of Justice, on "Competition in Contracting Act Developments," San Diego, Mar. 26.

Ronald Berger, Assistant General Counsel:
Addressed a seminar on computer acquisition and GAO's bid-protest experience and activities under the Competition in Contracting Act, Springfield, VA, Jan. 27.

Participated in a seminar on the Competition in Contracting Act and subsequent legislation, Washington, Feb. 4, and Las Vegas, Feb. 18.

Spoke before the Public Contracts Law Section, American Bar Association, on GAO and its protest experience under the Competition in Contracting Act, Baltimore, Feb. 7.

Ronald Wartow, Group Managing Attorney, addressed the Forest Service National Contracting Officers Conference on "Bid Protests Under the Competition in Contracting Act," Salt Lake City, Mar. 17.

Bertram J. Berlin, Senior Attorney, spoke before the Federal Interagency Health and Fitness Council, Office of Personnel Management, on the Comptroller General's decision on the "Federal Employees Fitness Program," Washington, Mar. 13.

Paul Edmondson, Senior Attorney, discussed GAO's role under the Balanced Budget and Emergency Deficit Control Act of 1985 (Gramm-Rudman-Hollings) at a law and public policy symposium at Catholic University School of Law, Washington, Mar. 26.

Office of Internal Evaluation

John Butcher, Senior Evaluator, was elected president of the Washington Chapter of the National Association of Accountants for 1986-87.

Joint Financial Management Improvement Program

David Dukes, Executive Director:
Participated in a panel discussion on cash management improvements before the National Assistance Management Association, Washington, Mar. 20.

Spoke at the Association of Government Accountants' Emerging Issues Conference on the future of financial management under the Balanced Budget and Emergency Deficit Control Act of 1985 (Gramm-Rudman-Hollings), Cleveland, Mar. 25.

Doris Chew, Assistant Executive Director, moderated and spoke at an Association of Government Accountants, Washington Chapter, and Joint Financial Management Improvement Program workshop on "Using Microcomputers in Accounting and Budgeting Operations," Washington, Jan. 23.

Personnel

Patricia A. Moore, Task Group Director, was appointed to the Merit System Board of the Maryland National Park and Planning Commission, 1986-90.

Dinah Griggsby, College Relations Officer, discussed recruitment and retention issues at the Virginia College Placement Association meeting for counselors, Fredericksburg, VA, Feb. 19, and the State Council for Higher Education and American Affirmative Action Association Conference, Richmond, VA, Feb. 20.

Regional Offices

Boston

Paul M. Greenley, Senior Evaluator, conducted a seminar on the rules of evidence and the detection of fraud for auditors from several Massachusetts state agencies, Jan. 14.

Harriet C. Ganson, Evaluator, received a Ph.D. in sociology from Ohio State University, Columbus, Mar. 21.

Jennifer Arns, Technical Information Specialist, chaired the Government Publications Librarians of New England's spring conference on "Sources of Government Information for Public Policy Research," Tufts University, Medford, MA, Mar. 21.

Denver

Pamela K. Tumler and Diane Sanelli, Writer-Editors, conducted workshops on "Understanding the Writing Process: Planning, Drafting, and Revising" at the annual Professional Development Session for senior and junior staff of the North Dakota State Auditor's Office, Bismarck, ND, Jan. 9-10.

Kansas City

David A. Hanna, Regional Manager, spoke before the Omaha Chapter of the National Association of Accountants on "The GAO: What It Has Done Lately and What It Plans to Do Next," Omaha, Feb. 25.

The Kansas City Regional Office was one of four organizations that received special recognition for outstanding contributions to the University of Iowa's Cooperative Education Program in 1984-85, Iowa City, Mar. 31.

Susanne Valdez, Executive Director, Mid-America Intergovernmental Audit Forum:

Spoke to a public administration class at the University of Missouri on performing program results reviews, St. Louis, Feb. 3.

Spoke to the University of Missouri Accounting Club about GAO and the Mid-America Intergovernmental Audit Forum, St. Louis, Mar. 7.

Denise Millet, Evaluator, spoke to a public administration class at the University of Missouri on GAO's methodology for program results reviews, St. Louis, Feb. 3.

Los Angeles

Vic Ell, Assistant Regional Manager:

Spoke before the National Accountants Association, San Gabriel Valley Chapter, on "The Changing Accounting Profession—Evolution or Revolution?" Jan. 16.

Participated in the California Society of Certified Public Accountants' Ethics Subcommittee meeting to draft a new code of ethics, Jan. 28.

Spoke at the Annual Educational Conference of the Association of Government Accountants on "How Many Auditors Are Enough?" Los Angeles and Buena Park, Jan. 29-30.

Spoke to the Exchange Club of Pasadena on "The GAO: Its Structure, Responsibilities, and Accomplishments," Feb. 26.

Frederick Gallegos, Manager, Management Science Group:

Spoke before the Institute of Internal Auditors, Orange Empire Chapter, on "Microcomputers and Their Use in Auditing," Anaheim, Jan. 29.

Spoke before the spring conference of the EDP Auditors Association, Los Angeles Chapter, on "EDP Audit Career Development Planning," Mar. 19.

Taught a session of the certified information systems auditor review on "Data Integrity Reviews and Information Systems Audit Management" for the EDP Auditors Association, Los Angeles Chapter, Mar. 22.

Taught a graduate course on advanced EDP audit for the EDP audit program at California State Polytechnic University, winter quarter.

Coauthored an article with Lorne Dear of the Air Force Audit Agency, "Planning for the Security of Local Area Networks," which was published by Auerbach Publishers, Mar. 1986.

Philadelphia

Fred Layton, Regional Manager:

Addressed the Central Pennsylvania Chapter of the Association of Government Accountants on the role of GAO in the congressional oversight process and current areas of key interest to the Congress, Harrisburg, PA, Nov. 19.

Addressed the National Association of Black Accountants regarding GAO and its work for the Congress, Philadelphia, Feb. 12.

Addressed the National Accounting Fraternity of Temple University regarding GAO's role in congressional oversight, Philadelphia, Mar. 25.

Charles Hodges, Evaluator, was selected to serve on the Policy Development Committee of the Hempfield Area School District, Greensburg, PA, 1986.

Richard Halter, David Pasquarello, and John Sabia, Evaluators, discussed GAO's continuing role in federal and state development of income eligibility verification systems (IEVS) at a state conference on IEVS cosponsored by the Department of Health and Human Services' Office of Family Assistance and the Department of Agriculture's Food and Nutrition Service, Mt. Laurel, NJ, Mar. 20.

San Francisco

Jim Mansheim, Assistant Regional Manager:

Spoke to the American Society of Military Comptrollers, San Francisco Chapter, on controlling fraud and abuse, Emeryville, CA, Mar. 5.

Was appointed to the 1986 Steering Committee of the San Francisco Bay Area Federal Financial Managers Council.

Seattle

Stephen J. Jue, Technical Assistance Group Manager, taught classes on "Systems Development Life Cycle" and "Application Systems Development and Maintenance Reviews" at the EDP Auditors Association's certified information systems auditor review course, Seattle, Feb. 3 and 10.

Carla J. Revell, Staff Manager:

Participated in a panel discussion of careers in foreign affairs at the Graduate School of Public Affairs, University of Washington, Seattle, Jan. 17.

Spoke on "The Effects of the Gramm-Rudman Act" before the Cascade Chapter, Federally Employed Women, Seattle, Jan. 30.

Keith C. Martensen, Senior Evaluator, was elected to a 3-year term on the National Council of the American Society for Public Administration (ASPA), Feb. 21. He will represent ASPA's Region IX members in Washington, Oregon, Idaho, Montana, Hawaii, and Alaska. Mr. Martensen also serves as Membership Chairman and Council Member for the Evergreen Chapter in western Washington.

Brian A. Estes, Evaluator:

Spoke on "The Presidential Management Internship: A GAO Perspective" before students of the Graduate School of Public Affairs, University of Washington, Seattle, Dec. 12.

Participated in a panel dicussion of career opportunities in program evaluation and policy analysis before students of the Graduate School of Public Affairs, University of Washington, Seattle, Jan. 17.

Washington

Gloria Mayer, Assistant Regional Manager for Planning and Reporting, chaired a session on "Whither Wage Parity: 'Comparable Worth' Made Easy" at the Federal Executive Institute Alumni Association forum, Washington, Mar. 27.

Elizabeth Toiya Nyang, Technical Information Specialist, spoke on "Using a Micro to Do Research and Keep Control of the Job" at the second international conference of the Application of Microcomputers in Information, Documentation, and Libraries, Baden-Baden, West Germany, Mar. 17.

Karen Smithwick and Fred Doggett, Evaluators, discussed executive agencies' audiovisual policies and practices before a meeting of representatives of the motion picture and videotape industry associations, New York, Mar. 13.

Senior Staff (continued from page 36)

Resources Division Director's Award in 1983, and a Comptroller General's commendation in 1985 for his PAQRS efforts.

F. Henry Barclay, Jr.

Mr. F. Henry Barclay, Jr., retired from GAO in January 1986, ending a government career that spanned 43 years. During his career, Mr. Barclay was promoted from the position of claims examiner to that of Associate General Counsel. He worked 41 years in the Office of the General Counsel and served at the Associate General Counsel level in the fields of transportation law, general government matters law, and personnel law.

Mr. Barclay attended the University of Virginia and and received his L.L.B. from the University of Maryland Law School. He was admitted to the Maryland Bar in 1939 and subsequently became a member of the Federal and American Bar Associations.

Mr. Barclay received numerous awards at GAO, including the Comptroller General's Award, the Distinguished Service Award, and the General Counsel's Award. He also achieved the rank of Meritorious Executive in the Senior Executive Service.

Ed. note: A photograph of Mr. Barclay was unavailable.

Census (continued from page 33)

Bibliography

Procedural History 1970 Census of Population and Housing. U.S. Department of Commerce, 1976.

U.S. Congress. Senate. Committee on Governmental Affairs. *Enumeration of Undocumented Aliens in the Decennial Census.* Hearings before the Subcommittee on Energy, Nuclear Proliferation, and Government Processes, 99th Cong., 1st sess., 1985.

U.S. Congress. House. Committee on Post Office and Civil Service. *Plans and Activities for 1990 Decennial Census.* Hearings before the Subcommittee on Census and Population, 99th Cong., 1st sess., 1985.

Procedures to Adjust 1980 Census Counts Have Limitations (GAO/GGD-81-38, Dec. 24, 1980). ■

Awards for the Best Articles Published in *The GAO Review*

Cash awards of $500 each are presented each year (see GAO Order 1551.1) for the best two articles written by GAO staff and published originally in *The GAO Review.* Staff through grade GS-15 at the time they submit the article are eligible for these awards. A noncash award is available for best article by a member of the Senior Executive Service (SES) or candidate pool. The awards are presented during the GAO Awards Program held annually in Washington, D.C.

The awards are based on recommendations of a panel of judges that is independent of *The GAO Review* staff. The panel of judges is chaired by the Director, Office of Policy (OP), who, together with the Director, Office of Public Information, serves as a permanent panel member. Two other SES-level panel members will be selected for a 1-year term by the Director, OP. These selections will be made from among the members of GAO's Office-wide Awards Committee. The judges evaluate articles from the standpoint of their overall excellence, with particular concern for the following:

- Originality of concepts and ideas. (The authors demonstrated imagination and innovation in selecting and developing a topic.)
- Degree of interest to readers. (The article, by virtue of the topic and its treatment or its relevance to GAO's mission, was of special interest to GAO staff.)
- Quality and effectiveness of written expression. (The article was well organized and written in polished prose.)
- Evidence of individual effort expended.

Statement of Editorial Policy

This publication is prepared primarily for use by the staff of the General Accounting Office and outside readers interested in GAO's work. Except where otherwise indicated, the articles and other submissions generally express the views of the authors and do not represent an official position of the General Accounting Office.

The GAO Review's mission is threefold. First, it highlights GAO's work from the perspectives of subject area and methodology. (The *Review* usually publishes inherently interesting or controversial articles on subjects generated by GAO audit work and articles related to innovative audit techniques.) Second, and equally important, the *Review* provides GAO staff with a creative outlet for professional enhancement. Third, it acts as historian for significant audit trends, GAO events, and staff activities.

Potential authors and interested readers should refer to GAO Order 1551.1 for details on *Review* policies, procedures, and formats.

Documents published by the General Accounting Office can be ordered from GAO Document Distribution (202) 275-6241.

The GAO Review

U.S. General Accounting Office
Washington, D.C. 20548

Second-Class Mail
Postage and Fees Paid
U.S. General Accounting Office
0016-3414

Designing Evaluations

Methodology
Transfer Paper 4

GAO
Program Evaluation And Methodology Division
United States General Accounting Office
July 1984

Designing Evaluations

Methodology Transfer Paper 4

Program Evaluation and Methodology Division

July 1984

PREFACE

This paper addresses the logic of program evaluation designs. It provides a systematic approach to designing evaluations that takes into account the questions guiding a study, the constraints evaluators face in conducting it, and the information needs of its intended user. Taking the time to design evaluations carefully is a critical step toward insuring overall job quality. Indeed, the most important outcome of a careful, sound design should be an evaluation whose quality is high in quite specific ways.

Evaluation designs are characterized by the manner in which the evaluators have

--defined and posed the evaluation questions for study,

--developed a methodological approach for answering those questions,

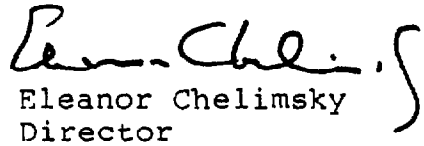
--formulated a data collection plan that anticipates problems, and

--detailed an analysis plan for answering the study questions with appropriate data.

Designing Evaluations is a guide to the successful completion of these design tasks.

Designing Evaluations also provides a detailed discussion of three kinds of evaluation questions--descriptive, normative, and causal--and various methodological approaches appropriate to each one. For illustration, the paper contains a narration of a recent design undertaken by the Program Evaluation and Methodology Division (PEMD) in response to a congressional request. To aid the understanding of the concepts in this paper, a workbook is being developed that will feature examples of the different design problems identified here.

Designing Evaluations is one of a series of papers issued by PEMD. The purpose of the series is to provide GAO evaluators with handy, clear, and comprehensive guides to various aspects of evaluation methodology, to explain specific applications and procedures, and to indicate where more detailed information is available. Other papers in the series include Causal Analysis and Content Analysis. Readers of Designing Evaluations are encouraged to send questions or comments about the contents of this paper to its authors--Ray Rist and Carl Wisler, both of PEMD.



Eleanor Chelimsky
Director

GLOSSARY

Bias
The extent to which a measurement or an analytic method systematically underestimates or overestimates a value.

Construct
An attribute, usually unobservable, such as educational attainment or socioeconomic status, that is represented by an observable measure.

Construct validity
The extent to which a measurement method accurately represents a construct and produces an observation distinct from that produced by a measure of another construct.

Covariation
The degree to which two measurements vary together.

Cross-sectional data
Observations collected on subjects or events at a single point in time.

External validity
The extent to which a finding applies (or can be generalized) to persons, objects, settings, or times other than those that were the subject of study.

Generalizability
Used interchangeably with "external validity."

Internal validity
The extent to which the causes of an effect are established by an inquiry.

Longitudinal data
Sometimes called "time series data"; observations collected over a period of time; the sample may or may not be the same each time but the population remains constant.

Measurement
A procedure for assigning a number to an object or an event.

Panel data
A special form of longitudinal data in which observations are collected on the same sample of respondents over a period of time.

Probability sampling
A method for drawing a sample from a population such that all possible samples have a known and specified probability of being drawn.

Program effectiveness evaluation
The application of scientific research methods to estimate how much observed results, intended or not, are caused by program activities. Effect is linked to cause by design and analysis that compare observed results with estimates of what might have been observed in the absence of the program.

Program evaluation
The application of scientific research methods to assess program concepts, implementation, and effectiveness.

Random assignment
A method for assigning subjects to two or more groups by chance.

Reliability
The extent to which a measurement can be expected to produce similar results on repeated observations of the same condition or event.

Representative sample
A sample that has approximately the same distribution of characteristics as the population from which it was drawn.

Simple random sampling
A method for drawing a sample from a population such that all samples of a given size have equal probability of being drawn.

Statistical conclusion validity
The extent to which the observed statistical significance (or the lack of statistical significance) of the covariation between two or more variables is based on a valid statistical test of that covariation.

Treatment group
The subjects of the intervention being studied.

CHAPTER 1

WHY SPEND TIME ON DESIGN?

According to a Chinese adage, even a thousand-mile journey must begin with the first step. The likelihood of reaching one's destination is much enhanced if the first step and the subsequent steps take the traveler in the correct direction. Wandering about here and there without a clear sense of purpose or direction consumes time, energy, and resources. It also diminishes the possibility that one will ever arrive. One can be much more prepared for a journey by collecting the necessary maps, studying alternative routes, and making informed estimates of the time, costs, and hazards one is likely to confront.

It is no less true that front-end planning is necessary to designing and implementing an evaluation successfully. Systematic attention to evaluation design is a safequard against using time and resources ineffectively. It is also a safeguard against performing an evaluation of poor quality and limited usefulness.

The goal of the evaluation design process is, of course, to produce a design for a particular evaluation. But what exactly is an evaluation design? Because there may be different views about the answer to this question, it is well to state what is understood in this paper. Evaluation pertains to the systematic examination of events or conditions that have (or are presumed to have) occurred at an earlier time or that are unfolding as the evaluation takes place. But to be examined, these events or conditions must exist, must be describable, must have occurred or be occurring. Evaluation is, thus, retrospective in that the emphasis is on what has been or is being observed, not on what is likely to happen (as in forecasting).[1] The designs and the design process outlined in this paper are focused on the observed performance of completed or ongoing programs.

To further characterize evaluation design, it is useful to look closely at the questions we pose and the answers we seek. Evaluation questions can be divided into three kinds: descriptive questions, normative questions, and cause-and-effect questions. The answers to descriptive questions provide, as the name implies, descriptive information about specific conditions or events--the number of people who receive Medicaid benefits in 1980, the construction cost of a nuclear power plant, and so on. The answers to normative questions (which unlike descriptive questions ask what should be rather than what is) compare

---

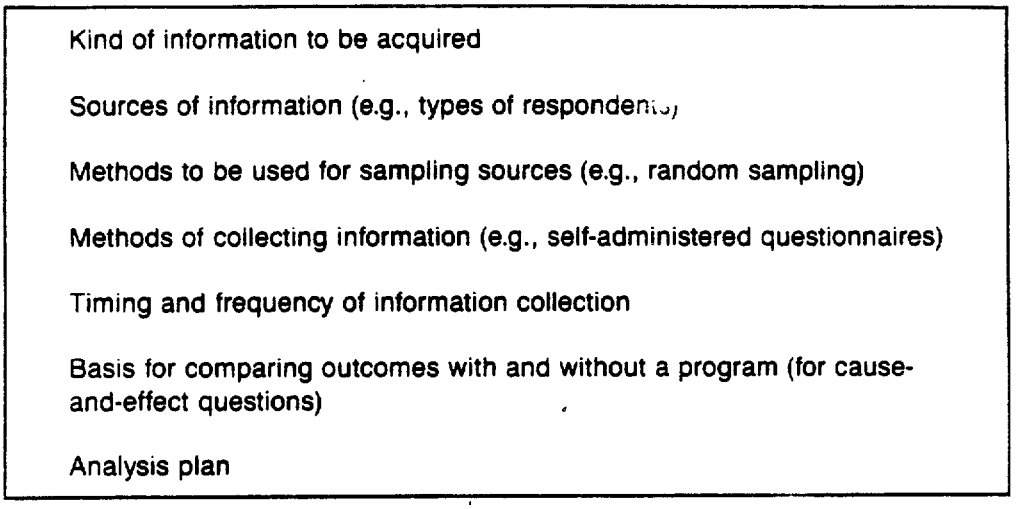
[1]Despite the retrospective character of evaluation, program evaluation findings can often be used as a sound basis for calculating future costs or projecting the likely effects of a program.

an observed outcome to an expected level of performance. An example is the comparison between airline safety violations and the standard that has been set for them. The answers to cause-and-effect questions help reveal whether observed conditions or events can be attributed to program operations. For example, if we observe changes in the weight of newborns, what part of those changes is the effect of a federal nutrition program? In sum, the design ideas presented here are aimed at producing answers to descriptive, normative, and cause-and-effect questions.

Given these questions, what elements of a design should be specified before information is collected? The most important elements are shown in figure 1. Taken together, these elements

Figure 1

Elements of Evaluation Design



form the basis on which a design is constructed. As will be seen, the choices that are made for each element are major determinants of the quality of the information that can be acquired, the strength of the conclusion that can be drawn, and the evaluation's cost, timeliness, and usefulness.

Before each component in this design process is identified and discussed, it would be well to address systematically why it is important to take the time to be concerned with job design. First, and probably most importantly, careful, sound design enhances quality. But it is also likely to contain costs and insure the timeliness of the findings, especially when the evaluation questions are difficult and complex. Further, good design increases the power and specificity of findings and recommenda-

tions, decreases vulnerability to methodological criticism, and improves customer satisfaction.

In thinking about these reasons for taking time to design an evaluation carefully, one may well find that guaranteeing job quality is the preeminent concern, the critical dimension of the design effort. Stated differently, the most important outcome of a careful, sound design should be that the overall quality of the job is enhanced in a number of specific ways.

An evaluation design can usually be recognized by the way it has

1. defined and posed the evaluation questions for study,

2. developed the methodological strategies for answering these questions,

3. formulated a data collection plan that anticipates and addresses the problems and obstacles that are likely to be encountered, and

4. detailed an analysis plan that will insure that the questions that are posed are answered with the appropriate data in the best possible fashion.

A well-designed evaluation will be more powerful and germane than one in which attention has not been paid to laying out the methodological strategy and planning the data collection and analysis carefully. It will also develop a stronger foundation and be more convincing in its conclusions and recommendations. Implementation also will be strengthened, because once the design has been established, less time will be lost in having to make ad hoc decisions about what to do next. Good front-end planning can substantially reduce the many uncertainties of a job. It helps provide a clear sense of direction and purpose to the effort.

Similarly, good front-end planning contains job costs by preventing (1) duplication of data collection, (2) unplanned data analysis in a search for relevant findings, (3) staff time being wasted on the collection and analysis of data that are irrelevant to the question, and (4) "down time" from making sporadic and episodic decisions on what to do next. It must be recognized that careful attention to design does take time and does necessitate front-end costs. However, the investment can save time and costs later in the job, and this is especially true for big, complex jobs. There is, of course, no assurance that careful work will require less expenditure of resources than ill-defined studies.

Attention to the design process also makes for high quality by focusing on the usefulness of the product to the intended

recipient. If attention is paid to the needs of the user in terms of information or recommendations, the design process can systematically address these needs and make sure that they are integrated into the job. In this way, the relevance of a job can be strengthened by tying it specifically to the concerns of its user. In addition, a concern with relevance is likely to increase the user's satisfaction with the product.

A sound design can help insure timeliness. A tight and logical design can reduce the time that accumulates on a job because of excessive or unnecessary data collection, the lack of a clear data analysis plan, or the constant "cooking" of the data as when the omission of a sound methodological strategy has made it impossible to answer the evaluation questions directly. The timeliness of findings with respect to the needs of the customer can make or break a technically adequate approach. It is not enough that a study be conducted with a high degree of technical precision to argue for its quality; the study must also be conducted in time to allow the findings to be of service to the user.

In summary, to spend the time to develop a sound design is to invest time in building high quality into the effort. Devoting attention to job design means that a number of considerations regarding job quality can be addressed and adequately met. Not allowing the time that is necessary for this vital stage of the job is, in the end, self-defeating. It can be a crippling, if not a fatal, blow to any job that skips quickly through this step. The pressure of wanting to get into the field as soon as possible has to be held in check while systematic planning takes place. The design is what guides the data collection and analysis.

Having looked at why it is important to design jobs well, we can turn our attention to the various components and processes that are inherent in job design. Our discussion is in five major parts: asking the right question, adequately considering the constraints, assessing the design, settling on a strategy that considers strengths and weaknesses, and rigorously monitoring the design and incorporating it into the management strategies of the persons who are responsible for the job.

CHAPTER 2

THE DESIGN PROCESS

ASKING THE RIGHT QUESTION

The first and surely the most fundamental aspect of every design effort is to insure that the questions that are posed for the job are the correct ones.[1] Posing a question incorrectly is an excellent way to lead a job in the wrong direction. It is obvious that one must ask the right question, but deciding what is exactly the "right question" is not necessarily easy. In fact, reaching agreement with the sponsors, users, program operators, and others on the contents and implications of a question can be difficult and challenging. Among the several reasons for the strenuousness of the task is that the formulation of a problem has preeminent importance in the remaining phases of the job. How a problem is stated has implications for the kinds of data to be collected, the sources of data, the analyses that will be necessary in trying to answer the question, and the conclusions that will be drawn.

Consider a brief example: juvenile delinquency and the question of what motivates young people to commit delinquent acts. The question about motivation could be posed in a variety of ways. One could ask about the personality traits of young persons and whether particular traits are associated with differences in who does or does not commit crimes. Asking the question this way entails data, data sources, and program initiatives that are different from those that are required in examining, for example, the social conditions of young persons; here, the focus might be on family life, schooling, peer groups, employment opportunities, or the like. To stretch the example further, each of these two ways of posing the question about what motivates juveniles to commit crime would lead to jobs quite different from either a job asking whether juveniles commit crimes because of a temporary hormonal imbalance or a job asking whether a youth culture uses crime as a "rite of passage" into adulthood.

Posing a question in four quite different ways shows clearly how the way in which a problem is stated has implications for an evaluation design. How an issue is defined influences directly how variables or dimensions are to be selected and examined and how the analysis will test the strength of the relationship between a cause and its expected consequence.

Question formulation is important also in that the concerns of the customer must be attended to. How a question is framed

---

[1] Often studies have more than one key question or a cluster of questions. Every question has to be given the same serious attention.

has to take the information needs and spheres of influence of the intended audience into consideration. Does the customer need to know the general effectiveness of a nationwide program? Or is the concern limited, for example, to individual problem sites or public attitudes to the program in those sites? The difference of type in these two questions is extremely important for evaluation design, and attention to the difference allows the evaluator to help make the job useful to its sponsor.

Clarifying the issue

Working toward the formulation of the right question has two phases (Cronbach, 1982, pp. 210-44).[2] In the first phase, the largest number and widest range of potential questions (and methods by which to address these questions) ought to be considered, even if they do not seem especially plausible or defensible. For example, congressional staff often begin with a very broad concern, so that it is necessary to try out a number of less sweeping questions in order to determine the priorities of the staff and to develop researchable questions. Thus, it is often useful for the evaluator and requestor to work through in detail which questions can be answered easily, which are more difficult, expensive, and time-consuming, and which cannot be answered at all and why. The evaluator is in a much stronger position to defend the final phrasing of a question if it is apparent that a number of alternatives have been systematically considered and rejected.

During this phase, the evaluator has several important aids for developing a range of questions. One is to imagine the various stages of the program--its goals, objectives, start-up procedures, implementation processes, anticipated outcomes--and to ask all the questions that could be asked about each stage. For example, in considering program objectives, the evaluator could ask questions about the clarity and precision of those objectives, the criteria that have been developed for testing whether the objectives have been met, the relationship between the objectives and program goals, and whether the objectives have been clearly transmitted to and understood by the persons who are responsible for the program's implementation.

Another aid is to focus on the nature of the program's objectives--on whether they are short term or long term, intense or weak, continuous or sporadic, behavioral or attitudinal, and so on. Yet another aid is to think of questions that would describe the program as it exists or that would judge the program against an existing norm or that would demonstrate which outcomes are a direct result of the program.

---

[2]Abbreviated bibliographic citations are expanded in the appendix on pages 53-54.

Each of these three kinds of question, which we discuss in chapter 3, necessitates a different design consideration. What is important for the evaluator is to separate a potential question into one of the three types and then to consider the implications of each type of question for the development of a design. To choose a set of evaluation questions is to choose a certain cluster of design options for answering them.

The second phase of formulating the right question is to match possible questions against the resources that will be available for the job. We discuss this in the following section.

Deciding which questions are feasible to answer

It is one thing to agree on which questions are most important and have highest priority. It is quite another to know whether the questions are answerable and, if so, at what costs in money, staff, and time. In the second phase of formulating the right question, the evaluator ought not to assume that a design developed to answer questions of highest priority can be implemented within the given constraints.

For example, the evaluator might determine that it would be very informative to collect data over several years, but the requirements of money, staff, and time might necessitate a less comprehensive or less complex design that could answer fewer questions, less conclusively, within given constraints. An alternative design that might be appropriate could focus on what a particular group of people remembers about a program or service during the years in which they were involved with it. Here, in place of the long-term, objective monitoring of events during years to come, the evaluator would substitute a look backward that is dependent on the memory and attitudes of the people involved with the program in the past.

Another less comprehensive alternative, of lower quality, would be to inquire of the group at only two future points in time rather than to make numerous inquiries over several points in time. In other words, the design option can influence the technical quality of the evidence and, hence, the expectations about what the evaluation can accomplish.

Meeting an information need reasonably

A large-scale and expensive evaluation is not likely to seem reasonable for a program that is small, diffuse, and short in duration. Similarly, a study that will allow national projections will probably require effort and resources quite different from those of a narrower study. To make national projections from a single case study, for example, is difficult, if not impossible. That is, whether or not an information need

can be reasonably met has to do with how conclusive the answer to the question being investigated has to be.

Questions that call for a high degree of conclusiveness in the answers will, of necessity, require stronger designs than questions for which brief descriptions or quick assessments are adequate answers. For example, to ask for a description of the children who receive services from an education program for migrants is quite different from asking whether those services are affecting their attendance in school, academic achievement, and proficiency in English. The first question could be answered descriptively with the collection and tabulation of demographic data, but the second is a cause-and-effect question that demands knowledge about, first, what is happening to similar children who are not in the program and, second, how the children who are in the program were performing before they joined it and, third, whether other possible causes for how the children are performing that have nothing to do with the program can be justifiably excluded.

The "strength versus weakness" issue

Strong evaluations employ methods of analysis that are appropriate to the question, support the answer with evidence, document the assumptions, procedures, and modes of analysis, and rule out the competing evidence. Strong studies pose questions clearly, address them appropriately, and draw inferences commensurate with the power of the design and the availability, validity, and reliability of the data. Strength should not be equated with complexity. Nor should strength be equated with the degree of statistical manipulation of data. Neither infatuation with complexity nor statistical incantation makes an evaluation stronger.

The strength of an evaluation is not defined by a particular method. Longitudinal, experimental, quasi-experimental, before-and-after, and case study evaluations can be either strong or weak. A case study design will always be weaker than a true experimental design in terms of its external validity. A simple before-and-after design without controls will always present problems of internal validity. Yet true experiments and longitudinal studies can be impossible for a variety of reasons. That is, the strength of an evaluation has to be judged within the context of the question, the time and cost constraints, the design, the technical adequacy of the data collection and analysis, and the presentation of the findings. A strong study is technically adequate and useful--in short, it is high in quality (Chelimsky, 1983).

Evaluators have considered the concept of strength at some length. Some argue that strong evaluations employ methods that allow the evaluator to make causal, as opposed to correlational, statements about a policy or program. It is argued that saying

that program intervention X caused outcome Y among the program's participants is a stronger statement than saying that X and Y are associated but it is not clear that X caused Y. In this argument, the notion of strength is related to the judgment that causal statements are more powerful than correlational statements.

Another argument is that the strength of a study or a method can be decided by comparing what was done with what was possible. An evaluation that stretches the data, modes of analysis, and opportunities for use to the limits should be judged strong even though what might have been a stronger design may not have been feasible.

Pilot versus full study

Formulating the right question is a necessary but not a sufficient condition for success. There is still the matter of translating the design and analytic assumptions into practice--into pragmatic decisions and patterns of implementation that will allow the evaluator to find the stipulated data and analyze them. In short, the evaluator must ask whether the design matches the area of inquiry. Answering this question is a "reality check" on whether the assumptions about the kinds and availability of data hold true, on whether the legislative and regulatory descriptions of the program bear any resemblance to what has been implemented, and on whether the proposed analysis strategies will answer the question conclusively.

At this stage of a job, the entire endeavor is still quite vulnerable and tentative. What if the data are not available? What if the program is nothing like its description in its documents or the grant application? What if the methodology will not allow for sufficiently conclusive answers to the evaluation questions? Any one of these situations could call an entire job into question.

That the condition of a job can be precarious in these ways argues for a limited exploration of the question before a full-scale, perhaps expensive, job is undertaken. This limited exploration is referred to as a "pilot phase," when the initial assumptions about the program, data, and evaluation methodology can be tested in the field. Testing the work at one or more sites allows the evaluator to confirm that data are available, what their form will be, and by what means they can be gathered.

Site selection for the pilot phase is important. Rather than choosing a site where the pilot could be easily conducted, it is critical to choose a site that represents an average, if not the worst, case. Choosing a noncontroversial site may hide the resistance an evaluator is likely to experience at other sites.

The pilot phase allows for a check on program operations and delivery of services in order to ascertain whether what is assumed to exist does. Finding that it does not may suggest a need to refocus the question to ask why the program that has been implemented is so different from what was proposed. This phase allows also for limited data collection, which provides an opportunity to assess whether the analysis methodology will be appropriate and what alternative interpretations of the data may be possible.

The study's pilot phase is very useful. It is an important opportunity to correct aspects of the design that can determine the success or the failure of the overall effort. To undertake a large-scale, full-blown study without this phase is a high-risk proposition. To allocate staff and financial resources and engage the time and cooperation of the persons in the programs to be studied without making as certain as possible that what is proposed will work is to court serious problems. It may well be that conducting a pilot will confirm what was originally designed, but to move ahead with this confirmation is preferable to merely assuming that everything will fall successfully into place.

To be sure, there are instances when a pilot is not possible: time pressures may not allow it, resources may be so scarce that there is but one opportunity for field work, or the availability of staff may be constrained. Yet the evaluator ought to recognize that not performing a pilot test increases the likelihood of problems and difficulties, even to the degree that the study cannot be completed successfully. The evaluator must give high priority to the pilot phase when considering time, resources, and staff.

A frequently posed question is how much pilot work is necessary before the large-scale evaluation is undertaken. There is no "cookbook" answer. The pilot is an evaluation tool that increases the odds that the effort will be high in quality. By itself, the pilot cannot provide a fail-safe guarantee. It can suggest alternative data collection and analysis strategies. It can also stimulate further thinking about and clarification of the job. The pilot is a strategy for reducing uncertainty. That uncertainty cannot be reduced to zero does not detract from the pilot's utility.

Perhaps the best answer to how extensive a pilot ought to be is a second question: How much uncertainty is the evaluator willing to tolerate as the evaluation begins? Only the evaluator can make the trade-off between the scope and resources of the pilot and problems on the job.

CONSIDERING THE EVALUATION'S CONSTRAINTS

Time is a constraint. It shapes the scope of the evaluation question and the range of activities that can be undertaken to answer it. It demands trade-offs and establishes

boundaries to what can be accomplished. It continually forces the evaluator to think in terms of what can be done versus what might be desirable. Because time is finite (and there is never enough of it), the evaluator has to plan the job in "real time" with its inevitable constraints on what question can be posed, what data can be collected, and what analysis can be undertaken.

A rule of thumb is that the time for a job and the scope of the question being addressed ought to be directly related. Tightly structured and narrow investigations are more appropriate when time is short. Any increase in the scope of a study should be accompanied by a commensurate increase in the amount of time that is available for it. The failure to recognize and plan for this link between time and scope is the Achilles heel of evaluation.

Linking scope and time in the study design is important because the scope is determined by the difficulty of the job, the importance of the subject, and the needs of the user and these are also determinants of time. Though it may be self-evident to say so, difficult jobs, important jobs, and jobs in which there is a great deal of interest will have different demands with respect to time than other jobs. No job is "too long" or "too short" within this context.

The need of the study's audience as a time constraint merits additional comment. Evaluations are requested and conducted because someone perceives a need for information. Producing that information without a sensitivity to the user's timetable diminishes its usefulness. For example, a report to the Congress may answer the questions correctly but will be of little or no use if it is delivered after the legislative hearings for which it is needed or after the preparation of a new budget for the program.

Cost is a constraint. The financial resources available for conducting a study partly determine the limits of the study. Having very few resources means that the evaluator will have to consider tight limitations on the questions, the modes of data collection, the numbers of sites and respondents, and the extent and elegance of the analysis. As the resources expand, the constraints on the study become less confining. Having more funds might mean, for example, either longer time in the field or the opportunity to have multiple interviews with respondents or to visit more sites or choose larger samples for sites. Each of these items has a price tag. What the evaluator is able to purchase depends on what funds are available.

It should be stressed that regardless of what funds are available, design alternatives should be considered. Cost is simply an important constraint within which the design work has to proceed. If only a stipulated sum is available, the evaluator has to determine what can be done with that sum in order to provide information that is relevant to the questions. The same

resources might allow three or four quite distinct approaches to a job. The challenge is to consider the strengths and weaknesses of the various approaches. Like the constraint of time, cost does not determine the design. It helps establish the range of options that can be realistically examined.

Even when resources can be expanded, cost is still a constraint. However, the design problem then becomes one of cost-effectiveness, or getting value for the dollar, rather than one of what can be done within a stipulated sum.

One other point: the quality of an evaluation does not depend on its cost. A $500,000 evaluation is not necessarily five times more worthy than a $100,000 evaluation. An expensive study poorly designed and executed is, in the end, worth less than one that costs less but addresses a significant question, is tightly reasoned, and is carefully executed. A study should be costly only when the questions and the means of answering them necessitate a large expenditure. As with the constraint of time, there is a direct correlation between the scope of a study and the money available for conducting it.

Staff expertise is a constraint. The design for an evaluation ought not to be more intricate or complex than what the staff can successfully execute. Developing highly sophisticated computer simulations or econometric models as part of an evaluation when the skills for using them are not available to the evaluation team is simply a gross mismatch of resources. The skills of the staff have to be taken into account when the design is developed.

It is perhaps too negative to consider staff expertise as only a constraint. In the alternative view, the design accounts for the range of available staff expertise and plans a study that uses that expertise to the maximum. It is just as much a mismatch to plan a design that is pedantic, low in power, and completely unsophisticated when the staff are capable of much more and the questions demand more as it is to create a design that is too complex for the expertise available. In either instance, of course, a design is determined not by expertise but by the nature of the questions.

A realistic understanding of the skills of the staff can play an important role in the kinds of design options that can be considered. An option that requires skills that the staff do not have will fail, no matter how appropriate the option may be to the evaluation questions. A staff with a high degree of technical training in a variety of evaluation strategies is a tremendous asset and greatly expands the options.

Some designs demand a level of expertise that is not available. When this happens, consultants can be brought into the study or the staff can be given short intensive courses or complex and difficult portions of the design can be isolated and

performed under contract by evaluators specializing in the appropriate type of study. In other words, the stress is on considering the options available. Preference should be given to building the capability of current staff. When this cannot be done, or time and cost do not allow it, expertise can be procured from outside in order to fulfill the demands of the design.

Location and facilities are secondary constraints in comparison to the others we have discussed, but they do impinge on the design process and influence the options. Location has to be considered from several aspects. One is the location of the evaluator vis-a-vis where the evaluation is to be conducted. Location is less critical for a national study, since most areas can be reached by air within a few hours, but it increases in importance if the study examines only a few individual projects. The accessibility and continuity of data collection may be jeopardized if the evaluator is on the east coast and the sites are in the South, in the Midwest, and on the west coast. A situation such as this may have to incorporate local persons as members of the evaluation team and may increase the utility of a mail questionnaire or telephone interviews compared to face-to-face interviews.

Another aspect of location has to do with the social and cultural mores of the area where the evaluation is to be conducted. For example, to gain valid and insightful data on attitudes toward rural mental health clinics, it may be wise not to send interviewers from urban areas. Good interviewing necessitates empathy between the persons involved, and it may be hard to generate between an interviewer and a respondent whose backgrounds are very different.

A third aspect of location is the stability of the population being studied. A neighborhood where residence is transient may necessitate a different strategy from a neighborhood where most people have lived in the same house for 40 years and have no intention of moving.

Finally, the evaluator must consider whether a trip to a site is justified at all. For example, if it costs $3,000 to travel to a remote town to ascertain whether a school there is using a $1,500-computer provided by a U.S. Department of Education grant, the choice of not going is defensible.

The constraint of facilities on the design options also has more than one aspect. One has to do with data collection and data processing. For example, if the study involves entering large aggregates of data into a computer, the equipment to do so must be available, or the money must be available for contracting the work. Similarly, if the design calls for data analysis at computer terminals with phone connections to the main computer, the equipment is a must. The absence of such

facilities limits both the kind and the extent of the data one can collect.

Another aspect is the need for periodic access to facilities that are not under the auspices of the project or program being studied. For example, to interview welfare clients in a welfare office about the treatment and service they are receiving there may be to risk highly biased answers. How candid can a client be, knowing that the caseworker who has made decisions on food, clothing, and rental allowances for the client's family is in the next room? "Neutral turf" cannot guarantee candid answers, but it may lessen anxiety and it can contribute to the authenticity of the evaluator's promise of anonymity and confidentiality. The example applies equally to interviews with persons who hold positions of power and influence.

ASSESSING THE DESIGN

Once a design has been selected, the impetus is to move full steam ahead into the execution of the study. However, the evaluator must fight this impulse and take time to look back on what has been accomplished, on what design has finally been selected, and on what the implications are for the subsequent phases of the study. The end of the design phase is an important milestone. It is here that the evaluator must have a clear understanding of what has been chosen, what has been omitted, what strengths and weaknesses have been embedded in the design, what the needs of the customer are, how usefully the design is likely to meet those needs, and whether the constraints of time, cost, staff, location, and facilities have been fully and adequately addressed.

Within GAO's Program Evaluation and Methodology Division, the director has developed and uses a job review system that includes a detailed and systematic assessment of the design phase. This system helps establish the basis for moving forward into implementation. It may be useful to other evaluators in judging their own designs. Five key questions figure prominently in the review system.

1. How appropriate is the design for answering the questions posed for the study? The evaluator ought to be able to match the design components systematically to the study questions in order to demonstrate that all key questions are being addressed and that methods are available for doing so. Even though this entails a judgment, the evaluator should assess the match between the strength of the design and the information necessary to answer the study questions. If the design is either too weak or too strong for the questions, serious consideration has to be given to whether the design ought to be implemented or whether the questions ought to be modified. This judgment about the appropriateness of the design is critical, because if the study begins with an inappropriate design, it is difficult to compensate later for the basic incongruity.

2. How adequate is the design for answering the questions posed for the study? The emphasis here is on the completeness of the design, the expected precision of the answers, the tightness of the logic, the thought given to the limitations of the design, and the implications for the analysis of the data. First, the evaluator should have reviewed the literature and give evidence of knowing what was undertaken previously in the area from both substantive and methodological viewpoints. That is, the evaluator should be aware of not only what kinds of questions have been asked and answered in the past but also what designs, measures, and data analysis strategies have been used. A careful study of the literature prevents "rediscovering" or duplicating existing work. Thus, in judging the adequacy of the design, the evaluator must link it to previous evaluations.

Second, the design should explicitly state what evaluation questions determined the selection of the design. Knowing which evaluation questions were thought germane and which were not gives the reader a basis for assessing the strength of the design. Since every evaluation design is constrained by a number of factors, recognizing them and candidly describing their effect provides important clues to whether the design can adequately answer the study questions.

Third, there is a need to be explicit about the limitations of the study. How conclusive is the study likely to be, given the design? How detailed are the data collection and data analysis plans? What trade-offs were made in developing these plans? The answers to these questions provide data on the design's adequacy.

3. How feasible is the execution of the design within the required time and proposed resources? Adequate and appropriate designs may not be feasible if they ignore time and cost--that is, if they are not practical. The completeness and elegance of a design can be quickly relegated to secondary importance if the design presents major obstacles in the execution. Further, asking about feasibility puts an important check on studies that simply cannot be done. For example, discovering that a particular evaluation with a true experimental design cannot be executed may prevent starting up a job that will fail.

4. How appropriate is the design with regard to the user's needs for information, conclusiveness, and timeliness? What kind of information is needed? How conclusive does it have to be? When does it have to be delivered? Being able to determine how well the design responds to the user's needs requires the evaluator and the user to be in close agreement and continuous consultation. In the absence of cooperation, the evaluator is left to presume what will be of relevance--and presumption is a poor substitute for knowledge. Since evaluations are undertaken because of a need for information, the degree to which they provide useful information is an inescapable and critical design consideration.

5. How adequate were the negotiations with the user regarding the relationship between the information need and the study design? It is one thing to know what the user needs and when it is needed. It is quite another to agree on how the questions ought to be framed so that the information can be gathered. If the user has causal questions in mind while the evaluator believes that only a descriptive study is feasible, and if the gap between these two perspectives is not resolved, the user's satisfaction with the final study is likely to be quite low and the ensuing report may not be used.

Further, the consideration of time is relevant to the size, complexity, and completeness of the evaluation that is finally undertaken. If the user is integrally involved in determining the project's timetables and products, the evaluator will know how to decide whether what is proposed can be accomplished. To ignore, or only guess at, rather than negotiate and agree on a timetable would be to risk the relevance of the whole effort. The negotiations with the user should be carefully scrutinized at the end of the design phase to make sure that there is common understanding and agreement on what is being proposed for the remaining phases of the evaluation.

CHAPTER 3

TYPES OF DESIGN

In chapter 2, we examined the factors to consider in arriving at an evaluation design. Here we take a systematic look at four major evaluation strategies and several types of design that derive from them (table 1). The discussion is brief and nontechnical. More details can be found in the references given under the heading "Where to look for more information" for each design type.

Evaluation strategies and designs can be classified in a variety of ways, each with some advantages and disadvantages in communicating a logical picture of the different forms of evaluation inquiry. We take the word "strategy," as the broader of the two concepts, to connote a general approach to finding answers to evaluative questions. A strategy embraces several types of design that have certain features in common.

Our classification scheme is similar to schemes used by Runkel and McGrath (1972), Kidder (1981), and Black and Champion (1976), but it is adapted to the work of the U.S. General Accounting Office. Sample surveys, case studies, field experiments, and the use of available data are useful strategies because they can be readily linked to the types of evaluation questions that GAO is asked to answer, and they explicitly accommodate evaluation strategies that are prominent in GAO's

Table 1

Evaluation Strategies and Types of Design

Strategy	Design
Sample survey	Cross-sectional Panel Criteria-referenced
Case study	Single Multiple Criteria-referenced
Field experiment	True experimental Nonequivalent comparison groups Before-and-after (including time series)
Use of available data	Secondary data analysis Evaluation synthesis

Table 2

Characteristics of Four Evaluation Strategies

Evaluation strategy	Type of evaluation question most commonly addressed	Availability of data	Design element		
			Kind of information	Sampling method	Need for explicit comparison base
Sample survey	Descriptive and normative	New data collection	Tends to be quantitative	Probability sampling	No[a]
Case study	Descriptive and normative	New data collection	Tends to be qualitative; can be quantitative	Nonprobability sampling	No[a]
Field experiment	Cause and effect	New data collection	Quantitative or qualitative	Probability or nonprobability sampling	Yes; essential to the design
Use of available data	Descriptive, normative, and cause and effect	Available data	Tends to be quantitative; can be qualitative	Probability or nonprobability sampling	May or may not be available

[a]In this classification sample surveys and case studies do not have an explicit comparison base by definition. This definition is not universal.

history. For simplicity, we speak only of program evaluation, but we imply the evaluation of policies also.

Some of the design elements we identified in chapter 1--in particular, kinds of information, sampling methods, and the comparison base--help distinguish the evaluation strategies. Table 2 shows the relationship between these three design elements and the four evaluation strategies, the types of questions, and the availability of data. In the rest of this chapter, we discuss this relationship in detail. Other design elements--information sources, information collection methods, the timing and frequency of information collection, and information analysis plans--are essential in specifying a design but are less useful in making distinctions among the major evaluation strategies.

Two points about the use of the classification scheme should be stressed. First, as we indicated in chapter 2, a program evaluation design emerges not only from the evaluation questions but also from constraints such as time, cost, and staff. Therefore, the scheme cannot be used independently as a "cookbook" for evaluation. Second, and related to the first point, every evaluation design is likely to be a blend of

several types. Often, two or more design types are combined with advantage.

Each of this chapter's sections on the four evaluation strategies is broken down into subsections on specific design types that may be applicable in GAO. For each type of design, we give several kinds of information: a description of the design, appropriate applications, planning and implementation considerations, and sources of more information. The last section of the chapter makes further connections between evaluation questions and the design types.

THE SAMPLE SURVEY

In a sample survey, data are collected from a sample of a population to determine the incidence, distribution, and interrelation of naturally occurring events and conditions.[1] The overriding concern in the sample survey strategy is to collect information in such a way that conclusions can be drawn about elements of the population that are not in the sample as well as about elements that are in the sample. A characteristic of the strategy is its method of probability sampling, which permits a generalization from the findings to the population. In probability sampling, each unit in the population has a known, nonzero probability of being selected for the sample by chance. The conclusions from this kind of sample can be projected to the population, within statistical limits of error.

Because of the aim to aggregate and generalize from the survey results, great importance is attached to collecting uniform data from every unit in the sample. Consequently, survey information is usually acquired from structured interviews or self-administered questionnaires. The three main ways of obtaining the data are by mail, phone, and face-to-face interviews.

The sample's units are frequently persons but may be organizations such as schools, businesses, and government agencies. A crucial matter in survey work is the quality of the "sampling frame" or list of units from which the sample will be drawn. Since the frame is the operational manifestation of the population, it does much to determine the generalizability and precision of the survey results.

Sample surveys have been traditionally used to describe events or conditions under investigation. For example, national opinion surveys report the opinions of various segments of the

---

[1]The special case in which the sample equals the population is called a "census." The word "survey" is sometimes used to describe a structured method of data collection without the goal of drawing conclusions about what has not been observed. We do not use the term in this narrow sense.

population about political candidates or current issues. A survey may show conditions such as the extent to which persons who support one side of an issue also tend to back candidates who advocate that side of the issue. In the interpretation of such relationships, there is usually no attempt to impute causality.

However, some analysts attempt to go beyond the purely descriptive or normative interpretations of sample surveys and draw causal inferences about relationships between the events or conditions being reported. The conclusions are frequently disputed, but there probably are circumstances in which causal inferences from sample survey data are warranted. Special data analysis methods are required for them, which do not silence methodological criticism but do allow appropriately qualified causal interpretations. In the rest of this section, we describe the designs from cross-sectional, panel, and criteria-referenced sample surveys.

The cross-sectional survey

A cross-sectional design, in which measurements are made at a single point in time, is the simplest form of sample survey.

> EXAMPLE: In 1971, a survey was made of 3,880 families (11,619 persons) to provide descriptive information on the use of and expenditures for health services. A probability sample was drawn from the total U.S. population outside institutions. Because of special interest in low-income, central-city residents, rural residents, and the elderly, these groups were sampled in numbers beyond their proportion in the population so that sufficiently precise projections could be made for these groups. Data were collected by holding interviews in homes, and some of this information was verified by checking other records such as those maintained by hospitals and insurance companies. A large amount of information, projected to the national population, was on topics such as where and why people receive health services, what kind of services they receive, how the services are paid for, and how much they cost.

Applications

When the need for information is for a description of a large population, a cross-sectional sample survey may be the best approach. It can be used to acquire factual information--such as the living conditions of the elderly or the costs of operating government programs. It can also be used to determine attitudes and opinions--such as the degree of satisfaction among the beneficiaries of a government program.

Because the design requires rigorous sampling procedures, the population must be well-defined. The kind of information

that is sought must be clear enough that structured forms of data collection can work. A sample survey design cannot be used when it is not possible to settle on a particular sampling frame before the data are collected. It is hard to use when the information that is sought must be acquired by unstructured, probing questions and when a full understanding of events and conditions must be pieced together by asking different questions of different respondents.[2]

A cross-sectional design can sometimes be used for imputing causal relationships between conditions, as in inferring that educational attainment has an effect on income. Other evaluation designs, such as the true experiment or nonequivalent comparison group designs, are ordinarily more appropriate, when they are feasible. However, practical considerations may rule out these and other designs, and the cross-sectional design may be chosen for lack of a better alternative. When the cross-sectional design is used for causal inferences, the data must be analyzed by techniques such as path analysis (U.S. General Accounting Office, 1982) and structural equation models, although the data collection procedures are the same as for descriptive applications.

Planning and implementation

Sampling. Having a sampling frame that closely approximates the population of interest and drawing the sample in accordance with statistical requirements are crucial to the success of the cross-sectional sample survey. The size of a sample is determined by how statistically precise the findings must be when the sample results are projected to the population.

Pretesting the instruments. To insure the uniformity of the data, the data collection instruments must be unambiguous and likely to elicit complete, unbiased answers from the respondents. Making one or more pretests of the instruments before using them in the survey is an essential preparatory step.

Nonrespondent follow-up. The failure of a sampling unit to respond to a data collection instrument or the failure to respond to certain questions may distort the results when the data are aggregated. Further attempts must be made to acquire missing information from the respondents, and the data analysis must adjust, as well as possible, for information that cannot be obtained.

Causal inference. The procedures for making causal inferences from sample survey data require hypotheses about how two

---

[2]A procedure that is suitable for this situation, called "multiple matrix sampling," applies to each respondent a subset of the total number of questions.

or more factors may be related to one another. Causal analysis methods use the hypotheses to test the consistency of the data. That is, the credibility of causal inferences from sample survey data rests heavily on the plausibility of the hypotheses. For plausible hypotheses, a premium is placed on broad literature reviews and a thorough understanding of the events and conditions in question.

Where to look for more information

Babbie, E. R. Survey Research Methods. Belmont, Calif.: Wadsworth, 1973.

Kidder, L. H. Research Methods in Social Relations, 4th ed. New York: Holt, Rinehart, and Winston, 1981.

Stuart, A. Basic Ideas of Scientific Sampling, 2nd ed. London: Charles Griffin, 1976.

The panel survey

A panel survey is similar to a cross-sectional survey but has the added feature that information is acquired from a given sample unit at two or more points in time.

EXAMPLE: The "panel study of income dynamics," carried out by the Institute for Survey Research at the University of Michigan, is based on annual interviews with a nationally representative sample of 5,000 families. The extensive economic and social data that are collected can be used to answer many descriptive questions about occupation, education, income, and family characteristics. Because follow-up interviews are made with the same families, questions can also be asked about changes in their occupation, education, income, and activities.

Applications

The panel design adds the important element of time to the sample survey strategy. When the survey is used to provide descriptive information, the panel design makes it possible to measure changes in facts, attitudes, and opinions.[3] For making decisions about government programs and policies, dynamic information--that is, information about change--is frequently more useful than static information.

---

[3]Change can also be measured by two or more cross-sectional, time-separated surveys if the samples and data collection procedures are consistent. However, it is possible to associate change on a measure not with an individual but with populations, so that the kinds of questions that can be answered are more limited than with the panel design.

The panel survey's use of time is also important when the survey data are used for causal inference. In this application, the panel design may help settle the question of whether, of two factors that appear to be causally related, one is the cause and the other is the effect.

Planning and implementation

Sampling, pretesting the instruments, nonrespondent follow-up, and causal inference. Panel survey designs are similar to cross-sectional designs in the need for attention to these activities.

Panel maintenance. To the extent that sample units leave the sample, changes in the sample may be mistaken for changes in the conditions being assessed. Therefore, keeping the panel intact is an important priority. When sample units are unavoidably lost, it is necessary to attempt adjustments to minimize distortion in the results.

Where to look for more information

The references in the discussion on cross-sectional survey designs are applicable.

The criteria-referenced survey

Sometimes the evaluation question is, How do outcomes associated with participation in a program compare to the program's objectives? Often, a normative question like this is best answered with a sample survey design (although a criteria-referenced case study design may sometimes be used).

> EXAMPLE: A soil conservation program has the objective of reducing soil loss by 2 tons per acre per year in selected counties. A panel survey could be designed in which actual soil loss on the land that is subject to the program could be compared to the criterion. That is, two measurements of soil depth 1 year apart could be recorded for a probability sample of locations in the targeted counties. Subject to the limitations of measurement and sampling error, the amount of soil loss in the counties could be estimated and then compared to the program objective.

This criterion-referenced survey design employs a probability sample to acquire information on the program's outcome, because a conclusion is sought about a representative sample of the program's population.

A normative evaluation question may also ask, How does actual program implementation match what was intended, or how well does it match a standard of operating performance? The attention is not on outcomes but on processes and procedures.

EXAMPLE: Federal policies require that commercial airlines observe certain safety procedures. A criteria-referenced design could produce information on the extent to which actual procedures conform to these criteria. A population of maintenance procedures--engine overhauls, for example--could be sampled to see if required steps were followed. The infraction rate, projected to the population, could then be compared to the standard rate, which might be zero.

In this example, the safety procedures are a means to an end--the passengers' safety--but the evaluation is focused not on the result but on the implementation of the program's policy on safety.

Applications

Whether dealing with outcomes or process, evaluators can use criteria-referenced designs to answer normative questions, which always compare actual performance to an external standard of performance. However, criteria-referenced designs do not generally permit inferences about whether a program has caused the outcomes that have been observed. Causal inference is not possible, because the criteria-referenced model does not produce an estimate of what the outcomes would have been in the absence of the program.

An audit model--the "criterion, condition, cause, and effect" model--is a special case of the criteria-referenced design that is widely used in GAO.[4] Outcomes, the condition, are often compared to an objective, or a criterion, and the difference is taken as an indication of the extent to which the objective has been missed, achieved, or exceeded. However, it is not ordinarily possible to link the achievement of the objective to the program, because other factors not accounted for may enter into failure or success in meeting the objective.

A variety of evaluation questions lead to the choice of the criteria-referenced design. For service programs, examples are questions about whether the right participants are being served, the intended services are being provided, the program is operating in compliance with legal requirements, and the service providers are properly qualified. Regulatory programs give rise to similar questions: whether activities are being regulated in compliance with the statutory requirements, inspections are being carried out, and due process is being followed.

Sometimes outcome questions are framed in terms of criteria. Did the missile hit the right target? Did the participants of

---

[4]The word "cause" in the audit model has a different meaning from the usual notion of causation. "Purported cause" would be a more accurate term, because the criteria-referenced design does not permit inference about causation.

the training program get jobs? Did the sale of timber yield the expected return? Did supplies of strategic minerals meet the quotas?

Whenever the evaluation questions are normative, criteria-referenced designs are called for. Frequently, but not always, a sample survey is embedded in a criteria-referenced design so that the conclusions can be regarded as representative of the population.

Planning and implementation

Consensus about the criteria. It is often difficult to gain consensus about the objectives of federal programs. When it is difficult, it is also hard to decide which criterion to use in an evaluation. The best way is usually to use not one criterion but several criteria, to allow for the objectives of the miscellaneous interests in the program--legislators, participants, taxpayers, and so on. The problem of consensus is usually of less concern with implementation criteria, because statutes and regulations are more likely to be specific about implementation requirements.

Measuring performance against the criteria. Just as it may be difficult to reach consensus on the objectives of a program, so there is likely to be debate about the procedures for measuring performance against the criteria. For example, Is the analysis of tests of military weapons that use simulated enemy targets a satisfactory way of estimating the probability that the weapons will hit real enemy targets? Similarly, views may differ about the appropriate way to measure performance against implementation criteria.

Where to look for more information

Herbert, L. Auditing the Performance of Management. Belmont, Calif.: Lifetime Learning Pub., 1979.

Popham, W. J. Educational Evaluation. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Provus, M. M. Discrepancy Evaluation. Berkeley, Calif.: McCutchan, 1971.

Rossi, P. H., and H. E. Freeman. Evaluation: A Systematic Approach, 2nd ed. Beverly Hills, Calif.: Sage, 1982.

Wholey, J. S. Evaluation: Promise and Performance. Washington, D.C.: Urban Institute, 1979.

THE CASE STUDY

The case study strategy is less well defined than the other evaluation strategies we have identified and, indeed, different

practitioners may use the term to mean quite different things. For GAO's purposes, a case study is an analytic description of an event, a process, an institution, or a program (Hoaglin et al., 1982). One of the most commonly given reasons for choosing a case study design is that the thing to be described is so complex that the data collection has to probe deeply beyond the boundaries of a sample survey, for example. The information to be acquired will be similarly complex, especially when a comprehensive understanding is wanted about how a process works or when an explanation is sought for a large pattern of events.

Case studies are frequently used successfully to address both descriptive and normative questions when there is no requirement to generalize from the findings. Cause-and-effect questions are sometimes considered, but reasoning about causality from case study evidence is much more debatable.[5]

We present three types of case study design: single case, multiple case, and criteria-referenced designs. Even in a study with multiple cases, the sample size is usually small. If the sample size is relatively large and data collection is at least partially structured, the case study strategy may be similar to the sample survey strategy, except that the latter requires a probability sample.

The single case

In single case designs, information is acquired about a single individual, entity, or process.

> EXAMPLE: The Agency for International Development fostered the introduction of hybrid maize into Kenya. An evaluation using a single case design acquired detailed information about the processes of introducing the maize, cultivating it, making it known to the populace, and using it. The evaluation report is a mini-history constructed from interviews and archival documents.

---

[5]The use of case studies to draw inferences about causality has been approached from diverse points of view. The scope of this paper permits only two examples. One approach is called "analytic induction" and involves establishing a hypothesis about the cause of an effect and then searching among cases for an instance that refutes the hypothesis. When one is found, a new hypothesis about a new cause is established, and the cycle continues until a hypothesis cannot be refuted. The cause associated with that hypothesis is then taken as a likely cause of the effect. Another is in "single case experimental" designs, originated largely in the area of psychology and related to field experiments. With substantial control over and manipulation of the hypothesized cause in a single case, inferences can be made about cause-and-effect relationships.

Single case evaluations are valued especially for their utility in answering certain kinds of descriptive questions. Ordinarily, much attention is given to acquiring qualitative information that describes events and conditions from many points of view, although it may be unstructured data. Interviewing and observing are the common data collection techniques. The amount of structure imposed on the data collection may range from the flexibility of ethnography or investigative reporting to the highly structured interviews of sample surveys. There is some tendency to use case studies in conjunction with another strategy. For example, case studies providing qualitative data might be used along with a sample survey to provide quantitative data. However, case studies are also frequently used alone.

Applications

Three applications of single case studies are illustrative, exploratory, and critical instance. These and other applications are not mutually exclusive categories. They simply draw attention to several common ways of using the case study strategy. Much more detail will appear in "Case Study Evaluations," forthcoming from the U.S. General Accounting Office.

An illustrative case study describes an event or a condition. A common application is to describe a federal program, which may be unfamiliar and seem abstract, in concrete terms and with examples. The aim is to provide information to readers who lack personal experience of what the program is and how it works.

An exploratory case study can serve one or another of at least two purposes. One is as a precursor to a possibly larger evaluation. The case study tells whether a program can be evaluated on a larger scale and how the evaluation might be designed and carried out. For example, a single case study might test the feasibility of measuring program outcomes, refine the evaluation questions, or help in choosing a method of collecting data for the larger study. The other purpose of an exploratory case study is to provide preliminary information, with no further study necessarily intended.

A single case study may also be used to examine a critical instance closely. Most common is the investigation of one problem or event, such as a cost overrun on a nuclear reactor. Here, the question is normative but the issue is probably complex, requiring an in-depth study.

Planning and implementation

Selecting a case. The choice of a case clearly presents a problem, except for the critical instance case study, in which the instance itself defines the study. In other applications, the results will depend to some degree on the case that is chosen. If it is expected that they will differ greatly from case to case, it may be necessary to use a multiple case design.

Impartiality. A case study that uses only qualitative data may present a problem of subjectivity. Subjectivity, in turn, can increase the possibility of systematic bias. The chance of bias should be minimized during the design phase.

Data reliability. Because there are often unstructured elements in the data collection for a case study, the reliability of the data may be doubted. The question is whether two data collection teams examining the same case could, without partiality, end up with quite different findings. Steps must be taken in the planning stages to avoid this form of unreliability.

Data analysis and reporting. Because analyzing and reporting qualitative data are relatively hard, the design for the single case study must have explicit plans for these tasks.

Where to look for more information

Babbie, E. R. The Practice of Social Research, 2nd ed. Belmont, Calif.: Wadsworth, 1979.

Bogden, R., and S. J. Taylor. Introduction to Qualitative Research Methods. New York: John Wiley and Sons, 1975.

Hoaglin, D. C., et al. Data for Decisions. Cambridge, Mass.: Abt Books, 1982.

Multiple cases

Single case designs are weak when the evaluation question requires drawing an inference from one case to a larger group. A multiple case study design may produce stronger conclusions. In our classification, an important distinction between the multiple case study design and sample survey designs is that the latter require a probability sample while the former does not.

> EXAMPLE: A program known popularly as the "general revenue sharing act" appropriated federal funds for nearly 38,000 state and local jurisdictions. An evaluation intended to answer both descriptive and cause-and-effect questions used the multiple case study design. Sixty-five jurisdictions were chosen judgmentally for in-depth data collection, including questionnaires, interviews, public records, and less formal observations. In selecting the sample, the evaluators considered some of the nation's most populous states, counties, and cities but also considered diversity in the types of jurisdiction. Budget constraints required a geographically clustered sample.

In this example, the evaluators balanced the need for in-depth information and the need to make generalizations, and they chose in-depth information over a probability sample. They tried to minimize the limitations of their data by using a relatively large and diverse sample.

Applications

The multiple case study design may be appropriate in evaluating either program operations or program results (and it can be useful for exploratory applications as described for single case designs). The aim is usually to draw conclusions about a program from a study of cases within the program, but sometimes the conclusions must be limited to statements about the cases. When the aim is to make inferences about a program, the best application is probably to base a description of the program's operations on cases from a very homogeneous program. The least defensible application is to try to determine a program's results from cases taken from a heterogeneous program.

Planning and implementation

Selecting cases. In our classification, the case study design does not involve probability sampling. The goal of sampling is shifted from getting a statistically defensible sample to getting variety among the cases. The hope is that insuring variation in the cases will avoid bias in the picture that is constructed of the program.

Uniformity of data. Even though data from several cases may not be aggregated, the frequent need to make statements about a program as a whole suggests the need for uniformity in the data collection. This may conflict with the in-depth, unstructured mode of inquiry that produces the rich, detailed information that can characterize case studies.

The concerns in single case studies about impartiality, reliability, analysis, and reporting apply to multiple cases.

Where to look for more information

The references in the section on single case designs apply.

The criteria-referenced case

Case studies can be adapted for answering normative questions about how well program operations or outcomes meet their criteria.

> EXAMPLE: Social workers must be able to rule out fraudulent claims under the Social Security Disability Insurance Program. To make sure of the uniform application of the law, program administrators have developed standard procedures for substantiating claims for benefits under the program. A case study could compare the social workers' procedures to the procedures that were prescribed by the program's administrators.

The examination of a number of cases might expose violations of prescribed claims-verification procedures. Unlike the criteria-

referenced survey design, the criteria-referenced c
would not permit an estimate of the frequency with
violations occur. It could show only that violatic
not occur and, if they do, it might give a clue as
course, if the number of cases is small and violati
the fact that there are violations may go undetecte
case study approach.

Applications

The applications of the criteria-referenced ca
design are similar to those of the counterpart desi
sample survey strategy. The major difference stems
fact that data from case studies cannot be statisti
projected to a population. However, for a fixed ex
resources, the case study may allow deeper understa
program's operations or outcomes and how these comp
criteria that have been set for the program. Since
can be expensive, care must be taken to insure the
cost estimates before choosing case studies over c:
Two applications are likely: an exploration toward
comprehensive project and a determination of the po
not the probability, that a criterion has not been

Planning and implementation

How to reach consensus on the criteria and how
performance against a criterion--issues that are im
criteria-referenced sample surveys--are considerati
criteria-referenced case studies. In addition, the
how to choose cases for study is crucial because th
may differ, depending on the sample of cases.

Where to look for more information

The references cited above for case studies anc
criteria-referenced survey designs are applicable

THE FIELD EXPERIMENT

The main use of field experiment designs is to dr
inferences about programs--that is, to answer cause-
questions. These designs allow the evaluator to com
example, a group of persons who are possibly affecte
program to others who have not been exposed to the p
evaluation question might be, Does the National Schc
Program improve children's health? To answer the qu
evaluator could compare a measurement of the health
participating in the program to a measurement of the
similar children who are not participating.

Field experiments are distinguishable from labo
experiments and experimental simulations in that fie
ments take place in much less contrived settings. C

inquiry in the field gives reality to the evaluation, but it is often at the expense of some precision in the results. From a practical point of view, GAO's only plausible choice among the three is experiments in the field.

True experiments, nonequivalent comparison groups, and before-and-after studies--the field experiment designs we outline below--have in common that measurements are made after a program has been implemented. Their major difference is in the base to which program participants' outcomes are compared, as can be seen in the first row of table 3. Two other important

Table 3

Some Basic Contrasts Between Three Field Experiment Designs

Basis for contrast	Design		
	True experiment	Nonequivalent comparison groups	Before and after
Measurements of program participants are compared to measurements of	others in a randomly assigned comparison group	others in a nonequivalent comparison group	same participants before program implementation
Persuasiveness of argument about the causal effect of program on participants is	generally strong	quite variable	usually weak except for interrupted series subtype
Administering the design is	usually difficult	often difficult	relatively easy

differences--the persuasiveness of causal arguments derived from the designs and the ease of administration--are shown in rows two and three.

True experiments

The characteristic of a true experimental design is that some units of study are randomly assigned to a "treatment" group and some are assigned to one or more comparison groups. "Random assignment" means that every unit in the population has a known probability of being assigned to each group and that the assignment is made by chance, as in the flip of a coin. The program's effects are estimated by comparing outcomes for the treatment group with outcomes for each comparison group.

EXAMPLE: The Emergency School Aid Act made grants to school districts to ease the problems of school desegregation. An evaluation question was, Do children in schools participating in the program have attitudes about

desegregation that are different from those of children in schools that are desegregating but not participating in the program? For each district receiving a grant, a list was formed of all schools eligible to participate in the program. The population consisted of the schools eligible to participate in the program. Within each school district, some schools were randomly assigned to receive program funds in the treatment group, and the remainder became the comparison group.

Although the true experimental design is unlikely to be applied much by GAO evaluators, it is an important design in other evaluation settings in that it is usually the strongest design for causal inference and provides a useful yardstick by which to assess weaknesses or potential weaknesses in a cause-and-effect design. The great strength of the true experimental design is that it ordinarily permits very persuasive statements about the cause of observed outcomes.

An outcome may have several causes. In evaluating a government program to find out whether it causes a particular outcome, the simplest true experimental design establishes one group that is exposed to the program and another that is not. The difference in their outcomes is attributed, with some qualifications, to the program. The causal conclusion works because, under random assignment, most of the factors that determine outcomes other than the program itself are evenly distributed between the two groups; their effects tend to cancel one another out in a comparison of the two groups. Thus, only the program's effect, if any, accounts for the difference.

Applications

When the evaluation question is about cause and effect and there is no ethical or administrative obstacle to random assignment, the true experiment is usually the design of choice. The basic design is used frequently in many different forms in medical and agricultural evaluations but less often in other fields.

The true experiment is seldom, if ever, feasible for GAO evaluators because they must have control over the process by which participants in a program are assigned to it, and this control generally rests with the executive branch. Being able to make random assignments is essential: the true experimental design is not possible without it. The obstacles might be overcome in a joint initiative between the executive branch and the evaluators, making a true experiment possible. Also, GAO occasionally reviews true experiments carried out by evaluators in the executive branch.

Planning and implementation

Generalization. If the ability to generalize is a goal, a true experimental design may be unwarranted. Generalization

requires that the units in the experiment be a random sample drawn from the population, but in a random sample, more than a few units are likely to refuse to participate.[6] In many true experiments, this limitation may not be serious, because either generalization from the results to a broad population is not a goal or the effects of treatment are expected to be reasonably uniform within the population, so that an attempt can be made to generalize even without a random sample from the population. The latter instance may be likely in some fields such as medicine, where relatively constant treatment effects maybe expected, but is less likely in evaluating government programs and policies.

Maintenance of experimental conditions. In order to apply the logic of random assignment to reasoning about cause and effect, the evaluator must see that the composition of the groups, and thus the integrity of the experiment, is maintained. One of the chief threats to causal reasoning from a true experiment is that the members of the treatment and comparison groups may drop out at different rates. If people drop out more from one group than from another--as they might if they find the treatment disagreeable, for example--then the evaluator's estimate of treatment effects may be distorted. Likewise, if the treatment is allowed to weaken or to vary from participant to participant or to spill over to a comparison group, the findings from the evaluation will be compromised.

Where to look for more information

Babbie, E. R. The Practice of Social Research, 2nd ed. Belmont, Calif.: Wadsworth, 1979.

Keppel, G. Design and Analysis: A Researcher's Handbook, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Kidder, L. H. Research Methods in Social Relations, 4th ed. New York: Holt, Rinehart, and Winston, 1981.

Nonequivalent comparison groups

As with the true experiment, the main purpose of the nonequivalent comparison group design is to answer cause-and-effect questions. A further parallel is that both designs consist of a treatment group and one or more comparison groups. Unlike the groups in the true experiment, however, membership

---

[6] It is important to bear in mind that a random sample from a population and random assignment to a treatment or comparison group are two quite different things. The first is for the purpose of generalizing from a sample to a population; random sampling helps insure external validity. The second is for inferring cause-and-effect relationships; random assignment helps insure internal validity.

in the nonequivalent comparison groups is not randomly assigned. This difference is important because it implies that, since the groups will not be equivalent, causal statements about treatment effects may be substantially weakened.

> EXAMPLE: Occupational training programs try to provide people with skills to help them obtain and keep good jobs. An evaluation question might be, Are the average weekly earnings of program graduates higher than would have been expected had they not participated in the training? Participants have ordinarily selected themselves for enrollment in such programs, which rules out random assignment. It may be possible to compare the participants with members of another group, but the members of the participant group and the comparison group will almost certainly not be equivalent in age, gender, race, and work motivation. Therefore, the raw difference in their earnings would probably not be an appropriate indicator of the effect of the training program, but other comparisons might be suitable for drawing cause-and-effect inferences.

This example is intended to show that although treatment effects can be estimated by comparing the outcomes of the treatment group to those of a comparison group, it is usually not possible to infer that the "raw" difference between the groups has been caused by the treatment. In other words, the two groups probably differ with regard to other factors that affect the difference in outcome, so that the raw difference should be adjusted to compensate for the lack of equivalence between the groups. Using adjustment procedures, including such statistical techniques as the analysis of covariance, may strengthen the evaluation conclusions.

Applications

Nonequivalent comparison group designs are widely used to answer cause-and-effect questions because they are administratively easier to implement than true experiments and, in appropriate circumstances, they permit relatively strong causal statements. Evaluations of health, education, and criminal justice programs can generally collect data from untreated comparison groups but cannot, as we noted above, easily assign subjects randomly to groups in a true experimental design. For example, an evaluation designed to look at the effects of correctional treatment on the recidivism of released criminals through a true experiment would probably not be feasible, because judges base their sentences on the severity of a crime, number of prior offenses, and similar factors, and they would not ordinarily be willing to randomize the correctional treatment that they declare.

Planning and implementation

Formation of comparison groups. The aim of a nonequivalent comparison group design is to draw causal inferences about

a program's effects. The evaluator's two most important considerations in doing this are the choice of the comparison groups and the nature of the comparisons. In the absence of random assignment, treatment groups and comparison groups may differ substantially. Great dissimilarity usually weakens the conclusions, because it is not possible to rule out factors other than the program as plausible causes for the results. For example, to evaluate a nutritional program for pregnant women, it might be administratively convenient to compare program participants in an urban area with nonparticipants in a rural area. This would be unwise, however, because dietary and other such differences between the two groups could easily account for differences in the status of their health and thereby exaggerate or conceal the effects of the program. Therefore, in most circumstances it is advisable to form treatment and comparison groups that are as alike as possible.[7]

Naturally occurring comparison groups. For many evaluations, the evaluator is not the one who formed the treatment and comparison groups. Rather, the evaluator is often presented with a situation in which some people have been exposed to the program and others have not. Although the presence of naturally constituted comparison groups somewhat limits the evaluator's options, the general logic of the design is the same.

Nature of the comparisons. The way in which treatment groups are compared to comparison groups involves statistical techniques beyond the scope of this paper. We can point out, however, that it is important that plans for the comparison be made early, because it will be necessary to collect data on precisely how the groups are not equivalent.

Where to look for more information

Anderson, S., et al. Statistical Methods for Comparative Studies. New York: John Wiley and Sons, 1980.

Cook, T. D., and D. T. Campbell. Quasi-Experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally, 1979.

Huitema, B. E. The Analysis of Covariance and Alternatives. New York: John Wiley and Sons, 1980.

Judd, C. M., and D. A. Kenny. Estimating the Effects of Social Interventions. Cambridge, Eng.: Cambridge University Press, 1980.

---

[7]The evaluator who has precise control over assignments to the group may prefer instead the "regression discontinuity," or biased assignment, design, in which the groups are distinctly different in known ways that can be adjusted for by statistical procedures.

Saxe, L., and M. Fine. Social Experiments: Methods for Design and Evaluation. Beverly Hills, Calif.: Sage, 1981.

Before-and-after designs

The distinguishing feature of before-and-after designs is that they compare outcomes for the units of study before the units were exposed to a program to outcomes for the same units after the program began or after they began to participate in it. There is no comparison group as it exists in the other designs.

> EXAMPLE: A training program was established to help increase the earnings of workers who had few job skills. For a random sample of trainees, an evaluation reported their average weekly income before and after their participation in the program.

Although this simple version of a before-and-after design can be used to answer questions about the amount of change that has been observed, it does not allow the attribution of that change to exposure to the program. This is because it is not possible to separate the effects of the training program from other influences on the workers such as the availability of jobs in the labor market, which would also affect their earnings. The absence of a comparison group sharply weakens the kinds of conclusions that can be drawn because comparison groups help rule out alternative explanations for the observed outcomes.

Before-and-after designs can be strengthened by the addition of more observations on outcomes. That is, instead of looking at a given outcome at two points in time, the evaluator can take a look at many points in time; with a sufficient number of points, an "interrupted time series" analysis can be applied to the before-and-after design to help draw causal inferences. (Such longitudinal data can also be used to advantage with the nonequivalent comparison group design: comparisons can be made between two or more time series.)

> EXAMPLE: After the development of a measles vaccine early in the 1960's, the Centers for Disease Control instituted a nationwide measles-eradication program. Grants were made to state and local health authorities to pay for immunization. By 1972, a long series of data was available that reported cases of measles by 4-week periods. The evaluation question was, What was the effect of the federal measles-eradication program on the number of measles cases? The answer, provided by a before-and-after design using interrupted time series analysis, required distinguishing the effects of the federal program from the effects of private physicians' acting in concert with state and local health authorities.

Before-and-after designs with a number of observations over time may provide defensible answers to cause-and-effect questions. Multiple observations before and after an event help rule out alternative explanations, just as comparison groups do in other designs.

Applications

GAO evaluators are most likely to apply before-and-after designs that employ interrupted time series analysis to data either collected by GAO or made available from other public sources. The Bureau of the Census, the National Center for Health Statistics, the National Center for Educational Statistics, the Bureau of Labor Statistics, and many other such agencies may provide data for investigating the effects of introducing, withdrawing, or modifying national programs. Evaluators will find that the best application is for studies in which a long series of observations has been interrupted by a sharp change in the operation of federal program.

Planning and implementation

Alternative causal explanations. The general weakness of before-and-after designs arises from the absence of comparison groups that could help rule out alternative causal explanations. However, using an interrupted time series can often help make causal arguments relatively strong.

Number of observations. The simple before-and-after design is seldom satisfactory for cause-and-effect arguments, although it may suffice for measuring change. The traditional rule of thumb for interrupted time series analyses says that at least 50 observations are required, but some analysis methods use fewer (Forehand, 1982).

Data consistency. When measurements are made repeatedly, definitions and procedures may change. Care must be taken to see that time series are free of definitional and measurement changes, because these can be mistaken for program effects.

Where to look for more information

Forehand, G. A. (ed.). Applications of Time Series Analysis to Evaluation. San Francisco: Jossey-Bass, 1982.

McCleary, R., and R. A. Hay, Jr. Applied Time Series Analysis for the Social Sciences. Beverly Hills, Calif.: Sage, 1980.

Posavac, E. J., and R. G. Carey. Program Evaluation: Methods and Case Studies. Englewood Cliffs, N.J.: Prentice-Hall, 1980.

THE USE OF AVAILABLE DATA

The evaluation strategies discussed above often involve the need to collect new data in order to answer an evaluation question. Because data collection is costly, it is always wise to see if available information will suffice. Even if the conclusion is that new data should be acquired, the analysis of data that are already available may be warranted for quick if tentative answers to questions that will be more completely addressed with new data at a later time. Available data may be used to address any kind of evaluation question; it need not be the one for which the data were originally collected. We discuss two approaches to the strategy of using available data: secondary data analysis and evaluation synthesis.

In the first approach, the evaluator may both have access to data and need to analyze them after others have done so. For example, secondary data analysis might answer an evaluation question by looking at decennial census data published by the Bureau of the Census and widely used by others.

In an evaluation synthesis, the evaluator combines a number of previous evaluations that more or less address the current question. For example, it might be possible to synthesize several evaluation findings on how behavior-modification programs affect juvenile delinquents in such a way that the synthesized finding is more credible than the finding of any of the several evaluations taken individually.

Secondary data analysis

We refer to secondary data analysis as an approach rather than a design because the data that are involved have already been acquired under an original design for data collection, using some technique such as self-administered questionnaires. If the first design was a sample survey, for example, the analysis might have produced descriptive statistics. The secondary data analysis might produce causal inferences with another method.

> EXAMPLE: Data from 11 sample surveys were used in a major secondary analysis that sought to describe the effects of family background, cognitive skills, personality traits, and years of schooling on personal economic success. The data that were available varied from survey to survey, but overall the investigation focused on American men 25 to 54 years old, and economic success was expressed as either annual earnings or an index of occupational status. Multivariate statistical methods were used to draw inferences about cause-and-effect relationships among the variables.

Applications

Probably the most common application of secondary data analysis in GAO is in answering questions that were not posed

when the data were collected. Many large data sets produced by sample surveys or as part of a program's administrative procedures are available for secondary analysis. The most likely answers in secondary data analysis are descriptive, but normative and cause-and-effect questions can be considered.

Planning and implementation

Access to data. Some data bases, such as those produced by the Bureau of the Census, are relatively easy to obtain. Others, such as those produced by private research firms, may be much more difficult or even impossible to acquire. Confidentiality and privacy restrictions may prevent access to certain data.

Documentation of data bases. There are generally two kinds of documentation problems. Automated data may be difficult to read if the information has been recorded idiosyncratically. The second problem arises when it is hard to understand how the data were collected. How were the variables defined? What was the sample? How were the data collected? How were the data processed and tabulated? How were composite variables, such as indexes, formed from the raw data? Misunderstanding such details can lead to a misuse of the data.

Data mismatched to questions. When the evaluator wants to answer an evaluation question with data collected for another purpose, it is very likely that the data will not exactly meet the need. For example, a population may be a little different from the one the evaluator has in mind, or variables may have been defined in a different way. The solution is to restate the question or to state proper caveats about the conclusions.

Where to look for more information

Boruch, R. F. (ed.). Secondary Analysis. San Francisco: Jossey-Bass, 1978.

Boruch, R. F., et al. Reanalyzing Program Evaluations: Policies and Practices for Secondary Analysis of Social and Educational Programs. San Francisco: Jossey-Bass, 1981.

Hoaglin, D. C., et al. Data for Decisions. Cambridge, Mass.: Abt Books, 1982.

The evaluation synthesis

Some evaluation questions may have been addressed already with substantial research. The evaluation synthesis aggregates the findings from individual studies in order to provide a conclusion more credible than that of any one study.

EXAMPLE: Many studies have been made of the effects of school desegregation. An evaluation synthesis

> statistically aggregated the results of 93 studies of students who had been reassigned from segregated to desegregated schools in order to answer the question of how the achievement of black students is affected when desegregation occurs by government action. The evaluation combined 321 samples of black students from 67 cities. Each of the original studies used some type of field experiment design.

An evaluation synthesis may take any one of several forms. At the opposite extreme of this example, a synthesis may be qualitative but beyond the limits of a typical literature review. The evidence is weighed and qualitatively combined, but there is no attempt to statistically aggregate the results of individual studies.

A variety of synthesis procedures have been proposed for statistically cumulating the results of several studies. Probably the most widely used procedure for answering questions about program effects is "meta-analysis," which is a way of averaging "effect sizes" from several studies. "Effect size" is proportional to the difference in outcome between a treatment group and a comparison group.

Applications

Some form of synthesis is appropriate when available evidence can answer or partially answer an evaluation question. When there is much information of high quality, a synthesis alone may satisfactorily answer the question. If the information falls considerably short, it may be useful to perform an evaluation synthesis for a tentative, relatively quick answer and to follow some other strategy for a more definitive answer

When an issue is highly controversial, the evaluation synthesis may help resolve it, because the synthesis takes account of the variable quality of conflicting evidence. The evaluations being reviewed for the synthesis may be graded for quality. Judgments may be made about what to include from them in the synthesis, or all usable information may be included, as in some forms of meta-analysis. For the latter, the relationship between quality and effect is statistically analyzed.

Syntheses almost always identify gaps in available information. Finding gaps is not the aim of the evaluation synthesis, but a dedicated search for information having revealed them, they can be useful in clarifying a debate. Of course, knowing about information gaps may usefully trigger the gathering of new evidence.

Planning and implementation

Choice of form. The nature of the evidence determines the appropriate form. Quantitative techniques, such as meta-analysis,

are probably the most stringent, but all syntheses require information about how the evaluations being examined were conducted. This means that the evaluator must become familiar with the literature before settling on a form to use.

Selection of studies. In synthesizing evaluations, the evaluator must make important decisions about how to define the population of applicable studies and how to insure that that population or an appropriate sample of it will be examined. Typically, the evaluator logically and systematically screens the population, selecting specific studies for consideration.

Reliability of procedures. A synthesis typically involves the detailed review of many studies, which may be undertaken by several staff members. When the work is divided among evaluators, attention must be given to the reliability of the synthesis procedures that the staff members use. Although consistency of procedure does not alone insure sound conclusions, reliability is necessary. Uniform procedures, such as the use of codebooks, must be established, and checks should be made to verify their effectiveness.

Where to look for more information

Glass, G. V, B. McGaw, and M. L. Smith. Meta-Analysis in Social Research. Beverly Hills, Calif.: Sage, 1981.

Hunter, J. E., F. L. Schmidt, and G. B. Jackson. Meta-Analysis: Cumulating Research Findings Across Studies. Beverly Hills, Calif.: Sage, 1982.

Jackson, G. B. "Methods for Integrative Reviews." Review of Educational Research, 50 (1980), 438-60.

U.S. General Accounting Office, Program Evaluation and Methodology Division. The Evaluation Synthesis. Washington, D.C.: 1983.

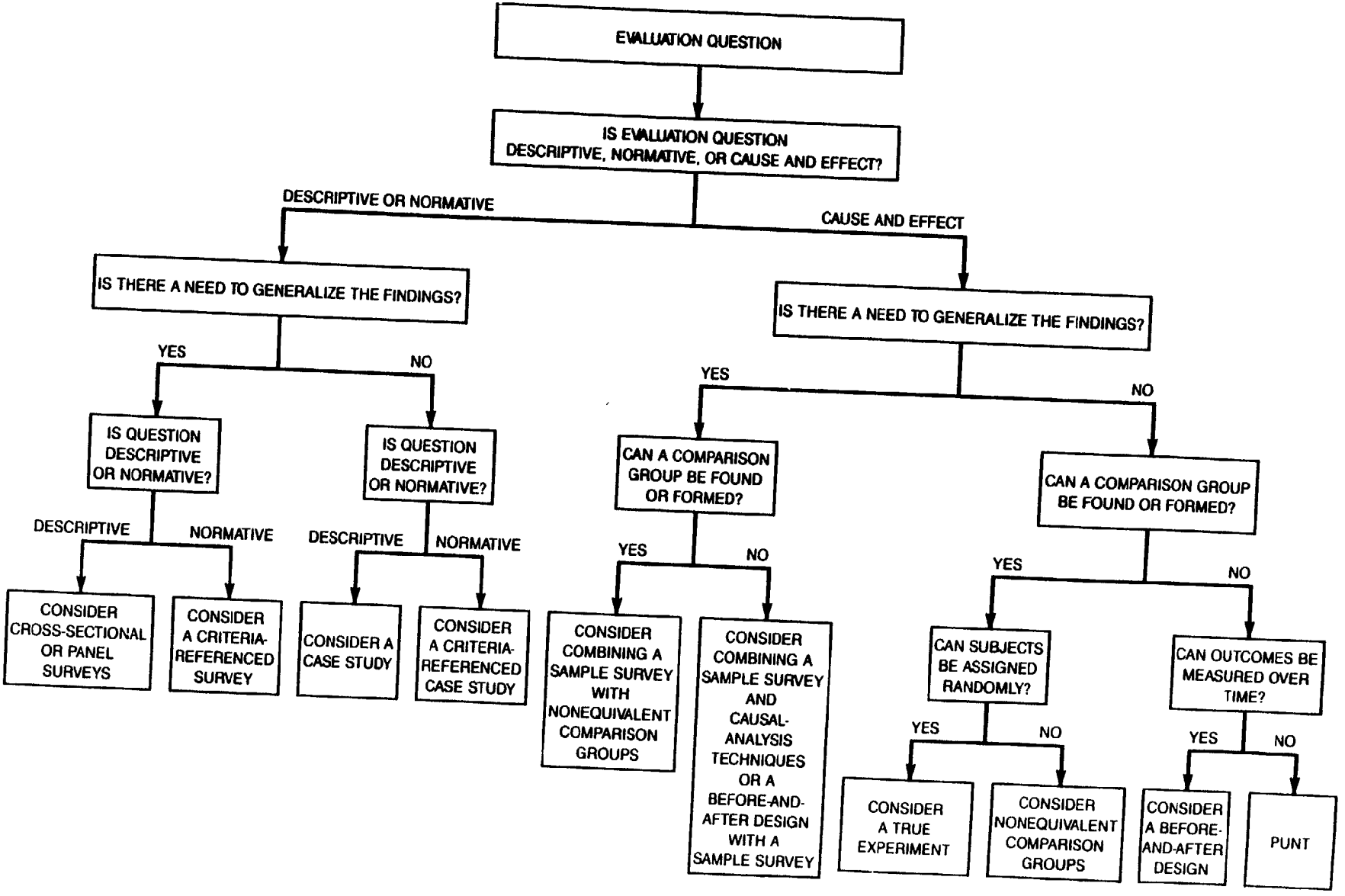
LINKING A DESIGN TO THE EVALUATION QUESTIONS

With particular strategies, designs, and approaches in mind, the evaluator should consider the type of evaluation question being asked and a number of design-screening questions in order to narrow the choices. The point of departure is the evaluation question. Is it descriptive (about how a high-tech training program was implemented)? Is it normative (about whether the job-placement goals of the high-tech training program were met)? Is it causal (about whether the high-tech training program had an effect on job-placement rates)? The answer will partly determine the design or approach to choose.

The choice of what design or approach to settle on is further narrowed with the help of several design-screening

Figure 2

Linking a Design to the Evaluation Questions



questions about the definitiveness needed in the conclusions and the kind of constraints that are expected. An example of the former is, Must we be able to generalize from what we examine in the evaluation to some larger class of things? Examples of the latter are, Can a comparison group be formed? Do we have 6 months or 18 months in which to perform the evaluation?

Figure 2 is a "decision tree" that illustrates this process of choosing an evaluation design. The "branches" at the top of the figure point the way to the answer about the type of evaluation question (descriptive, normative, or causal). Branches further down in the figure point out the place at which to ask design-screening questions (Do we want to generalize the findings? Can a comparison group be found or formed? Can subjects be randomly assigned to groups? Can outcomes be measured over time?).

It must be stressed that the design-screening questions in figure 2 are illustrative and that the figure presents only selected technical matters; for example, approaches using available data have been omitted. Other, equally important factors in choosing a design have also been omitted. They include the availability of resources, the intended use of the evaluation, and the date when the evaluation report is expected. When these factors represent constraints, they put boundaries around what can be done.

As a design evolves, and as the evaluation questions become more specific and research possibilities more narrow, the evaluator must balance the technical considerations against the constraints. For example, it might be necessary to choose between collecting new data, which might answer the evaluation questions comprehensively, and using available data, which is usually the least expensive course and the quickest but may leave some avenues unexplored.

The decision tree almost always ends with the instruction to consider a particular type of design. However, we emphasize the tentativeness in "consider," because we do not want to suggest that there is only one way of designing evaluations. Answers to design-screening questions are not usually as clear-cut as the decision tree suggests, and the relative importance of even these questions may be debated. Furthermore, most evaluations must answer several questions, and where there are several questions, there may be several design types. Even with one question, it may be advisable to employ more than one design. The strengths and the weaknesses of several designs may offset one another. Thus, the decision tree is not a rigid procedure but a conceptual guide for a systematic consideration of design alternatives (McGrath, Martin, and Kulka, 1982).

CHAPTER 4

DEVELOPING A DESIGN: AN EXAMPLE

We have been stressing a consistent theme--that the development of an evaluation design is a systematic process that takes time, thought, and craft. The evaluator must pay careful attention to the formulation of questions and the means of answering them. This painstaking work can be lengthy at the start of a job, but postponing or eliminating it is an invitation to costly delays, incomplete or mediocre data collection, and uncertain analysis. To generate a design is to think strategically; it is to see the link between the questions being asked and the way in which to collect and analyze the data for answering them. Our theme is exemplified in the narrative that follows about the development of a design for a congressionally requested evaluation of the effects of 1981 changes to the Aid to Families with Dependent Children (AFDC) program.

THE CONTEXT

The Omnibus Budget Reconciliation Act of 1981 mandated important changes to AFDC, a major welfare program at the center of debate about welfare and work. On the one hand were people who suggested that providing welfare income reduces a recipient's motivation to work and creates dependence on welfare and a permanent underclass of nonworkers; these people favored strict eligibility criteria for the program and work requirements for welfare recipients. On the other hand were some who suggested that work incentives and work requirements are irrelevant to a welfare population composed largely of households headed by women with small children, who either cannot find work or cannot find work that pays enough to meet their daycare, transportation, or medical expenses.

The AFDC program had grown during the 1960's from 3.0 million to 7.3 million in recipients and from $1.1 billion to $3.5 billion in costs. By 1980, the caseload was 11.1 million persons and the yearly costs were $12.5 billion. Throughout the period, attempts were made to slow the growth.

For example, AFDC's expansion during the 1960's, both in the level of benefits and in the categories of eligibility, had been accompanied by a movement to encourage mothers who were receiving benefits to work. In 1962, a community work and training program had emphasized voluntary training and social services as an alternative to prolonged participation in AFDC.

Another strategy had been to reduce the 100-percent federal tax on the earnings of AFDC families, a tax that was seen as a "disincentive" to work because each dollar earned was a welfare dollar lost. Modifying this strategy in 1967, the Congress incorporated an "earned-income disregard" provision into the

AFDC program, allowing recipients to earn $30 each month with no reduction in benefits--a tax rate of 0 percent--and disregarding one third of all additional earnings.

Along with this change, the Congress enacted the Work Incéntive (WIN) program, in which AFDC recipients could volunteer to receive training services. During the 1970's, however, as the caseload continued to grow, registration in WIN was made mandatory for some AFDC households.

The changes in the AFDC regulations that were specified in the 1981 Omnibus Budget Reconciliation Act focused again on work requirements by allowing the states to operate mandatory "workfare" programs. Other amendments to the legislation changed the policy of allowing working welfare families to accumulate more income than that available to nonworking welfare families. One of the key provisions limited the earned-income disregard to 4 months and the total income of an AFDC household to 150 percent f the AFDC need standards established by each of the states.

THE REQUEST

In June 1982, the House Committee on Ways and Means asked GAO to study the 1981 modifications of the AFDC program. The changes were expected to remove many working AFDC families from the program's rolls, causing many of them to lose their eligibility for Medicaid. Other families would be able to remain on the rolls but with significantly reduced benefits. One concern of the committee was that, faced with the prospect of losing benefits or seeing them greatly diminished, the families would simply choose to work less or quit working entirely. By cutting back on work, they could retain their eligibility for AFDC and Medicaid. However, faced with the loss of benefits, families might instead increase their work effort in order to compensate for the loss.

The committee specifically asked GAO to ascertain (1) the economic well-being, 6 to 12 months after the act's effective date, of the AFDC families that had been removed from the rolls and that had had their benefits reduced and (2) whether families losing benefits had returned to the rolls or compensated for their welfare losses by cutting back on work.

If working families who would lose AFDC or have their grants reduced were to lessen their work effort in order to stay on the rolls, projected budget savings from the legislated changes would be negated or diminished. Therefore, GAO was asked to estimate the budgetary effect of the program changes. The request also required GAO to find out whether the changes had affected family or household composition and to provide information about the demographic, income, and resource characteristics of the AFDC families both before and after the change and the frequency with

which they moved on and off the rolls. The committee asked GAO to make its report early in 1984, which it did with the April 2 report entitled An Evaluation of the 1981 AFDC Changes: Initial Analyses, issued by the Program Evaluation and Methodology Division.

DESIGN PHASE 1: FINDING AN APPROACH

The evaluators began by exploring ways of stating the key questions and strategies for answering them. They reviewed the substantive and the methodological literature and acquired information on the program's operations. They explored the relevance of available data, and they consulted with the committee's staff and other experts.

The literature review centered on welfare dependence, the effects of earlier changes in the program, and the methods other researchers had used to address questions of similar scope and complexity. A systematic reading of the voluminous literature on these topics generated a number of important insights that guided further thinking and refinement of the study. For example, the reading on welfare dependence led to three hypotheses on the 20-year growth of the AFDC caseload. Similarly, the review pointed out areas where information is lacking, such as on the rate at which people leave welfare programs and do not return within a specified time.

The evaluators found that the literature on program effects stressed the need for a longitudinal perspective. They found that the reports relating work to changes in the AFDC tax rates were informative on design approaches as well as on findings. In reviewing the earlier research methods, the evaluators were interested in identifying both designs and measures that fell short or were especially vulnerable and those that were successful. Thus, the review indicated what not to do and suggested strategies that were promising and worth further consideration.

The evaluators also explored the relevance of available data. The ability to make use of existing data sets has the advantage of cutting the cost of collecting, organizing, verifying, and automating information. Five data sets were identified and carefully scrutinized.

The consultation with experts included contact with committee staff, economists, political scientists, social welfare analysts, policy analysts, evaluation specialists, and statisticians. Discussions ranged over a wide number of substantive and methodological issues, and they were held frequently to allow an ongoing critique of the design as it was being formulated. The consultation continued throughout the study, suggesting valuable leads to pursue and dead ends to avoid.

In acquiring information on the operation of the AFDC program, the evaluators paid attention to broad operational

procedures but also concentrated on three areas. The first was how the states determined AFDC benefits before and after the 1981 act and when and how the changes were implemented. The second was how the program was related to other programs from state to state. The third was the relationship in the states between the participation of AFDC families and local economic conditions. Clearly germane to the questions posed by the committee, these interests were stated as questions in language sufficiently general to allow the exploration of multiple ideas and sources of information. The goal was not to foreclose prematurely on potentially useful material that might lead to a thorough understanding of the program's history, how it changed when federal policy was translated to the local level, and whatever would increase the possibility of making cause-and-effect statements.

After about 6 weeks, this group of evaluators, as a design team, began to feel confident about two of several possible designs. Then they began to link alternative designs to evaluation questions.

DESIGN PHASE 2: ASSESSING ALTERNATIVES

The constraints that came to light in phase one shaped subsequent thinking about the job and sharpened the assessment of various alternatives. This allowed the evaluators to refine the evaluation questions, which they did in phase two, so that they could settle on a strategy and a final design.

The first of the constraints began to influence the design when the discussions with experts and numerous visits to the states made it readily apparent that the "national" AFDC program is actually 50 different AFDC programs, one for each state. The heterogeneity was evident in the fact that each state develops its own payment levels and procedures for setting work and child-care expense deductions within the framework of the federal regulations.

For example, the evaluators found considerable variation with respect to two-parent families in requirements about the presence of an unemployed parent, "need" standards, the percentage of the need standard being paid to recipients, and deductions allowable for child-care and work expenses. The variations meant that quite dissimilar grant payments were being made to families whose composition and financial circumstances were identical. The circumstance placed pronounced limitations on the evaluators' ability to generalize from individual states to the nation.

A second constraint was that the states had not timed their implementation of the changes uniformly. Most states began to implement most of the changes in October 1981, but some states did not implement some provisions until 6 months later, in spring 1982. The variation meant that an aggregation of data from all states would be problematic and that generalizations would be

limited. Consequently, the baseline for making comparisons would have to shift from state to state.

Another constraint was that the study could not be predicated on the simple assumption that AFDC recipients would make choices between welfare funds and employment funds. AFDC provides direct income support but also enables the recipients to draw on a number of services, most notably health care under the Medicaid program. Any study of why people choose to stay in or leave the AFDC program has to account for the other benefits. They could play an important, if not decisive, role in influencing financial decisions.

A constraint of a different type had to do with the size of the population of working AFDC recipients. The changes in the legislation were of immediate relevance to working families, but their proportion is small in relation to the total caseload. Nationally, the 1979 figure was about 14 percent, but in some states it was as low as 6 or 7 percent. The small percentages meant that data would have to be collected in a way such that the numbers of earners would be high enough to make statistical projections meaningful.

These and other constraints told the evaluators that to refine the evaluation questions, they would have to pose a study within, rather than between, the states. Similarly, the evaluators began to see the degree to which the study would be able to isolate the effect of the legislative changes from other causal factors, particularly when addressing AFDC recipients' decisions to stop working and stay on the rolls or to remain off the rolls and seek to support themselves through their own earnings. That is, the 1981 changes to the program were initiated at a time when state economies varied widely, so that the economy could not be "held constant," or presumed to be comparable among the states. Thus, it had to be considered a possible cause in earners' decisions. The evaluators also found that their questions would have to account for reductions in other social welfare programs.

As the design team refined the questions, given the constraints on answering them, it was able to examine data collection and analysis strategies. That is, what the evaluators had learned about the questions, and the considerations of time, cost, staff availability, and user needs, enabled the design team to pull together and assess methods for gathering and analyzing data. The evaluators saw two broad strategies, one that would primarily analyze available data and one that would require the collection of original data.

It was thought that using one of the five available data sets would be an economical and quick way to report early findings to the Congress. A data set called the "Job Search Assistance Research Project" (JSARP) was the most promising for a

study of the effects of the changes in the legislation. JSARP was begun by the U.S. Department of Labor late in 1978 as a large-scale effort to measure the effects of job-search assistance, public-service employment, and job training on the employment, earnings, and welfare dependence of low-income persons (not all of whom were AFDC participants). Ten jurisdictions under the Comprehensive Employment and Training Act of 1973 were chosen as "treatment" sites, where special demonstration programs were established to improve the employment opportunities of the target population. Each site was matched with a comparison site as similar as possible in racial and ethnic composition, unemployment rate, primary industries and occupations, size, and location. The researchers interviewed 30,000 respondents in spring 1979, when the demonstration programs were being initiated. Slightly fewer than 3,000 of the respondents had been AFDC recipients for at least part of the year prior to the interview. In 1980, a follow-up interview with 5,700 of the original respondents used substantially the same interviewing instrument; among these respondents were all who had indicated earlier that they had AFDC support, and a large proportion had incomes below 225 percent of the poverty line. Thus, JSARP provides a lengthy record of earnings, other income, work behavior, job search, job training, and family composition for a large sample prior to the institution of the 1981 changes to AFDC.

The evaluators therefore thought that using a before-and-after design and the JSARP data, they could interview the same respondents (or others selected for their similarity to the JSARP respondents) with the same or nearly the same data collection instrument to find out their experiences of the 1981 changes. This would provide for a comparison of work and welfare patterns before and after the program change, although it would not establish with certainty whether the 1981 act was the sole cause of any difference between the two interview periods. Nevertheless, statistical analyses might lead to defensible conclusions about cause.

The alternative strategy, the one that was eventually selected, involved collecting before-and-after data at five sites across the country, making interviews at the five sites with members of working AFDC households who were terminated from AFDC when the 1981 act was implemented, and analyzing national before-and-after data on AFDC caseloads and costs. Of the designs we discussed in chapter 3, this approach included three designs--a nonequivalent comparison group design, a one-group before-and-after design, and a national interrupted time series design.

The plan for the nonequivalent comparison group design was to identify at each site two samples of AFDC recipients, one from a year and a month before the changes and one from the month immediately preceding them. The earlier group would provide a baseline from which to look at the dynamics of work and

welfare both immediately before and after the implementation of the act. Both samples would allow for separate subsamples of working and nonworking AFDC recipients. Depending on the completeness of case records at the sites, the following information could be compared: length of participation in AFDC, percentage of AFDC households with earnings at different times, percentage of households leaving and then returning to the rolls, average dollar amounts of AFDC benefits and earned income, percentage of households drawing on various other welfare benefits, and reasons for the termination of AFDC payments. Thus, the comparisons could be both within and between groups and of several types across three points in time (the baseline and before and after implementation). The evaluators could compare the static characteristics of earners and non-earners, the employment status of the various groups, and the relationship between changes in administrative practices and the behavior of the respondents in terms of the time they spent on AFDC's rolls, their average net earnings, and what they did because of changes in AFDC benefits.

Having decided on this approach, the evaluators constructed interviews within the case study component that were intended to collect data on and assess the economic well-being of the persons who were removed from the rolls, how they coped with the loss of benefits, and whether they worked more to keep up an income. Here, the comparisons were to be within groups of households before and after the program changes. For example, the evaluators could compare household composition, employment status, earnings, and total disposable income. Of particular interest would be data on whether people increased their work effort or shifted their reliance for support to other programs such as General Assistance or Unemployment Insurance.

The national analysis component, with its interrupted time series analysis, would rely on data provided by the U.S. Department of Health and Human Services and by state welfare departments on the operation of AFDC programs, including the implementation of the 1981 provisions, and on caseloads and outlays for AFDC and related programs. The objectives that were planned were to document the degree to which the 1981 AFDC provisions represented change from past practices, to explore their effects on national AFDC caseloads and costs, and to determine whether some states tried to negate or reduce the effects of certain provisions. The design team planned for a request of all the states to provide GAO with the results of their own independent evaluations.

Two smaller and complementary components were also posited. One would use archival data and the other would require conducting interviews with state and local program officials and staff. The archival data would include information on AFDC caseload fluctuations and local economic conditions. Collecting these data would explore the degree to which different patterns of dependence on AFDC in three periods might be the product of

events other than the AFDC changes, such as a deteriorating labor market.

DESIGN PHASE 3: SETTLING ON A STRATEGY

In the end, a choice has to be made between competing design options. The difficulty for the evaluator making this choice is in assessing the alternatives. Each one will have strengths and weaknesses, so that the decision comes to what will be both most feasible and most defensible. In the AFDC study, the choice was made in favor of the multi-strategy approach. The JSARP approach using available data and interviewing a sample of the original respondents was dropped.

To be sure, both approaches had strengths, and strong arguments were made for both. The scales tipped against the simpler approach when it came to weaknesses. There were several reservations about using the JSARP data. They had problems with respect to accuracy, precision, and completeness (largely because the respondents' reports of AFDC participation were retrospective to as far as 18 months). There was a possibility of bias, since 23 percent of the original respondents did not turn up for the second set of interviews, and the difficulty of finding the respondents for the new study could be even greater. There were not enough earners in the sample. And, finally, practical problems included the fact that the JSARP data were not for public use and might not be either obtainable or useful, complete, or accurate.

In light of all this, the multi-strategy approach was adopted. Even with it, there was concern about the availability of case records, finding respondents who had left the AFDC program, the extensive time required to code case records at sites that did not have automated data, the ability to control for disparate economic conditions site by site, and the sheer volume of data that would have to be gathered, coded, analyzed, and reported. However, compared to the concern about JSARP, which tended to be analytical, these problems were more simply procedural. In the end, it was concluded that the analytical problems were a greater threat to the ability to answer the study questions than the procedural ones.

BIBLIOGRAPHY

Anderson, S., et al. Statistical Methods for Comparative Studies. New York: John Wiley and Sons, 1980.

Babbie, E. R. Survey Research Methods. Belmont, Calif.: Wadsworth, 1973.

-----. The Practice of Social Research, 2nd ed. Belmont, Calif.: Wadsworth, 1979.

Black, J. A., and D. J. Champion. Methods and Issues in Social Research. New York: John Wiley and Sons, 1976.

Bogden, R., and S. J. Taylor. Introduction to Qualitative Research Methods. New York: John Wiley and Sons, 1975.

Boruch, R. F. (ed.). Secondary Analysis. San Francisco: Jossey-Bass, 1978.

-----, et al. Reanalyzing Program Evaluations: Policies and Practices for Secondary Analysis of Social and Educational Programs. San Francisco: Jossey-Bass, 1981.

Chelimsky, E. "The Definition and Measurement of Evaluation Quality as a Management Tool." Management and Organization of Program Evaluation, ed. by R. G. St. Pierre, pp. 113-26. San Francisco: Jossey-Bass, 1983.

Cook, T. D., and D. T. Campbell. Quasi-Experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally, 1979.

Cronbach, L. J. Designing Evaluations of Educational and Social Programs. San Francisco: Jossey-Bass, 1982.

Forehand, G. A. (ed.). Applications of Time Series Analysis to Evaluation. San Francisco: Jossey-Bass, 1982.

Glass, G. V, B. McGaw, and M. L. Smith. Meta-Analysis in Social Research. Beverly Hills, Calif.: Sage, 1981.

Herbert, L. Auditing the Performance of Management. Belmont, Calif.: Lifetime Learning Pub., 1979.

Hoaglin, D. C., et al. Data for Decisions. Cambridge, Mass.: Abt Books, 1982.

Huitema, B. E. The Analysis of Covariance and Alternatives. New York: John Wiley and Sons, 1980.

Hunter, J. E., F. L. Schmidt, and G. B. Jackson. Meta-Analysis Cumulating Research Findings Across Studies. Beverly Hills, Calif.: Sage, 1982.

Jackson, G. B. "Methods for Integrative Reviews." Review of Educational Research, 50 (1980), 438-60.

Judd, C. M., and D. A. Kenny. Estimating the Effects of Social Interventions. Cambridge, Eng.: Cambridge University Press, 1980.

Keppel, G. Design and Analysis: A Researcher's Handbook, 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Kidder, L. H. Research Methods in Social Relations, 4th ed. New York: Holt, Rinehart, and Winston, 1981.

McCleary, R., and R. A. Hay. Applied Time Series Analysis for the Social Sciences. Beverly Hills, Calif.: Sage, 1980.

McGrath, J. E., J. Martin, and R. A. Kulka. Judgment Calls in Research. Beverly Hills, Calif.: Sage, 1982.

Popham, W. J. Educational Evaluation. Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Posavac, E. J., and R. G. Carey. Program Evaluation: Methods and Case Studies. Englewood Cliffs, N.J.: Prentice-Hall, 1980.

Provus, M. M. Discrepancy Evaluation. Berkeley, Calif.: McCutchan, 1971.

Rossi, P. H., and H. E. Freeman. Evaluation: A Systematic Approach, 2nd ed. Beverly Hills, Calif.: Sage, 1982.

Runkel, P. J., and J. E. McGrath. Research on Human Behavior: A Systematic Guide to Method. New York: Holt, Rinehart, and Winston, 1972.

Saxe, L., and M. Fine. Social Experiments: Methods for Design and Evaluation. Beverly Hills, Calif.: Sage, 1981.

Stuart, A. Basic Ideas of Scientific Sampling, 2nd ed. London: Charles Griffin, 1976.

U.S. General Accounting Office, Program Evaluation and Methodology Division. Causal Analysis: A Method to Identify and Test Cause-and-Effect Relationships in Program Evaluation. Washington, D.C.: 1982

-----. The Evaluation Synthesis. Washington, D.C.: 1983.

Wholey, J. S. Evaluation: Promise and Performance. Washington, D.C.: Urban Institute, 1979.

(973177)

*U.S. GOVERNMENT PRINTING OFFICE : 1985 0-528-526/30745