

## Chapter 1

### ACCOUNTABILITY

In 1978 Proposition 13 was passed by California voters and limited the growth of state government. Other states such as Massachusetts passed similar legislation to curb the expansion of state services. Five months after the passage of Proposition 13, the *Washington Post* conducted a poll indicating that American taxpayers were not nearly as angry with their tax bills as they were with how little they got from government in return.

The public is uninformed about how tax dollars are spent. The press periodically publishes an exposé on the misuse of government funds. An employee in the private sector may be smug about his or her alleged superiority in productivity and may joke about the inefficiency of government bureaucracies and the public servants who work for them. "Tax money," said the *Post* story announcing poll results, "is seen as largely wasted by local, state and federal governments that have padded payrolls and employees who are overpaid, lazy, discourteous and inefficient" (quoted in Greiner et al., 1981: xv).

There is some basis for the perception of public employees as overpaid, lazy, and inefficient. However, it represents a bias that is largely based on a convenient stereotype. The perpetuation of this stereotype has many positive consequences for those who promote it. For the press, it is a convenient news item, a straw man approach to sensationalistic news coverage and an easy alternative to responsible journalism. For the private sector it is a basis for maintaining a false sense of superiority in the management of their organization. For the public it has been a scapegoat for focusing the frustration of mandat-

ory taxes individuals feel powerless to direct or contain. As a result, they remain largely ignorant of the services their tax dollars purchase. This is beginning to change as a result of the growing tax burden.

The perception of an unaccountable government will only change when individuals are provided with information and have personal experiences with agencies that are accountable to the taxpayer. This book is an example. It is written by a taxpayer who has experienced the process of bringing accountability to the public sector. It is an exciting story. It is a story of how some service agencies have created an environment for complete accountability. Most agencies stop short of accounting for what happens to their clients. They do not indicate the difference they made in the lives of the clients they have served; they only account for how many they served and at what cost. This is limited or incomplete accountability.

The primary focus of this book is on human services programs such as day-care, foster care, child and adult protective services, adoption, delinquency services, counseling, rehabilitation, and employment services. In addition to these there are a number of programs that seek to improve the lives of clients through intervention strategies. Table 1.1 lists the 1982 funding for many of the human services programs.

### **THE REAGAN ADMINISTRATION'S NEW FEDERALISM**

President Reagan has provided a national context for the review and assessment of the proper role of government in the U.S. economy. His proposed program includes several dimensions:

- to reduce the regulatory role of the federal government by eliminating many of the federal regulations;
- to reduce the level of federal funding for many social programs by consolidating them into block grants to be distributed at consecutively lower levels;
- to transfer entire programs to states in return for the federal takeover of others; and
- to encourage volunteerism and private sector initiatives to substitute for government funding of social programs.

These dimensions shift the focus of accountability away from the federal level and reinforce state and local level accountability

mechanisms. It is now even more important to review these mechanisms to ensure that they are appropriate, available, and used by key decision makers in allocating increasingly scarce public resources.

### ACCOUNTABILITY MEASURES

There are many kinds of accountability measures, basically falling into two categories: efficiency and effectiveness. *Efficiency* refers to the number of units of output per resource. An example would be the number of services delivered per worker. *Effectiveness* refers to the number of desired outcomes as a result of outputs. An example of effectiveness is the number of delinquents who are served who do not

**TABLE 1.1** Federal Funding of Key Human Services Programs Administered by State and Local Governments (in millions of dollars)

<i>Program</i>	<i>Funding</i>	<i>FY 1982 Appropriation</i>
Alcohol, Drug Abuse, and Mental Health Block Grant		432.0
Social Services Block Grant		2400.0
Community Services Block Grant		366.1
Administration on Aging		636.5
Developmental Disabilities		58.7
Rehabilitation Services Grants to States		863.0
Foster Care		299.8
Adoption Assistance		5.0
Child Welfare Services		156.3
Child Welfare Training		3.8
Child Abuse and Neglect		16.2
Runaway Youth		10.5
Head Start		911.7
Work Incentive (WIN)		280.8
Refugee and Entrant Assistance		670.7
Adolescent Family Life		11.0
Total		7122.1

SOURCE: American Public Welfare Association, "Lame-Duck Congress Protects Human Services from Budget Ax," *Washington Report*, Vol. 18, No. 1, p. 4. Copyright © 1983 by the American Public Welfare Association. Reprinted by permission.

become recidivists. I believe effectiveness should receive primary emphasis.

The accountability measure employed varies according to where an individual is located in the decision-making process. The closer a staff member is to the actual delivery of service to clients, the more likely he or she is to use effectiveness as a criteria. Generally, administrators emphasize efficiency rather than effectiveness. A recent review of efficiency and effectiveness criteria used at the state level illustrates that effectiveness measures are not often assigned much importance.

Aurora Associates and Westat, Inc., two Washington-based research firms, reviewed the criteria used to assess the efficiency and effectiveness of state social service block grant programs in Iowa, Florida, Maine, Texas, Washington, and Wisconsin. The researchers interviewed legislative appropriation committee members and their staff, agency commissioners, and program managers. Tables 1.2 and 1.3 display effectiveness and efficiency measures from the level of legislative decision maker to county or regional administrators. Table 1.4 summarizes both types of accountability measures by availability and use by decision makers.

Ann Majchrzak, the project director, concluded that various efficiency and effectiveness measures were used. However, processing and cost measures (i.e., number of clients served), were systematically collected and used, but effectiveness measures were gathered on a special one-time study basis. States have institutionalized the collection of the number and cost of the services that are delivered but have not institutionalized the collection of the data on the consequences to clients of the delivery of these services. As a result there is no systematic review of the impact of the services provided but, instead, the focus is on how to deliver more or cheaper units of service as they currently exist. Once a program gets an appropriation, it continues without serious review of its impact; it is assumed to be helpful because it is funded. Various interest groups evolve to perpetuate the program, and they become much more important politically as a reason to continue the program than the effectiveness of the program itself.

### ETHIC OF INTRINSIC GOODNESS

There is limited accountability in terms of effectiveness because there is a basic unquestioned ethic of the intrinsic goodness of government programs by those who promote them.

TABLE 1.2 Illustrative Definitions of Effectiveness

<i>Organizational Levels</i>	<i>Definitions</i>	<i>Example Measures</i>
Legislature	Are programs reaching targeted populations?	Number and type of services delivered; Number and type of clients served
	Are services needed?	Not available
	How is state doing compared to other states (in terms of inputs, processes and outputs)?	Length of time to complete adoption process relative to other states
	Providing services in cheapest and easiest way	Total cost and services provided
	Meeting need for services	Not available
	Appropriateness of services	Public/client complaints
Agency Executives	Appropriateness of services	Number of services provided for each program area; Number of court-ordered services for each program area
	Degree to which services are delivered	Number of services delivered
	Appropriateness of services	Citizen/consumer complaints about program
	Optimal mix of programs within state	Not available
	Processing of clients through social service system appropriately	Compliance with standards of care; client complaints
	Meeting objectives for program outputs	Number of services provided by each program
Program Directors	Delivery of appropriate services	Timeliness, eligibility
	Meeting objectives for program outputs	Number of completed service plans; halt growth in institutional care

*(continued)*

TABLE 1.2 Continued

<i>Organizational Levels</i>	<i>Definitions</i>	<i>Example Measures</i>
	Changes in client outcomes	Recidivism rate of foster care children; reduction in recurrence rate of child abuse
	Compliance with quality standards	Response to child abuse referrals within 24 hours
	Delivery of services to clients most in need	Number and type of clients served
County/Regional Social Service Administrators	Did services achieve goals of client service plan?	Reasons for case closure
	Compliance with quality standards	Length of time child in foster care residential care facilities
	Proper case management	Review of cases for compliance with procedural requirements
	Appropriateness of services	Public/client complaints

SOURCE: A. Majchrzak, A. Schroeder, and R. Patchen, *Assessing and Improving State Social Service Programs*. Contract No. DHHS-HEW 100-81-003. Washington, DC: Aurora Associates and Westat, Inc. Reprinted by permission of Aurora Associates and Westat, Inc.

Dr. William Benton, previous Deputy Secretary for the Maryland Department of Human Resources, summarized the reasons why there has been little serious assessment of the effectiveness of human services. In a keynote address to the International Council on Social Welfare, Benton contended that there has never been a question of effectiveness because of the "ethic of intrinsic goodness."

Since their inception in the United States, personal social services have tended to be encapsulated in an ethic of intrinsic goodness. Perhaps, due to their philanthropic origins, social services have been widely viewed as a symbolic commitment of a society to do good. As a result, the extent of a society's commitment has traditionally been measured in the size of the investment in programs to meet the economic and social needs of deserving individuals, families and communities. That is, the more we would spend on social programs, the better.

**TABLE 1.3 Illustrative Definitions of Efficiency**

<i>Organizational Levels</i>	<i>Definitions</i>	<i>Example Measures</i>
Legislature	More services for same cost	Program costs by total services delivered
	Delivery of services at lowest cost	Unit cost of services; size of workload
	Indirect versus direct costs	Total administrative costs for agency
	Comparative program costs	Total program costs versus costs of other programs
Agency Executives	How well mandated job done?	Timeliness of services delivered
	Maximum service for lowest dollars	Total caseload by overall costs
	Delivery of most beneficial services	Projected versus actual program outputs
	Controlling program costs	Total costs for total clients served; average cost per client per unit of service
Program Directors	Cost of service	Ratio of supervisors to caseworkers
		Unit cost of service
		Cost per client per service
		Time to deliver services
	Workload standards	
	Monthly expenditures	
	Serve clients in least costly, most effective manner	Serve clients in community care rather than institutional care
	Compliance with procedural service standards	Monitoring of required services
	Caseworker productivity	Ratio of caseworker time in direct service to time spent on paperwork

*(continued)*

TABLE 1.3 Continued

<i>Organizational Levels</i>	<i>Definitions</i>	<i>Example Measures</i>
	Avoid overdelivery of services	Minimize duplication; service hours authorized versus service hours used
	Controlling service costs	Projected versus actual expenditures
County Social Service Administrators	Provision of services in a timely manner	Caseworker time for different activities
	Meeting specifications of procedural regulations	Review of cases to ensure compliance with quality standards

SOURCE: A. Majchrzak, A. Schroeder, and R. Patchen, *Assessing and Improving State Social Service Programs*. Contract No. DHHS-HEW 100-81-003. Washington, DC: Aurora Associates and Westat, Inc. Reprinted by permission of Aurora Associates and Westat, Inc.

The ethic of intrinsic goodness has had a profound effect on the development of personal social services at two levels.

First, on an individual level, the notion that personal social services are inherently good has created a corps of professionals whose altruism is unrivaled.

Second, at a societal level, the ethic of intrinsic goodness has undoubtedly resulted in an investment in personal social services which is much larger than it might otherwise have been.

In spite of its important and beneficial role in the evolution of the personal social services, there are substantial limits to the ethic of intrinsic goodness.

First, the ethic has tended to reduce social policy to arguments of "more" versus "less." Since the ethic of intrinsic goodness accepts the efficacy of social services as an article of faith, the most logical way to make a program better has been to increase our investment in it. As long as additional resources have been forthcoming, the ethic of intrinsic goodness has precluded much systematic questioning of the need for human service reform.

Second, the ethic has led to an inordinate investment in "symbolic" programs. Personal social services have often been established to respond to the needs (real or perceived) with little understanding of the technology involved in effecting the conditions which brought about the service in the first place. Since the ethic of intrinsic

TABLE 1.4 Criteria of Effectiveness and Efficiency Useful to States

<i>Possible Types of Criteria</i> <sup>1</sup>	<i>Observations from Site Visits</i>
Performance of program over time	Highly preferred by state officials, especially for effectiveness defined as changes in client outcomes (e.g., decreased institutional placements) and efficiency defined as amount and costs of services delivered. However, historical trend data using similar service definitions were rarely available.
Performance of program across different organizational units (e.g., counties)	Considered by state officials to be particularly useful for identifying problem areas (e.g., regions) or providers.
Performance of program relative to a prespecified performance target	Used particularly for effectiveness defined in terms of processes (e.g., standards of care) or program outputs (e.g., "performance objectives") and for efficiency assessments using projected costs, a standard unit of service, or workload standards.
Performance of program relative to comparative programs	Primarily used to compare costs of different service program areas. However, state officials indicated that any such comparison is difficult because of the noncomparability of services and client population.
Performance of different treatment models within program	Typically used on an ongoing basis for assessments of efficiency only (e.g., residential versus in-home treatments); used to assess relative effectiveness on a special study basis only.
Performance of comparative client types served by similar programs	Not generally mentioned as a useful criterion, possibly because most states do not collect relevant data

1. Adapted from Millar and Millar (1981).

goodness views personal social services as worthwhile by definition, their efficacy has been beyond question. Indeed, rigorous evaluation of these programs has tended to be viewed as a form of attack on the service under review. As a result of the ethic's inability to differentiate between the symbolic and the substantive, credibility on the whole has suffered.

Most critically, the ethic of intrinsic goodness is irrelevant to conditions of scarce and ever-declining resources. The debate over pro-

posals to reduce Federal spending for human services is informed not one whit by an ethic which enables us to say only that budget cuts are "bad" and that if reductions could be avoided that would be "good" [1981: 1-2; reprinted by kind permission of William Benton].

In short, the ethic of intrinsic goodness has limited the need to collect effectiveness information. In addition, efficiency data are routinely collected, are available, and thus are used by decision makers in holding public agencies accountable.

When effectiveness data are requested, they are generally collected as parts of special studies, which are often called "program evaluations." These studies are limited in resources, timeliness, and usefulness.

### LIMITS OF PROGRAM EVALUATIONS

Program evaluation uses the scientific method to study practical problems and answer administrative questions about the efficiency and effectiveness of programs.

Program evaluation has limitations imposed by the agency and/or environment that reduce the likelihood of the use of the findings. Most of these limitations restrict the appropriate use of scientific methodology. In order to infer a cause-effect relationship and thus prove that a service caused a desired client impact, the program evaluation research design must include

- random assignment of subjects into
- experimental and control groups and
- a large enough sample to draw reliable statistical inferences from
- data having high validity

There are several different scientific research designs that vary in their ability to draw inferences about cause and effect. Donald Campbell and Julian Stanley (1963) detailed these experimental and quasi-experimental designs. Different approaches to evaluation use these designs and thus vary in their ability to draw inferences about cause and effect.

Few of the four conditions necessary for a true experimental research design (let alone a stable enough environment for the groups to remain relatively constant in their differences) exist in most public

agencies (Carter, 1980). In addition to the limits of conducting true experiments, there are other more serious timing and resource issues restricting the use of program evaluation as a means of collecting effectiveness data.

First, program evaluations usually take from two to three years to complete. Generally, the most pressing resource allocation issues require more immediate resolution in a constantly changing, highly unpredictable environment. In order to address an issue of this nature an evaluation study must be completed in six to twelve months.

Second, program evaluations are more costly than most organizations are willing to expend for outcome information. This is particularly true during periods of severe budget cuts.

Third, program evaluation findings are usually drawn from a sample of offices, workers, or clients and are open to criticism of questioning the representativeness of the sample. Theoretically, this can be overcome by claims of scientific random sample methodology. Practically, opponents can always claim that *something* has changed and the findings no longer apply.

Fourth, the persuasiveness of an argument is much more powerful than scientific facts in resolving competing positions.

Fifth, in order to change the perceptions of administrators and/or legislators it is important to have a long enough history of consistent information. Program evaluation studies seldom have the historical trend information because they focus on one-shot special studies (Cox, 1977).

*Given these limitations, it is important not to follow the program evaluation approach but instead to institutionalize the collection of effectiveness measures just as efficiency measures have been institutionalized.* If regularly collected, the results will be available on a continuous basis. If available, they will be used along with efficiency measures for a complete account of program funds.

### **BENEFITS OF REGULARLY COLLECTING EFFECTIVENESS DATA**

Most employees of a public agency care about their program. However, few know whether their work has an impact on the clients they serve. This book describes how everyone who cares can also know if they are truly being effective.

Some of the many benefits of regularly collected effectiveness data are that

- workers experience increased morale as a result of knowing which clients were helped;
- the public has increased confidence in government when they know that programs are effective;
- public administrators have confidence in the extent to which services are successful. Over time, such information provides a baseline for assessing new program initiatives that determine if they are more or less effective than former programs.
- better public relations result when increased client impact information enables an agency to respond more accurately to sensational newspaper articles.

The balance of the book provides you with a variety of strategies to implement a complete accountability system. It is written from my experience as a state administrator and is primarily designed to assess barriers commonly encountered at the state level. Many of these same variables, however, exist in all agencies and can be applied by every administrator.

### EXERCISES

- (1) Ask a sample of friends, neighbors, relatives, or other tax payers to list ten public agencies and to describe what kinds of services they perform. Then ask them if they believe any of them are doing a good job. Finally, ask them how they know the agency is doing a good job. The results of this exercise will reveal the public's general absence of knowledge about agencies that are funded by their tax dollars.
- (2) Create a list of the major departments funded by the state budget. Ask a sample of the public to distribute in percentages how they would prefer to have their tax dollar spent and why. The results will give you some perspective on the general public and the criteria they use in determining which agencies are performing an important function.

- (3) Ask a sample of the public about their perceptions of human services agency employees and their basis for the perceptions.
- (4) Pick a specific service program (i.e., day-care, runaway services, domestic violence) and find out the current measures of efficiency and effectiveness. Then interview the program or agency director and ask how each of these measures are used in monitoring performance levels.
- (5) Interview a program manager, first-line worker, or supervisor and ask which informal or indirect measures of success they use to know when an intervention strategy is effective. These nonmeasured or unofficial indicators are often preferred by workers because they are perceived as a more relevant outcome.
- (6) One of the criteria for a scientific experimental evaluation study is the use of a random sample of cases that are divided into an experimental group (those who get a specific service) and a control group (those who do not get any service). Choose an experimental human services program and design a study to select clients to be placed in both groups. Begin the process by obtaining a list of all clients to be assigned. Describe how you would assign clients so that each group would be almost identical before you even begin the experiment. What kinds of objections do you expect to encounter from the agency in order to actually accomplish this initial step in the establishment of the classical experimental research design?
- (7) One of the criteria for the classic experimental research design is that the data should be highly reliable and valid. *Valid* refers to the accuracy of the information and *reliability* refers to the consistency of the process of gathering the information. Choose a form that is currently used by a human services agency to record information about clients. Interview five workers in terms of their understanding of the meaning of the data elements. If all five have the same meaning, then the form has high reliability. Then ask the workers if the same data are accurate and how they know the elements are accurate. This will give you some understanding of the validity of information. The exercise is important because it illustrates the general lack of consensus on the use of any given data element.

- (8) Locate an evaluation that has been done in an agency and find out how long it took to complete it and at what cost. If possible, talk to the administrator of the agency and ask how the results were used.