

Third Edition

PROGRAM EVALUATION

Methods and Case Studies

EMIL J. POSAVAC

RAYMOND G. CAREY



Third Edition

PROGRAM EVALUATION

Methods and Case Studies

EMIL J. POSAVAC

Loyola University of Chicago

RAYMOND G. CAREY

Parkside Associates, Inc.

Prentice Hall, Englewood Cliffs, New Jersey 07632

Acknowledgments

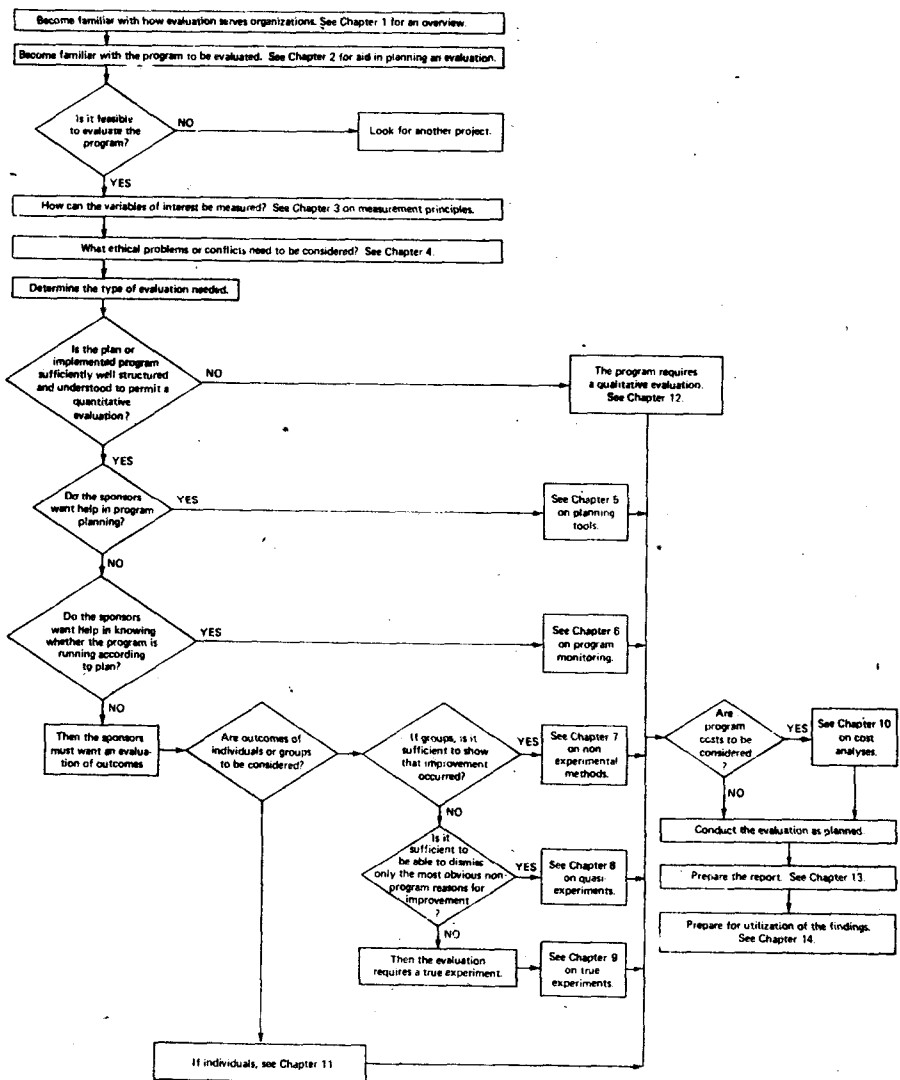
The following organizations and journals are thanked for permission to re-print or adapt material for this text: Institute for Program Evaluation, Inc., Roanoke, Virginia; *General Hospital Psychiatry*; *Journal of Alcohol & Drug Education*; *Evaluation Review*; *Evaluation*; *Evaluation & The Health Professions*; and American Psychological Association.

A number of secretaries and assistants have helped us in the preparation of the manuscript. At various times the following people have given much appreciated help: Carolyn Foraker, Karen Gianfortune, Bernadette Jaroch-Hagerman, and Susan Borkowski Lueger.

The authors are grateful to their respective institutions for encouragement during the writing of this text. The first author's contribution to the first edition was greatly facilitated by a sabbatical leave awarded by Loyola University of Chicago. The second author acknowledges with thanks the explicit support received from Lutheran General Hospital.

The authors also wish to express their appreciation for the helpful comments of the following reviewers: Amado M. Padilla, University of California at Los Angeles; Alan Siman, School of Social Work, San Diego State University; Ross F. Conner, University of California at Irvine; Brian T. Yates, American University; Myron Mast, School of Public Service, Grand Valley State College; Clara Mayo, Boston University; Richard M. Wolf, Teachers College of Columbia University; Robert J. Calsyn, University of Missouri at St. Louis; Francis G. Caro, Institute for Social Welfare Research of the Community Service Society of New York; and Fred B. Bryant, Loyola University of Chicago. For the third edition: James R. Cook, The University of North Carolina at Charlotte; Mark W. Lipsey, Claremont Graduate School; and Melvin M. Mark, Pennsylvania State University.

The first author thanks his wife, Wendy Cook Posavac, for reading and commenting on the manuscript and for providing unfailing encouragement through all stages of this text's preparation. He also thanks his son, Steve, for accepting his father's obsession. The second author acknowledges with gratitude the contributions to his personal and professional development made by his parents, the late Elizabeth Carey and the late Raymond G. Carey.



STEPS IN PROGRAM EVALUATION

Contents

Acknowledgments

xiii

PART I: AN OVERVIEW OF PROGRAM EVALUATION 1

1 Program Evaluation and Organizations 3

- Why Program Evaluation Developed 3
 - Quality of Service Not Assumed 3
 - Difficulty in Defining and Measuring Results 4
 - Regulation of Human Services 4
- How Program Evaluation Is Used in Organizations 5
 - Types of Organizations Requiring Evaluations 5
 - Various Kinds of Programs 6
 - Reasons to Evaluate Programs 8
 - What Program Evaluation Is Not 10
 - Common Types of Evaluations 10
- Purpose of Program Evaluation 14
- Trends Making Evaluation Necessary 16
 - The Consumer Movement 16
 - Professional Concern 16
 - Managerial Effectiveness 16
 - Limitations on Resources 17
 - Legislative Mandates 17
- The Roles of Evaluators 18
- Work Settings 18

Consultants Compared to In-House Evaluators	20
Conflicts Between Evaluators and Organizations	22
Getting a Job as an Evaluator	23
Evaluation and Other Activities of an Organization	24
Summary	25
Study Questions	25
Further Reading	25
2 Planning an Evaluation	26
Six Steps in Planning an Evaluation	27
Step 1: Identify Relevant People	27
Step 2: Arrange Preliminary Meetings	28
Step 3: Assess the Evaluability of the Program	31
Step 4: Examine the Literature	33
Step 5: Determine the Methodology	33
Step 6: Present a Written Proposal	38
Potential Sources of Resistance to Program Evaluation	39
Expectations of a "Slam-Bang" Effect	39
Self-Styled Experts in Evaluation	40
Fear That Evaluations Will Inhibit Innovation	40
Fear That the Program Will Be Terminated	40
Fear That Information Will Be Abused	41
Fear That Qualitative Understanding May Be Supplanted	41
Fear That Evaluation Drains Program Resources	42
Fear of Losing Control of the Program	42
Fear That Evaluation has Little Impact	42
Summary	43
Study Questions	44
Further Reading	44
3 Measurement Principles and Tools	45
Sources of Data for Evaluations	45
Program Records	45
Program Participants	46
Staff	48
Significant Others	49
Evaluator Observations	49
Community Indexes	50
Which Sources Should Be Used?	50
Principles in Choosing Assessment Procedures	51
Multiple Variables	51
Important Variables	52
Sensitivity to Change	53
Valid Measures	54
Reliable Measures	55
Nonreactive Measures	57
Cost-Effective Measures	57
Types of Measures	58
Written Surveys Completed by Program Participants	58
Ratings of the Program Participants by Others	59
Interviews	59
Behavioral Observations	60

Achievement Tests	60
Published versus Specially Constructed Instruments	61
Preparing Special Measures	64
Summary	67
Study Questions	67
Further Reading	68

4 Ethical Standards of Conducting Program Evaluations 69

Standards for the Practice of Evaluation	70
Ethical Issues Involved in the Treatment of People	71
Assignment to Program Groups	71
Informed Consent	72
Confidentiality	73
Role Conflicts Facing Evaluators	73
The Scientific Quality of the Evaluation	75
Valid Measurement Instruments	75
Skilled Data Collectors	75
Appropriate Research Design	76
Adequate Descriptions of Programs and Procedures	77
Recognizing the Needs of All Stakeholders	77
Program Managers Are Concerned with Efficiency	78
An Evaluation Can Help the Staff	78
Clients Can Be Served by an Evaluation	78
Community Members as Stakeholders	79
Financial Backers	79
Avoiding Possible Negative Side Effects of Evaluation Procedures	79
Can Someone Be Hurt by Inaccurate Findings?	80
Statistical Type II Errors	81
Pay Attention to Unplanned Program Effects	82
Unexamined Values Held by the Evaluator	82
Summary	83
Study Questions	84
Further Reading	84

PART II: EVALUATION IN PROGRAM PLANNING AND MONITORING 85

5 Evaluation in Program Planning 87

The Evaluator and the Planning Committee	88
What Planning Committees Want to Accomplish	88
Evaluators Can Help Program Planners	88
General Issues to Address in Planning	89
Identifying the Intended Participants	89
Identifying the Needs to Be Met	90
The Program's Resources	91
Expectations for the Program	91

Information Sources for the Assessment of Need	93
Census Data	94
Existing Services and Providers	94
Community Resident Survey	96
A Decision Theory Approach to Planning: The Multi-Attribute Utility (MAUT) Method	98
Basic Purpose and Design of the MAUT Method	98
Ten Steps of the MAUT Method	101
An Application of the MAUT Method	103
Use of the MAUT Model to Guide Evaluation	107
Summary	107
Study Questions	107
Further Reading	108
6 Monitoring the Operation of Programs	109
The Program Audit as a Means of Program Evaluation	110
What to Summarize in an Audit of Services	112
Relevant Information	112
Actual State of Program	112
Providers of Services	113
Program Participants	113
An Approach to Providing Repeated Audits	114
Existing Records of a Program	114
Increasing the Usefulness of Records	115
How Records Can Be Used in Process Evaluations	116
Following the Course of Service	120
A Threatening Use of Process Evaluation	124
Implementing an Information System	128
Avoiding Major Problems in Implementing an Information System	129
Avoid the Duplication of Records	129
Avoid Narrow-Minded Technology	129
Avoid Serving the Needs of Only One Group	129
Summary	130
Study Questions	130
Further Reading	131

PART III: EVALUATION OF OUTCOME 133

7 Nonexperimental Approaches to Outcome Evaluation	135
Single-Group Designs	136
One Set of Observations	136
One Group—Two Sets of Observations	136
Uses of the Pretest-Posttest Design	137
Did the Program Participants Change?	137
Did the Participants Change Enough?	138
Relating Change to Services and Participant Characteristics	139
Threats to Internal Validity	142
Actual but Nonprogram-Related Changes in the Participants	143

Who Was Observed	144
Methods of Obtaining Observations	147
Interactions of These Threats	148
Internal Validity Threats Are Double-Edged Swords	148
Chance and Creativity in Single-Group Designs	149
Usefulness of Single-Group Designs	150
Assessing the Usefulness of Further Evaluations	150
Correlating Improvement with Other Variables	151
Preparing the Facility for Further Evaluation	151
Summary	152
Study Questions	153
Further Reading	153
8 Quasi-Experimental Approaches to Outcome Evaluation	154
Distinguishing Change from Random Fluctuations	155
Time-Series Designs	156
Analysis of Time-Series Designs	160
Observing Other Groups	162
Nonequivalent Control Group Designs	162
Selecting Comparison Groups	163
Weakness of the Design	163
Regression-Discontinuity Design	166
Observing Other Dependent Variables	167
Combining Methods to Increase Internal Validity	168
Time-series and Nonequivalent Comparison Groups	168
The Patch-Up Design	169
Summary	172
Study Questions	172
Further Reading	173
9 Analysis of Causes of Change	174
Experimentation in Program Evaluation	174
The Logic of Experimentation	174
Experimental Designs	176
Objections to Experimentation	178
Don't Experiment on Me!	179
Clinical Prerogatives	179
Time and Effort Involved	180
The Costs of Not Doing an Experiment	180
The Most Desirable Times to Conduct Experiments	181
When a New Program Is Introduced	182
When Resources Are Scarce	184
When There Is Controversy About Program Effectiveness	184
When Change Is Desired	185
Preserving an Experimental Design	186
Precautions Before Data Collection Begins	186
Precautions While the Experiment Is Occurring	187
Summary	189
Study Questions	190
Further Reading	190

PART IV: ADDITIONAL APPROACHES TO PROGRAM EVALUATION 191

10	Analysis of Costs and Outcomes	193
	Cost Analyses 194	
	Variable versus Fixed Costs 194	
	Incremental versus Sunk Costs 194	
	Recurring versus Nonrecurring Costs 195	
	Hidden versus Obvious Costs 195	
	An Example 195	
	Direct versus Indirect Costs 196	
	The Necessity of Examining Costs 196	
	Comparing Outcomes to Costs 197	
	The Essence of Cost-Benefit Analysis 197	
	The Essence of Cost-Effectiveness Analysis 200	
	When Outcome Cannot Be Quantified 200	
	Some Details of Cost Analysis 201	
	Units of Analysis 201	
	Alternative Uses of Investments 202	
	Future Costs and Benefits 203	
	Who Bears the Costs and Who Reaps the Benefits? 204	
	Using Cost-Benefit and Cost-Effectiveness Analyses 205	
	Major Criticisms of Cost Analysis 206	
	Summary 207	
	Study Questions 208	
	Further Reading 208	
11	Assessing the Achievement of Individualized Goals	209
	Developing Outcome Goals 210	
	Problems in Setting Performance Goals 211	
	The Approaches to Be Presented 212	
	Individual Treatment Goals 213	
	Goal Attainment Scaling 213	
	An Example of the Use of GAS 214	
	Management by Objectives 218	
	Preparing MBO Goals 218	
	Institutionalizing MBO Goals 219	
	Peer Review 221	
	Summary 222	
	Study Questions 223	
	Further Reading 223	
12	Qualitative Evaluation of Need, Process, and Outcome	224
	Evaluation Settings Best Served by Qualitative Evaluations 226	
	Admission Procedure with Subjective Decision Rules 227	
	Dissatisfaction with a Library Collection 227	
	A Political Campaign 228	
	Gathering Qualitative Data 229	
	The Central Importance of the Data Gatherer 229	
	Observational Methods 230	
	Interviewing to Obtain Qualitative Data 233	

Plan of Naturalistic Evaluations	235
Establishing Trust in Qualitative Data	237
Criteria of Trustworthy Measurement Procedures	238
Criteria of Trustworthy Evaluation Designs	239
Authenticity of Evaluation	240
Mixing Qualitative and Quantitative Methods	242
The Setting	242
Switching Between Approaches	242
The Evaluation Question	243
Cost of Evaluation	243
Summary	243
Study Questions	244
Further Reading	244

PART V: EFFECTIVE APPLICATION OF FINDINGS 245

10 The Evaluation Report: How to Get It Read	247
The Content of a Report	247
Who Were the Participants	249
How Was the Evaluation Conducted?	250
What Was Observed?	251
What Action Should Be Considered?	254
Report Outline	255
The Style of Reports	257
In-house Reports versus Journal Articles	257
Quantitative versus Qualitative Style	258
The Importance of Attractive Packaging	258
Pitfalls in Reporting	259
Overgeneralization from Inadequate Sampling or Return Rate	259
Conclusions That Are Too Strong for Modest Differences	260
Ignoring the Possibility of Type II Errors	260
Analyzing Specific Rather than Total Effects of a Program	261
Being Honest to the Point of Being Tactless	262
Manner of Distribution	263
Respecting both Sponsors and Program Personnel	263
Personal Presentation to Key Persons	264
Interim Reports	264
Distribution Outside an Organization	264
Feedback to Participants	265
Summary	266
Study Questions	266
Further Reading	267
 14 A Favorable Evaluation Climate:	
How to Encourage Utilization	268
Encouraging Proper Attitude in Staff and Policy Makers	268
Some Innovations Will Not Succeed	268
Innovations Can Be Improved	270
Applications May Be Broader Than Expected	270
Findings Are Working Hypotheses	271

Dealing With Negative Findings Constructively	272
There Are Many Reasons to Continue a Program	272
Common Sense Supports Some Innovations	273
Some Elements of a Program May Be Valuable	274
Dealing With Mixed Results	275
Allowing Others to Interpret Mixed Results	276
The Easy Way Out	276
The Adversary Approach	277
Using the Evaluation to Improve the Program	277
Obstacles to Effective Utilization of Findings	278
Value Conflicts Among Evaluators	278
The Weight of Nonempirical Influences	279
Preexisting Polarization of Interested Parties	280
Misapplied Methodology	281
Metaevaluation	281
An Essay Review	282
Secondary Analysis	283
Quantitative Integration	283
Summary	284
Study Questions	285
Further Reading	285

PART VI: CASE STUDIES 287

Case Study 1: An Evaluation of an Intervention to Change a Work Environment: An Example of a Burn Unit	289
Case Study 2: Responsible Alcohol Use Media Campaign: Can We Get College Students' Attention?	294
Case Study 3: Keeping the Marines in the Field	300
Case Study 4: The Costs and Benefits of Title XX Funding for Family Planning Services in Texas	309
Case Study 5: An Evaluation of a Family Practice Residency Program Using a Naturalistic Inquiry Paradigm	316
Epilogue	321
References	324
Name Index	339
Subject Index	345

part I

AN OVERVIEW OF PROGRAM EVALUATION

Education at Harvard University, job training at Leavenworth Prison, marriage counseling, blood pressure screening, school hot lunch services, the installation of safety equipment, and day care for preschoolers are formally organized ways that help people achieve their goals.

Of a myriad of services and products available in modern industrial societies, the greater proportion of workers and professionals are involved in meeting human needs for health care, education, training, counseling, emotional support, rehabilitation, and so on. The field of human services includes programs and facilities to help people lead more satisfying, healthier, safer, and more productive lives. Many of these needs have traditionally been met through the extended family or through voluntary charity; however, such needs increasingly are being met through organized, publicly funded efforts. These efforts are often called programs, and the methods used to plan, monitor, and improve such programs make up the field of program evaluation. The evaluation methods presented in this book are applicable to both the activities of nonprofit public and private agencies and to profit-making organizations.

Although scientific and engineering problems are important in industrial settings, success in business is often a question of success in managing human resources. Attempts to improve productivity, employee morale, or product quality all depend on good planning techniques, accurate feedback on the impact of the plan adopted, and utilization of that feedback. Furthermore, research on the need for and the effectiveness of a new product requires skills similar to those used by evaluators. Although this

text and the examples it uses reflect the fact that most well-known program evaluations have been in publicly funded organizations, the principles of program evaluation have wide applicability.

The first four chapters of this book provide an orientation to the field and an overview of the methods and goals of program evaluation. In Chapter 1 we describe organizations and the types of programs and services offered. The people responsible for these organizations have many questions about how to plan the programs and judge their effectiveness. The contributions of program evaluation to the various stages of program development and implementation are illustrated.

An overview of the approach evaluators take in preparing to conduct program evaluations is provided in Chapter 2. Often program evaluations succeed or fail due to decisions made during the initial contacts with program managers or sponsors. The principles described apply to all the types of evaluations discussed and illustrated in later chapters.

Because program evaluation is essentially a science of information and its utilization, evaluators must be prepared to assess the quality of the information forming the basis of their work. In Chapter 3 the reader is shown the major ways of gathering information on service programs and how to judge the quality of information obtained. This presentation reflects the need for multiple sources of information in evaluation work; various sources have different strengths and weaknesses. The chapter should expand the reader's appreciation of measurement principles in applied settings.

When research is conducted in organizations serving people's educational, health, employment, emotional, and other needs, special care is exercised to protect the interests of the students, patients, employees, and clients of these organizations. Chapter 4 deals with such ethical issues as well as with the conflicts of interest that can arise when data are collected in organizational settings.

How This Book Is Organized

The chapters of this book are arranged in a systematic fashion that reflects the steps followed in conducting a program evaluation. The flow chart that precedes this introduction provides a preview of what is to come. The reader can keep track of how each chapter fits into the organization of the text and can identify the major aspects of the program evaluator's responsibilities by referring to this figure now and as the course progresses.

We hope this text imparts some of the excitement of participating in the growing, developing field of program evaluation: the field is not fully defined, its techniques are varied, and conducting a program evaluation is challenging. We believe that it will continue to have a positive influence on our institutions and governmental agencies.

1

Program Evaluation and Organizations

Program evaluation is a new and exciting applied social science. A productive society recognizes both the importance of human resources and the importance of providing human services in an efficient manner. Program evaluators provide information about human services in the same way accountants and auditors provide information about financial resources. The mission of program evaluation, or simply evaluation, is to assist in improving the quality of human services. More specifically, program evaluation is a collection of methods, skills, and sensitivities necessary to determine whether a human service is needed and likely to be used, whether it is sufficiently intense to meet the need identified, whether the service is offered as planned, and whether the human service actually does help people in need without undesirable side effects. Through these activities evaluators seek to help improve programs, utilizing concepts from psychology, sociology, administrative and policy sciences, economics, and education. Before discussing specific applications of program evaluation methods, we will describe the field as a whole and illustrate how evaluation and evaluators fit into organizations.

WHY PROGRAM EVALUATION DEVELOPED

Quality of Service Not Assumed

No longer is it assumed that well-meaning individuals or groups who institute a new health, education, training, rehabilitation, or other service actually help people. The rejection of the assumption that all human services help people

receiving the service has made the efforts of many people who teach, counsel, provide medical care, or work with community groups more challenging—but possibly more productive. Today innovative programs as well as expansions of standard services can seldom be funded without some means of demonstrating that the costs of the service are justified by the improved state of the clientele. Although Congress initiated the requirement that services supported by federal grants be evaluated (see, for example, Neigher et al., 1982), the source of a program's funding has little effect on whether the program must be evaluated: It must.

The increasing use of social science methods to improve the effectiveness of human service programs and institutions made evaluation a "growth industry" during the 1970s (Guttentag, 1977). There were several reasons for this growth. Well-meaning, expensive, and ambitious attempts to overcome the effects of disadvantaged backgrounds during the middle and late 1960s were by and large ineffective; at least the impact of these efforts did not measure up to the optimistic expectations held by many program developers, government officials, as well as the general public (Cook and Shadish, 1986). During the 1970s considerable caution was expressed about beginning national programs whose effectiveness had not already been demonstrated. Boruch et al. (1983) show, however, that such warnings were often not heeded.

Besides the need to show that proposed programs will be effective, there are persistent needs to demonstrate that existing programs are worth having and are managed efficiently. Frequently, government programs are not explicitly terminated. Instead, new approaches to problems are implemented alongside of old programs. Some have suggested that programs be authorized for a specific length of time—five years, for example. After the specified period of time expires, the program goes out of existence unless its success has been documented and it has been reauthorized (Chelimsky, 1978). Although federal legislation has not been passed requiring this degree of evaluation of federal programs, this idea illustrates a rigorous use of program evaluation methods to guide legislation.

Difficulty in Defining and Measuring Results

Another reason why the demand for the evaluation of people-oriented services grew is that it is difficult to describe the intended outcomes of human service organizations as compared to those of product-oriented organizations. Whenever the intended outcome is hard to define, assessments of success are very complicated. Industrial firms make things that can be seen, weighed, and counted. Department stores, supermarkets, and bakeries sell merchandise. To evaluate the success of such organizations is easy—at least in theory. To judge success, it is necessary to determine whether the products were in fact made and sold, and whether the amount collected for the products exceeded the cost of making and selling them. In practice, these questions are complex, due to the nature of large businesses; however, the approach to judging the ultimate success of such an organization is easy to define and widely ac-

cepted. Of course, the ways of improving the profitability of a company require one to pay attention to the human resources of the organization.

The use of human resources is just as important as the use of financial resources; however, for many human services, the nature of the "products" is as difficult to describe as are the methods of assessing their quality. What are the goals of a remedial education program? How do we know that a child has obtained what a program was designed to offer? When is an injured person rehabilitated? What level of rehabilitation should be hoped for? If the state provides hot lunches to a school, but the children throw out the vegetables, is the program a failure? Should we change the menu or teach nutrition to the parents? If, after marriage counseling, a couple gets a divorce, was the counseling inadequate?

Regulation of Human Services

If human services were available only on a fee-for-service basis, evaluating them might be left to free-market forces. The services that were purchased would, by definition, be successful; unwanted services would fail because there would be few, if any, buyers. The free-market approach has not been permitted to govern the field of human services. There are two major reasons why this is so. First, it has been assumed that the public cannot easily differentiate between poor and good human service providers. How many patients, for example, can really tell when a physician is competent? For this reason, physicians and many other service providers are required to obtain a state license before providing services to people. Traditionally, qualifications for offering services were based on training; however, a trend is developing that would require a demonstration of skill before licensing.

A second reason why free-market forces do not control human services is that the recipient usually pays for the service indirectly. Private insurance companies, large charitable organizations, local governments, and, increasingly, federal and state governments cover the costs of many human services. Once a service program is developed, the intended recipients have little to say about how it is administered or what is provided. However, it is important that insurance and public funds be used wisely and productively. The need to demonstrate the effectiveness of human services is critical.

HOW PROGRAM EVALUATION IS USED IN ORGANIZATIONS

Types of Organizations Requiring Evaluations

The preceding paragraphs have been very general, and the reader may have had some difficulty knowing just what sort of activities are being discussed. The following is a brief discussion of the types of human services that can and ought to be evaluated.