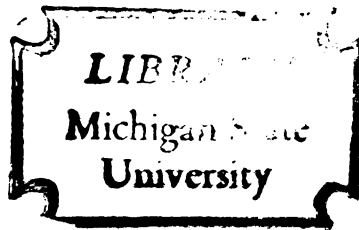


THESIS



This is to certify that the
thesis entitled
THE EVALUATION ENTERPRISE

presented by

MARY STUTZMAN PATRICK

has been accepted towards fulfillment
of the requirements for
Ph.D. degree in Political Science

Charles F. Crundale

Major professor

Date May 9, 1980



OVERDUE FINES:

25¢ per day per item

RETURNING LIBRARY MATERIALS:

Place in book return to remove
charge from circulation records

140
S
\$ 23.00 cc
R-304

THE EVALUATION ENTERPRISE

By

Mary Stutzman Patrick

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Political Science

1980

ABSTRACT

THE EVALUATION ENTERPRISE

By

Mary Stutzman Patrick

The practice and study of program evaluation activities have increased dramatically in recent years. Although program evaluation is maturing as a field of scholarly inquiry, the literature tends to be issue-oriented, to be scattered, and to lack a conceptual base. In an effort to address these perceived shortcomings, the dissertation relies on a rational choice perspective to examine the demand, supply and consumption aspects of the evaluation enterprise as well as integrate these aspects into a proposed model. By making plausible assumptions, constructing simplified models, and developing logical underpinnings, a conceptual approach to the evaluation enterprise offers a number of potential benefits: (1) Organizing an eclectic literature, (2) integrating and explaining past empirical findings, and (3) suggesting new paths for empirical inquiry.

Program evaluation is defined, for dissertation purposes, as a process which produces information assessing program implementation and/or impact. The process used to generate this information is a formal and systematic one--a process based on scientific methods of inquiry. Since the evaluation enterprise is approached from a choice perspective, a model of a program's planning-implementation-assessment cycle is constructed. Three plausible rational choice

assumptions are made concerning the behavior of various actors involved with program choices during stages of the program cycle: An individual is motivated by his self-interest, pursues strategies consistent with this self-interest, and operates under various states of uncertainty with respect to program processes and impact.

Each aspect of the evaluation enterprise is examined from this rational choice perspective in an effort to integrate previous literature and to develop testable propositions for future research endeavors. This analysis proceeds from conditions and motives underlying the demand for and sponsorship of a program evaluation (demand aspect), through the factors affecting the total supply of program assessments with which program evaluation products compete (supply aspect), and to the factors contributing to use of the final program evaluation product (consumption aspect).

In addition to applying a rational choice perspective to each aspect separately, a major task of the dissertation is to construct a model of the evaluation enterprise linking together demand, supply, and consumption variables cited in the literature. The proposed model of the evaluation enterprise identifies the sources of conflicts and constraints on a program evaluation effort and examines their impact on characteristics of the final product which affect utilization. These variables are linked by relying on a rational choice perspective and testable propositions are offered.

While the primary emphasis of this dissertation is conceptual, the utility of the rational choice approach lies not only in terms of the plausibility of its assumptions and ability to generate

propositions, but must also be judged in terms of its testability.

Thus, a research strategy is offered, operationalizations of concepts developed, and results of a pilot study reported.

ACKNOWLEDGMENTS

I would like to thank a number of individuals contributing to my graduate education and this dissertation effort. The dissertation guidance committee was composed of Professor Charles F. Cnudde, Chairperson, Professor Charles W. Ostrom, Jr., and Professor Charles Press. Each provided valuable guidance and support throughout the dissertation effort as well as during my graduate work at Michigan State University.

I am grateful for the assistantships Professor Cnudde provided throughout my graduate work as well as his help in securing a Graduate Office Fellowship that enabled me to conduct the pilot study. More importantly, however, Professor Cnudde's commitment to increasing conceptual as well as methodological rigor enhanced my professional development during graduate school and encouraged me to take a conceptual approach in the dissertation. In addition, I would like to thank him for setting aside his valuable time as departmental chairperson to help shape the dissertation, read and discuss the drafts, and chair the dissertation committee.

Professor Ostrom also played a central role in terms of my dissertation and graduate education. While all members of my dissertation committee generously shared their time and insights with me, I would like to thank Professor Ostrom for the inordinate

amount of time he spent reading and critiquing the arguments made. More importantly, however, Professor Ostrom's thought-provoking course in program evaluation, insightful criticisms, and demanding standards helped me to refine and organize embryonic dissertation ideas in a more thorough and systematic manner.

Professor Press also played a central part during my graduate work and dissertation effort. I am grateful for his continued support and guidance during my coursework, comprehensives, and dissertation writing. His commitment to teaching, professional advice, and help during the writing of this dissertation have contributed greatly to my professional development.

In addition to members of my committee, I would like to acknowledge the help and support of my family. I would like to thank my parents for proofreading the final draft of the dissertation. However, I would especially like to thank my husband, Michael, who always strongly encouraged my educational goals, provided continual intellectual and emotional support during our graduate school years, and has sacrificed some of his own important professional goals so that I can continue to develop professionally.

TABLE OF CONTENTS

Chapter	Page
LIST OF FIGURES	xi
I. PUBLIC PROGRAM EVALUATION: AN OVERVIEW	1
History of Program Evaluation Efforts	3
Federal Evaluation Efforts.	3
State Evaluation Efforts.	6
Evolution of Program Evaluation as An Academic Field of Inquiry.	8
Trends in Solidifying Program Evaluation as A Discipline	8
The Program Evaluation Literature: Topics and Issues.	13
Traditional Concerns in The Evaluation Literature.	13
Defining Program Evaluation.	13
Emerging Topics in The Evaluation Literature . . .	17
Neglected Areas in The Evaluation Literature . . .	19
Purpose and Plan of The Dissertation	22
The Framework.	22
The Evaluation Enterprise: Demand, Supply, and Consumption Aspects	23
The Evaluation Enterprise: A Proposed Model . . .	24

Chapter	Page
II. THE BASIC FRAMEWORK: A RATIONAL CHOICE APPROACH. . .	25
Program Evaluation: A Conceptualization.	28
An Evolutionary Perspective	28
The Method Used	29
Focus on The Program.	30
Timing in The Program Cycle	31
Integrating Conceptualizations of Program Evaluation	33
Program Evaluation as Information	36
The Action Context: Constraints on Basic Research And Setting for Program Choices	38
Constraints on Basic Research	39
A Program's Planning-Implementation- Assessment Cycle	41
The Rational Choice Approach.	44
Behavioral Assumptions.	44
The Uncertainty Assumption.	45
Conclusion.	50
III. THE DEMAND ASPECTS OF THE EVALUATION ENTERPRISE . . .	52
Examining Demand Aspects from A Rational Choice Perspective	54
Motives and States of Uncertainty	54
Dissatisfaction and Puzzlement.	57
Generating Positive or Negative Evidence. . .	58
Organization of The Literature.	59
The Decision to Sponsor A Formal Program Evaluation	59
Altering One's Own Program Uncertainty. . . .	62
Altering Other's Program Uncertainty.	62
Programs Subject to Formal Evaluation Efforts . .	63
Summary	66

Chapter	Page
The Contextual Variables.	68
The Demand Variables.	68
The Evaluation Sponsor.	69
Type of Program Decision Implied.	70
Form of The Demand.	71
The Organizational Variables.	72
Type of Program Agency.	73
Funding-Sponsorship Mix	73
Program Manager Characteristics	74
Conclusion.	76
IV. THE SUPPLY ASPECTS OF THE EVALUATION ENTERPRISE . . .	78
Examining Supply Aspects from A Rational Choice Perspective	81
The State of The Evaluation Literature.	81
Evaluation Methodologies.	81
Evaluator Characteristics	81
The Type of Evaluation Unit	82
The Evaluation Environment.	82
Producers of Program Information.	83
Self-Interest	84
Production Costs.	85
The Decision to Produce Program Information .	87
The Total Supply of Program Information	88
The Nature of Program Information Supplied. . . .	90
Selected Facts.	91
Technical Quality	92
Useability Potential.	92
Summary	94
Towards A Model of The Evaluation Enterprise	
The Supply Variables.	97
Production of Information Variables	97
Resource Variables: Time and Funds	98
Time Allocated for An Evaluation Effort . .	98
Funds Allocated for An Evaluation Effort. .	98

Chapter	Page
Institutional Arrangement Variables:	
Type of Unit and Evaluator Skills	99
Evaluation Unit Structure	100
Evaluator Skills	100
The Evaluation Environment	102
Conclusion	103
V. CONSUMPTION ASPECTS OF THE EVALUATION ENTERPRISE . .	105
State of The Literature	105
Plan of The Chapter	108
Examining The Consumption Aspect from A Rational Choice Perspective	110
The Program Decision Calculus	111
Discounting A Program Evaluation Product	112
Source of Information	113
Interpretive Bias	114
Policy Relevance	114
Reliability and Validity	114
The Implications of The Program Decision Calculus Model	115
Explaining Past Empirical Findings	115
Utilization and Decision Costs	117
Search Costs	118
Processing Costs	119
Summary	120
Towards A Model of The Evaluation Enterprise:	
The Consumption Aspect Variable	122
The Producer of A Program Evaluation Product . .	123
The Nature or Implications of Findings	125
Useability Potential Characteristics	126
Addressing Pertinent Policy and Program Issues	127
Releasing Information When Needed	128
Offering Fiscally Concervative Program Options .	128
Stating Conclusive Findings	129
Format and Recommendations for Action	129

Chapter	Page
Technical Quality Characteristics	130
Reliability	130
Validity.	131
Concern with Technical Quality.	135
Conclusion.	137
VI. A MODEL OF THE EVALUATION ENTERPRISE.	140
The Evaluation Enterprise: A Proposed Model.	141
A Description of The Proposed Model	143
Supply Aspect Variables and Technical Quality and Useability Potential Characteristics.	144
Characteristics Affecting Use	146
Institutional Arrangements: Affects on Technical Quality and Useability Potential Characteristics.	147
Evaluator Skills and Training	148
Type of Evaluation Unit: Organizational Incentives and Sanctions	151
Internal Evaluation Unit.	155
Governmental Evaluation Unit.	156
Entrepreneurial Evaluation Unit	157
Academic Evaluation Unit.	160
Institutional Arrangements and Mixtures of Technical Quality and Useability Potential Characteristics in the Final Product	162
Resource Variables--Constraints on Rational Strategies	166
The Evaluation Environment--Constraints on Rational Strategies	169
Mixtures of Technical Quality and Useability Potential Characteristics.	171
Demand Aspect Linkages to The Allocation of Resources, Choice of An Evaluation Unit, and The Evaluation Environment.	176
Allocation of Resources	176
Funds	176
Time.	180

Chapter	Page
Choice of An Evaluation Unit.	182
The Evaluation Environment.	186
The Evaluation Sponsor.	188
Form of Evaluation Demand	189
Type of Program Decision Implied.	191
Program Funding Mixtures.	192
Type of Program Agency.	193
Program Manager Characteristics	194
Institutional Arrangement Autonomy.	195
Summary	197
Conclusion.	198
VII. THE EVALUATION ENTERPRISE: A PILOT STUDY	200
The Pilot Study	202
The Program--Evaluation Setting and Political Context.	204
The Demand, Supply, and Consumption Aspects of The Evaluation Enterprise	207
Demand Propositions	209
Supply Propositions	209
Consumption Propositions.	210
The Model of The Evaluation Enterprise.	212
Linkages between Supply Variables and Product Characteristics.	212
Allocation of Resources	214
The Choice of An Evaluation Unit.	214
The Evaluation Environment.	215
Assessment of The Measures and Research Approach.	217
The Measures and Survey Instruments	217
The Research Strategy	219
Conclusion.	220
Organizing The Literature	220
Integrating and Explaining Past Empirical Findings	221
New Paths for Research.	221

Chapter	Page
APPENDIX A	224
APPENDIX B	253
APPENDIX C	266
BIBLIOGRAPHY	277

LIST OF FIGURES

Figure		Page
1-1	The Program Evaluation Literature: Traditional Concerns, Emerging Topics, and Neglected Areas.	14
2-1	A Typology of Dimensions for Examining Conceptualizations of Program Evaluation.	28
2-2	The Program Planning-Implementation-Assessment Cycle and The Action Context.	42
2-3	The Evaluation Enterprise: Categories of Variables . .	51
3-1	Typology of Motives	57
3-2	The Evaluation Enterprise: Demand Aspect Variables . .	77
4-1	The Evaluation Enterprise: Supply Aspect Variables . .	104
5-1	A Program Decision Calculus Model	111
5-2	The Discount Element and The Corresponding Program Evaluation Characteristic	122
5-3	Scientific and Evaluation Research Methods: Validity and Reliability Issues	134
5-4	The Evaluation Enterprise: The Consumption Aspect Variable	138
6-1	A Model of The Evaluation Enterprise.	142
6-2	Institutional Arrangement: Tehcnical Quality and Useability Potential Characteristics.	165
6-3	Resource Variables: Technical Quality and Useability Potential Characteristics.	168
6-4	The Evaluation Environment: Technical Quality and Useability Potential Characteristics.	172
6-5	Type of Program Decision Implied: Increasing Uncertainty	178

Figure		Page
7-1	An Inventory of Completed and Planned Program Evaluations in The Department.	203
7-2	The Evaluation Environment	216
C-1	Research Designs: Control of Threats to Internal Validity	272

CHAPTER ONE
PUBLIC PROGRAM EVALUATION:
AN OVERVIEW

The practice and study of public program evaluation activities have increased dramatically in recent years. While perspectives often differ, program evaluation, broadly construed, refers to systematic assessments made about public programs relying on systematic collection of data and quantitative analyses. This Chapter begins by providing a history of program evaluation activities. After emphasizing the increasing importance of formal evaluation efforts to federal and state officials, the focus shifts to the evolution of program evaluation as a field of academic inquiry. Here, trends in solidifying program evaluation as a legitimate field of scholarly inquiry, such as the recent proliferation of texts and journals, are outlined. More importantly, however, the state of the literature concerning public program evaluation is reviewed.

This review is organized in terms of traditional concerns, topics of increasing importance, and neglected areas in the program evaluation literature. Although program evaluation is maturing as a field of scholarly inquiry, the literature tends to be scattered, to be issue-oriented, and to lack a systematic framework. In an effort to address these perceived shortcomings, the dissertation relies on a rational choice approach to examine the demand, supply,

and consumption aspects of the evaluation enterprise and to integrate these aspects into a proposed model. The concluding section of this Chapter outlines the plan for the dissertation.

HISTORY OF PROGRAM EVALUATION ACTIVITIES

Formal efforts to assess public programs based on social science methodology have mushroomed at both the federal and state levels of government over the past decade--signalling a growing commitment to program evaluation activities by practitioners. Federal expenditures for non-defense program evaluation activities reflect phenomenal growth. For example, in 1969 the federal government invested a relatively modest \$24.3 million for program evaluation activities. By 1977, expenditures for non-defense federal evaluations increased tenfold to a staggering \$243 million (Rein & White, 1977; Office of Management and Budget, 1977). No comparable figures for state program evaluation expenditures are available. Yet increasingly, federal legislation mandates formal assessments of public programs conducted by the states. In addition, many states sponsor program evaluations through state legislative mandates or budgetary requirements. The brief history which follows notes the importance of budgeting innovations to the entrenchment of program evaluation activities at both state and federal levels of government.

Federal Evaluation Efforts

Although evaluation activities began to flourish during the mid-1960s, attempts at systematically assessing public programs paralleled development of social science methods during the 1920s and 1930s. During this period, a number of psychologists undertook extra-laboratory studies focusing on such diverse topics as social

influence, work group norms, and leadership characteristics (Freeman, 1975:141). These studies demonstrated the possibilities of applying social science methodology to the study of public programs. Thus, the development of social and psychological measurement techniques made systematic assessment of public programs technically feasible. Much of this instrumental focus, equating program evaluation to the science of instrument development and interpretation, lingers today in the field of program evaluation (Guba, 1972).

In 1935, Stephan advocated applying the "new social research methods" to analyze public programs and policies (Stephan, 1935:518). Writing during the New Deal, he perceived insightfully the possibility of applying the experimental methods developing in the social sciences to the new social program strategies emerging. Besides advocating "social experimentation", Stephan also suggested that the government should conduct social audits and sponsor impact evaluations in order to evaluate the numerous New Deal programs.

Yet, few New Deal programs were evaluated systematically in terms of their impact. The performance of the economy, not the effects of social programs, was the major concern of government officials. Government support of evaluation research languished until the mid-1960s when the Johnson Administration made massive commitments to solve social problems (Wholey et al, 1971:21). In 1964, the Johnson Administration's War on Poverty programs were launched in order to attack chronic poverty problems. Unlike much previous legislation, the Economic Opportunity Act of 1964 contained a demonstration project and research provision for Community Action Programs. As the political controversy grew

concerning various poverty program strategies, the Office of Economic Opportunity sponsored a number of evaluations for such programs as Head Start and the Job Corps. Thus, the War on Poverty legislation and its subsequent uncertainty regarding program effectiveness served as a precedent for writing evaluation requirements into social legislation. In 1974, forty legislative acts contained an evaluation component. Although the evaluation requirement is quite broad in some pieces of legislation, 20 percent of the legislation requiring periodic evaluation stipulates the methods for collecting the data (Rein & White, 1977:242).

Today, many program evaluations are often attached to the budgetary cycle. Budget innovations implemented during the Johnson Administration also helped to increase program evaluation activity at the federal level of government. In 1965, President Lyndon Johnson instituted a planning-programming-budgeting (PPB) system for all federal executive agencies. This budgeting system requires agencies to evaluate program success according to specified criteria when submitting budgetary requests. Subsequent administrations initiated other executive budgetary reforms such as management-by-objectives (MBO) and zero-based budgeting. Like PPB, both of these budgeting approaches include an evaluation of agency and program performance.

Evaluation activities were further entrenched at the federal level of government by Congress. Reacting in part to executive centralization of program funding and management, Congress passed the Legislative Reorganization Act of 1970. Under this Act, Congress ordered five-year cost projections for all new programs,

provided for standardized program data, and directed the General Accounting Office (GAO) to evaluate governmental programs (Newland, 1976). In an effort to tighten further its control over program accountability and effectiveness, Congress authorized congressional committees to conduct evaluations on their own or to require agencies to perform evaluations in 1974.^{1/} Thus, since the mid-1960s, both the executive and legislative branches of the federal government have increasingly demanded and supported program evaluation activities. Evaluation activities at the state level of government have increased as well.

State Evaluation Efforts

Many programs conducted in the states currently receive evaluative scrutiny. The increase of evaluation activity in states is due in part to federal requirements or requests for evaluations of programs funded largely by federal dollars. Yet states, like the federal government, are requiring increasingly formal assessments of the programs which they design and administer. Many states employ budgeting procedures such as PPB or zero-based which require some type of formal assessment of program effectiveness. A relatively new thrust for state evaluation efforts is embodied in sunset legislation adopted by Colorado in 1976 and spreading subsequently

^{1/}The Congressional Budget and Impoundment Act of 1974

to other states. The sunset approach limits by statute the duration of a program and requires a systematic evaluation in order to justify program renewal.

This brief history presented, then, depicts an increasing commitment to program evaluation activities by both federal and state officials. Concomitant with this growth in evaluation activities, scholarly attention to the methods and issues involved with program evaluation has increased as well. An overview of the broad developments concerning program evaluation within the academic community is presented next.

EVOLUTION OF PROGRAM EVALUATION AS AN ACADEMIC FIELD OF INQUIRY

Trends in Solidifying Program Evaluation as A Discipline

Program evaluation is beginning to emerge as a specialized field of inquiry in the social sciences. This has not always been the case. In the academic community, psychologists and sociologists were early advocates of governmental efforts to assess public programs using social science methods and technology; they also tended to dominate the early writings in the program evaluation literature. However, as program evaluation activities have increased, so too have the scholars writing about program evaluation from other disciplines such as political science, education, and the like. While program evaluation is still considered a subfield in many disciplines such as psychology, sociology, education, and political science, there seems to be increasing evidence that program evaluation may be maturing into a discipline itself.

Perhaps the increasing importance of program evaluation within the academic community can best be evidenced by the proliferation of texts, journals, and academic programs devoted to program evaluation. Only one major text outlining the procedures for conducting a program evaluation guided those evaluating the War on Poverty programs during the late 1960s--Edward Suchman's Evaluative Research (1967). An early glimpse into the contextual problems involved while conducting evaluations was provided by the Urban Institute's Federal Evaluation Policy. This monograph analyzed

the War on Poverty evaluation efforts as well as outlined procedures for conducting a program evaluation (Wholey et al, 1970). Like many of the texts which emerged at a later date, these early efforts emphasized the scientific basis of evaluation and discussed the steps necessary for conducting systematic assessments of program implementation and/or impact:

- (1) Identify implementation/program goals;
- (2) develop measures for the goals identified;
- (3) design the research strategy;
- (4) collect systematically data concerning program inputs/and/or effects;
- (5) code the data and analyze the findings;
- (6) present the findings to elected decision-makers and program managers.

In the early 1970s, Carol Weiss published Evaluation Research (1972) and a companion compendium of program evaluation readings (Weiss, 1972b). Unlike previous works published, however, Weiss's works stressed contextual issues involved in program evaluation rather than merely review the scientific steps necessary to assess public programs. Although the number of academic texts and collections of readings guiding evaluation efforts were rather limited in the early 1970s, their numbers have increased dramatically over the last decade (eg., Franklin & Thrasher, 1976; Rutman, 1977; Scioli & Cook, 1975; Rossi et al, 1979; Nachmias, 1979; Guttentag & Struening, 1975; Glass, 1976).

While the basic methodology advocated and the contextual issues examined have changed little from the early evaluation literature, the proliferation of evaluation readers and texts signal the growth of scholarly activities as well as the emergence of program evaluation as a specialized field of inquiry. Yet, emergence of program evaluation as a specialized field is not only reflected by the growth of texts and collections of readings; journals devoted exclusively or in large part to program evaluation issues have appeared as well in recent years.

Professional journals provide a forum for analyzing program evaluation issues and experiences. More importantly, perhaps, professional journals report empirical findings concerning evaluation for peer review. Thus, journals help solidify program evaluation as a legitimate scholarly area. Evaluation Review, a journal devoted exclusively to program evaluation topics, published its first issue in February 1977. Other professional journals providing a forum for program evaluation issues and research, yet not restricted entirely to the field of program evaluation, include Policy Studies, Policy Sciences, Policy Analysis, and Knowledge: Creation, Diffusion, Utilization. Most of these journals commenced publication during the late 1970s.

The development of evaluation courses and programs also signal increasing academic commitment to program evaluation activities. These courses and programs provide training for individuals who aspire to become professional program evaluators. In large part, the growth of academic courses and programs devoted to training program evaluators can be traced to the growing demand for formal

assessments of public programs by government officials. However, like the scholars studying program evaluation activities, the training offered reflects interdisciplinary approaches. Traditional academic disciplines offering specializations and/or courses in program evaluation include psychology, sociology, education, and political science. Yet, training is also offered by professional schools such as medicine, criminal justice, urban planning, social work, and the like. Thus, training for evaluators offered by academic institutions often reflects a rather fragmented approach; course offerings often tend to be dispersed among many disciplines and departments. However, there appears to be increasing interest in establishing separate programs for training evaluators.

The emergence of the Evaluation Research Society in 1978 provides an organizational structure for integrating the multidisciplinary approach found in the academic community. A national professional organization, the Evaluation Research Society sponsors conferences, publishes a newsletter, produces a directory of training available at various institutions, plans to publish a journal, and encourages exchanges among academics, evaluators, and government officials. In addition, this society is in the process of formulating a code of ethics to help guide the conduct of formal program evaluations. The establishment of a professional organization is yet another indicator of the maturing of program evaluation as a field of academic specialization.

The increasing number of texts and empirical studies focussing on program evaluation, the emergence of professional journals and

courses devoted to program evaluation topics, and the formation of a professional organization indicate substantial interest in program evaluation within the academic community. Perhaps the most important contribution the academic community makes to the practice of program evaluation, however, lies in analysis and empirical study of the evaluation enterprise. This literature serves to guide the conduct of evaluations as well as to shape further scholarly inquiry. Yet when turning to the literature, one is often confronted with an eclectic collection of conventional wisdom and sporadic empirical findings.

THE PROGRAM EVALUATION LITERATURE: TOPICS AND ISSUES

The program evaluation literature has evolved in a rather serendipitous manner often focussing on narrow topics and issues. Figure 1-1, displayed on the following page, organizes these scattered topics and issues according to traditional concerns, emerging topics, and neglected issues in the program evaluation literature.

Traditional Concerns in The Evaluation Literature

Traditional concerns in the evaluation literature revolve around (1) defining program evaluation, (2) discussing methods for conducting a program evaluation and the issues involved, and (3) exploring the organizational setting for and the politics of program evaluation.

Defining Program Evaluation. What constitutes legitimate program evaluation activities has occupied a substantial proportion of the evaluation literature. There seems to be general agreement that program evaluation involves systematic assessments of public programs; yet, there is often much effort in the literature expended differentiating among types of evaluations. Some argue that impact studies, assessing program outputs and effectiveness, comprise legitimate program evaluation efforts (eg., Deming, 1975; Riecken, 1972; Suchman, 1972; Weiss, 1972a). Others include process studies, assessing program implementation, within their definitions of program evaluation (eg., Cain & Hollister, 1972; Hatry et al, 1973; Pressman, 1973; Wholey et al, 1971).

TRADITIONAL CONCERNS

Defining Program Evaluation Activities

Methods for Conducting Evaluations & Issues Involved

- Identifying program goals
- Data collection
- Research Design
- Appropriate statistical tools
- Dissemination of findings

Organizational Setting & Politics of Program Evaluation

- Applied vs. basic research controversy
- Basis differences between evaluators and practitioner
- Sources of tension for evaluation activities
- Alternative arrangements for conducting an evaluation

EMERGING TOPICS

Characteristics of The Evaluation Product

- Quality aspects
- Useability aspects

Utilization of Evaluation Products

NEGLECTED AREAS IN EVALUATION

Demand for An Evaluation

Linkages among Aspects of The Evaluation Enterprise

Figure 1-1. The Program Evaluation Literature:
Traditional Concerns, Emerging Topics,
and Neglected area

While general texts outline quantitative methods appropriate for conducting systematic evaluations (eg., C. Weiss, 1972a; Rossi et al, 1979), separate bodies of literature have developed around each step of the program evaluation process:

- (1) Identifying Program Goals (eg., Scriven, 1969; Morris & Fitz-Gibbons, 1978)
- (2) Data Collection Methods & Problems (eg., Mann, 1965; Rutman, 1977; DeNeufville, 1978)
- (3) Research Designs for Evaluation (eg., Rossi et al, 1979; Cook & Campbell, 1979)
- (4) Appropriate Statistical Tools (eg., Bernstein, 1976; Guttentag, 1973; Scioli & Cook, 1975)
- (5) Strategies for Disseminating Findings (Larsen et al, 1972; Davis & Salasin, 1975)

Another area traditionally receiving a large amount of scholarly attention is the organizational setting of and politics involved when conducting a program evaluation. Some argue that evaluation research differs from basic scientific research in terms of the organizational setting (eg., Coleman, 1972; Rossi, 1969; Mann, 1969). Many evaluation scholars have speculated about and analyzed the organizational setting of program evaluation and organizational constraints on applying systematic research principles. They often conclude that the public sector setting of evaluation constrains basic research methodology. Like much of the literature in the field, however, the approach to organizational problems and their solutions is often piecemeal.

There is a group of writings enumerating and discussing the differences between evaluators and practioners and their subsequent impact on the evaluation process (eg., Rodman & Kolodny, 1964; Campbell, 1969; C. Weiss, 1973). In general, this literature

portrays evaluators and program managers as possessing divergent training and orientations about the evaluation enterprise; it is these differences, it is argued, that often generate tension while conducting an evaluation of a public program. Other works focus more generally on the politics and tensions arising during an evaluation effort due to a number of variables such as time constraints, demands made by program advocates, and the like (eg., Carter, 1971; Weiss, 1970).

Another traditional focus within the evaluation literature concerns the type of arrangement chosen to conduct an evaluation: inside or outside the program agency (eg., Suchman, 1972; Weiss, 1972b). Much of this literature suggests that an internal evaluation unit minimizes tension yet raises questions of independent program assessments. Conversely, external evaluation units tend to increase tension yet decrease doubt about the subjectivity of program assessments.

Traditional topics and issues discussed in the evaluation literature tend to focus on defining program evaluation, outlining strategies and problems when conducting evaluations, and reflecting on the organizational setting and subsequent problems generated while conducting evaluations. Much of this literature is speculative in nature--based on conventional wisdom, casual observation, and sporadic case studies. However, the topics and issues receiving increased activity in recent years, the final evaluation product and its utilization, depart from the normative and qualitative tradition characterizing much of the early program evaluation literature.

Emerging Topics in The Evaluation Literature

Although the final product of the formal evaluation process has always received some attention by scholars, in recent years there has been increased empirical activity assessing characteristics and utilization of the evaluation product. Two general areas of inquiry tend to dominate the current writings and research concerning program evaluation: (1) Characteristics of the final evaluation product and (2) utilization of evaluation products. More importantly, however, this recent literature tends to be based on empirical studies rather than mere speculation by evaluation scholars.

With respect to the characteristics of final program evaluation products, two major categories have been identified as important: quality aspects and useability aspects. When speaking of the quality of an evaluation product, scholars refer implicitly or explicitly to its methodological rigor. During the early 1970s, impressionistic judgments concerning the quality of evaluations concluded that evaluation products were unsatisfactory, poor, or mediocre at best (Scriven, 1971; Weiss, 1972c). Empirical research in recent years tends to substantiate these early impressions about poor quality (Bernstein & Freeman, 1975; Weiss & Bucuvalas, 1978; Minnesota Systems Research, Inc., 1975).

The useability aspect of a program evaluation product is another major area of current empirical activity. Although most of the attention has focused on quality aspects, more recent investigations of evaluation products focus on useability aspects such as policy relevance of the findings, style of the evaluation report, and the like (Brown et al, 1978; Weiss & Bucuvalas, 1978).

Besides analyzing characteristics of evaluation products, utilization of the final product has received increasing theoretical and empirical consideration within the literature as well. Early efforts merely speculated about the use of program evaluation findings by government officials (Weiss, 1972c; Downs, 1965; Riecken, 1972). Much of this early literature lamented that evaluation findings seemed to have little observable impact on program decisions; other works recommended strategies to increase utilization. Within the past few years, however, more systematic studies of utilization and the factors affecting use have appeared in the program evaluation literature (eg., Caplan et al, 1975; Weiss & Bucuvalas, 1978; Alkin et al, 1979; Patton, 1978). These empirical studies of utilization cite timing, quality, useability, dissemination practices, threatening findings, evaluator characteristics and user characteristics as important variables affecting the actual use of evaluation products by various government officials.

As the number of formal program evaluation efforts sponsored by federal and state government has grown dramatically during the past decade, so too has the literature analyzing and studying program evaluation. While the early literature was mainly speculative, qualitative, or normative, efforts analyzing the evaluation product and its use ushered in a more empirical approach to analyzing program evaluation issues and topics. Although the amount of scholarly literature devoted to program evaluation seems to be increasing at almost a geometric rate, there are still some neglected topics and serious shortcomings present.

Neglected Areas in The Evaluation Literature

There is a conspicuous lack of attention to some issues and topics in the program evaluation literature: scattered and minimal attention is given to the demand aspects of program evaluation; more notably, however, there is a dearth of models and integrative approaches found in the literature.

While there is some speculation about the motives which underlie the demand for a program evaluation (eg., Weiss, 1975; Downs, 1965; Suchman, 1972; Rein & White, 1977), no scholars tend to devote their sole attention to motives stimulating systematic assessments of public programs. Rather, the issue of motives for sponsoring and conducting evaluations is generally mentioned merely in passing when discussing other program evaluation issues. Likewise, there are few articles devoted to identifying program variables which affect the decision to evaluate a specific public program. Not only is there a lack of literature discussing the demand aspects of the evaluation enterprise, the literature which does explore briefly the motives and rationale for conducting an evaluation or the features of programs giving rise to evaluation demands is often speculative and generally lacks an empirical base. To date, there are no empirical studies that investigate the motives and types of programs that lead to formal evaluation efforts.

A conspicuous lack of an integrative approach and theoretical thinking concerning program evaluation is perhaps the most serious shortcoming in the program evaluation literature. The literature

devoted to program evaluation tends to be scattered and issue-oriented. Few have attempted to link systematically the evaluation enterprise together from the initial demands for evaluative scrutiny, through the conduct of evaluation research, to the utilization of the final evaluation product. Instead, aspects of the evaluation enterprise are generally treated separately: There is a literature which describes methods for implementing evaluation efforts and suggests strategies to overcome methodological problems; a separate body of writings discusses the politics of evaluation; and still another major grouping of the literature focuses on characteristics associated with the final evaluation product--quality and useability potential.

Besides the fragmented nature of the literature, many of the works are merely speculative or consist of empirical findings lacking a theoretical base. Although the need for integrative thinking about program evaluation was acknowledged early (Guba, 1972; Scriven, 1971), theoretical developments have been slow to emerge. Those few developments found tend to parallel the fragmented focus which characterizes much of the evaluation literature. Some models structuring the process for conducting an evaluation have been attempted (Alkin, 1972; Bennett & Lumsdaine, 1975b; Wholey et al, 1975); others focus on the evaluation product and utilization (Cook & Gruder, 1978; Davis and Salasin, 1975).

Lack of integration among issues and topics concerning the evaluation enterprise results in a confusing body of literature for academics training evaluators as well as for practitioners seeking

guidance. Moreover, lack of empirical research based on a larger model contributes to the fragmented focus and lack of explanatory power found in much of the literature. It is this gap in the program evaluation literature to which this dissertation is addressed.

PURPOSE AND PLAN OF THE DISSERTATION

In order for program evaluation to mature as a field of inquiry, it seems that a systematic approach to the evaluation enterprise needs development. By making plausible explicit assumptions, constructing simplified models, and developing logical underpinnings, a systematic approach to the evaluation enterprise offers a number of potential benefits: (1) Organizing an eclectic literature; (2) explaining past empirical findings and conventional wisdom; and (2) suggesting new paths for empirical inquiry.

The central purpose of this dissertation is to examine each aspect of the evaluation enterprise based on a systematic approach and to build a model of the evaluation enterprise linking the demand, supply, and consumption aspects. To this end, a rational choice perspective is taken. The rational choice perspective is one that has been used with some success by political scientists and economists in an effort to build integrative models (eg., Buchanan & Tullock, 1962; Tullock, 1965; Bish, 1971; Riker & Ordeshook, 1973).

The Framework

Chapter Two defines program evaluation, posits a simplified model of a program's planning-implementation-assessment cycle, and presents the rational choice assumptions made. Program evaluation is defined as a process which produces information assessing program implementation and/or impact. The process used to generate

this program information is a formal and a systematic one--a process based on scientific methods of inquiry. Since the evaluation enterprise is approached from a choice perspective, a model of a program's planning-implementation-assessment cycle providing the setting for generating and making program choices is constructed. Three plausible rational choice assumptions are made concerning the behavior of various actors involved with program choices during stages of the program cycle: An individual is motivated by his self-interest, pursues strategies consistent with this self-interest, and operates under various states of uncertainty with respect to program processes and impacts. Chapter Two, then, presents the basic framework that will be applied to various aspects of the evaluation enterprise: Demand, supply, and consumption aspects.

The Evaluation Enterprise: Demand, Supply, and Consumption Aspects

Each aspect of the evaluation enterprise is examined from the rational choice perspective developed in Chapter Two in an effort to integrate previous literature and to develop testable propositions. These Chapters examine the evaluation enterprise from the conditions and motives stimulating demand for program evaluation (Chapter Three), through the variables affecting the supply of program evaluation (Chapter Four), to utilization of the final program evaluation product (Chapter Five). The relevant literature is reviewed in each of these chapters.

In addition to applying a rational choice framework to each aspect separately, a major task of the dissertation is to construct a model of the evaluation enterprise linking together demand, supply, and consumption variables. To this end, major variables associated with each individual aspect are identified and discussed briefly at the ends of Chapters Three, Four, and Five.

The Evaluation Enterprise: A Proposed Model

The full model of the evaluation enterprise, linking the variables identified in the three previous Chapters, is presented and discussed in Chapter Six. This Chapter explores the linkages among the various aspects and develops a number of testable propositions. While the rational choice framework helps to organize and integrate a widely scattered literature, provides insight into the various aspects of the evaluation enterprise, and generates a number of interesting propositions, the utility of this approach must ultimately be tested. Therefore, the results of a pilot study testing the empirical feasibility of a rational choice approach to the evaluation enterprise are reported and discussed in the concluding chapter. The proposed research strategy (Appendix A) and survey instruments (Appendix B) are appended. The concluding chapter (Chapter Seven) also reflects on the utility and future research directions suggested by applying a rational choice framework to the evaluation enterprise.

CHAPTER TWO

THE BASIC FRAMEWORK: A RATIONAL CHOICE APPROACH

As program evaluation activities have proliferated over the past decade, so too has the scholarly literature. While the literature has contributed much to our knowledge about the evaluation enterprise, one of the serious shortcomings in the field is the lack of a systematic approach integrating the various issues and topics discussed. The central purpose of this Chapter is to develop a systematic framework and approach for examining the evaluation enterprise. The framework offered here emphasizes the informational aspects of program evaluation, constructs a model of a program's cycle as a setting for choice, and posits assumptions about individual's behavior based on a rational choice perspective.

This Chapter begins by tracing the evolution of various meanings attached to the term program evaluation. In order to analyze the various conceptualizations offered in the literature and to trace the development of the concept, a typology of dimensions often used to differentiate program evaluation efforts from other activities is constructed. When defining program evaluation, scholars tend to encompass three major dimensions: (1) The method used, (2) the focus on the program, and (3) the timing in the program cycle.

While perspectives concerning what constitutes a program evaluation effort often differ in terms of these dimensions, most scholars tend to agree that a program evaluation provides informational assessments of a public program. Yet, few have explored systematically the implications of program evaluation as information. The perspective taken for dissertation purposes emphasizes and builds upon program evaluation as a process to generate information--information generated by a systematic process based on the tenets of social science methods. Integral to the definition offered is the environment for program evaluation efforts--the action context.

The action context, discussed in the second section, imposes constraints on executing rigorous social science research as well as provides the setting for program choices. After discussing the constraints imposed by the action context, a model of a program's planning-implementation-assessment cycle is constructed and discussed. This program cycle, the setting for program choices, provides a basic framework for examining the demand, supply, and consumption of information produced by a formal program evaluation effort.

By making three plausible rational choice assumptions about individuals, the model of the program cycle becomes dynamic: An individual is motivated by self-interest, pursues strategies consistent with this self-interest, and operates under various states of uncertainty with respect to program processes and outcomes. The specific assumptions made concerning the self-interest that various individuals in the program cycle and its action context

pursue are discussed in the third section. The uncertainty assumption makes information critical in terms of shaping the strategies pursued by various individuals and the dynamics of the evaluation enterprise: There are risks involved when making program choices that may damage an individual's self-interest and the information produced by a program evaluation effort may increase or reduce this risk. That is, when individuals operate under uncertainty, there is a demand, supply, and consumption of program information.

PROGRAM EVALUATION: A CONCEPTUALIZATION

An Evolutionary Perspective

Much of the program evaluation literature is devoted, at least in part, to defining the concept of program evaluation. Often however, there is no clear agreement among scholars concerning the definition of program evaluation. Yet when reviewing various conceptualizations of program evaluation, it seems that scholars tend to encompass three major dimensions: (1) The method used, (2) Focus on the Program, and (3) timing in the program cycle. Figure 2-1 displays these three dimensions and the types of restrictions often employed by scholars when delineating program evaluation activities from other types of processes used to generate assessments of public programs.

METHOD USED	FOCUS ON PROGRAM	TIMING IN PROGRAM CYCLE
Systematic Social Science - Quantitative - Qualitative	Implementation - Process Studies Program Outcomes - Impact Studies Implementation & Outcome - Comprehensive Studies	Formative - Before Program Implementation Summative - After Program Implementation

Figure 2-1. A Typology of Dimensions for Examining Conceptualizations of Program Evaluation

The sections which follow explain further the three dimensions and their components and trace the evolution of various definitions of program evaluation employed in the literature.

The Method Used. Some confusion concerning what activities constitute legitimate program evaluation activities has always been present. Yet, in terms of the prescribed methodology, scholars as well as practioners tend to agree that the program evaluation process relies on a systematic approach--the scientific method. This methodological dimension found in nearly all conceptualizations offered, distinguishes program evaluation activities from other less formal types of program assessments such as casual observation, past experiences, and the like. Instead of collecting opinions about a program in a rather serendipitous manner, program evaluation relies on the scientific model of formulating hypotheses and measures, collecting information in as unbiased a manner as possible, and presenting the results.

When employing the methodological dimension in their conceptualizations, many scholars implicitly or explicitly restrict their definitions further to systematic quantitative methods (eg., Weiss, 1972a; Suchman, 1972; Wholey et al, 1971; Nachmias, 1979). That is, emphasis is on empirical research that proceeds from an experimental or quasi-experimental design, gathers data on important variables, and analyzes the results statistically. It seems that this quantitative approach stressed in most conceptualizations of program evaluation can be traced to the early influence of psychologists and sociologists writing in the field (eg., Stephan, 1935; Suchman, 1967; Weiss, 1972a).

In the early 1970s, only a few scholars included a systematic qualitative component within the methodological dimension when defining program evaluation (Stake, 1967; R. Weiss, & Rein, 1972). Throughout most of the 1970s, definitions of and approaches towards program evaluation were restricted exclusively to quantitative social science methods. Yet, there seems to be new interest in combining both systematic quantitative and qualitative methods when conceptualizing program evaluation (eg., Cook & Reichardt, 1979; Patton, 1980). Qualitative analysis, such as a case study, in-depth interviewing, field observation, and the like, stresses systematic description and appraisal of many non-quantifiable variables in a program. In terms of the methodological dimension, then, it seems that most perspectives in the field stressed and continue to stress systematic quantitative methods when defining program evaluation. Yet there seems to be renewed interest in including systematic qualitative methods as well when defining program evaluation.

Focus on The Program. While most scholars offering definitions or analyses in the field tend to agree broadly in terms of the methodological dimension, differences can be found in definitions and perspectives with respect to the program aspect receiving systematic assessment: the program implementation processes and/or program outcomes. Early in the development of the program evaluation literature, scholars emphasized assessing program outcomes: The scientific method was engaged to link systematically program inputs and program outputs. These "impact studies" investigate

a program's effectiveness in achieving its stated goals (Deming, 1975; Riecken, 1972; Suchman, 1967; Weiss, 1972a). Impact studies resemble closely the classical hypothesis testing and theory building models advocated in methodology texts: Program input variables such as monetary resources and staffing patterns serve as independent variables that in turn affect the dependent variable, program outcomes such as effectiveness and number of clients served.

Although many writing in the early 1970s restricted their definitions of program evaluation in terms of analyzing program outcomes, others argued that studies focusing on program implementation constituted program evaluation as well (Wholey et al, 1971; Cain & Hollister, 1972; Hatry et al, 1973; Pressman, 1975). The term "process studies" was introduced to describe systematic assessments of program implementation. Unlike an impact study, which focuses on the linkages between program inputs and outputs, a process study focuses on program inputs. A process study examines whether or not a program is implemented according to its stated guidelines and often assesses the efficiency of program operations. When referring to a systematic program evaluation encompassing both program implementation and impact, the term "comprehensive study" is often invoked (Bernstein & Freeman, 1975).

Timing in The Program Cycle. The third dimension used in various conceptualizations of program evaluation involves the timing of the systematic assessment in terms of a program's cycle. In 1972, Scriven offered a distinction between formative

and a summative evaluation. A formative study produces an assessment that feeds into the development phases of a public program; summative evaluations, on the other hand, provide assessments that feed into the implementation and outcome phases of a public program.

The distinction made in terms of the timing during a program's cycle leads to emphasis on different program variables and research strategies. To some degree, separate bodies of literature have developed around this summative and formative distinction. Until recently, it seems that many scholars implicitly or explicitly limited their perspectives to summative evaluations: Program evaluations assess the implementation and/or impact of an ongoing program. Thus, variables such as participation by target populations, efficiency of staff utilization, measures of program success, effectiveness of service delivery, and the like become important variables to study in a summative evaluation. Given that summative evaluations generally commence after the introduction of a public program, research strategies are often constrained. Thus, much of the literature proceeding from a summative perspective towards program evaluation enumerates problems and strategies associated with eliminating alternative explanations for evaluation findings.

Based on the formative perspective towards program evaluation, however, a somewhat separate literature has evolved discussing the important variables and research strategies involved during a formal evaluation. Formative evaluations tend to concentrate on variables useful for planning purposes such as identifying

target populations, assessing resources available, and the like. Research techniques advocated within the formative context include developing social indicators (eg., deNeufville, 1975; Bauer, 1966), surveying target populations, and conducting "social experiments" prior to widescale implementation of a public program (eg., Riecken & Boruch, 1974; Rivlin, 1971; Rivlin & Timpayne, 1975). Although the distinction made by Scriven between formative and summative evaluations resulted in the development of two separate bodies of articles, research, and texts, there is some evidence that an effort is being made to encompass both types of evaluation in recent texts (Rossi et al, 1979).

Integrating Conceptualizations of Program Evaluation. The typology of dimensions used when defining program evaluation presented in Figure 2-1 provides some structure when turning to the widely scattered program evaluation literature and attempting to reconcile the various types of program evaluations often enumerated. Although there are often conflicting perspectives in terms of what activities comprise program evaluation in terms of the program focus (process/impact) and the timing in the program's cycle (formative/summative), there seems to be general consensus that program evaluation relies on systematic scientific methods. This consensus concerning the scientific basis of program evaluation is not surprising given the intellectual and historical roots of the field.

In the 1930s, psychologists and sociologists involved in behavioral research recognized possibilities in applying their new methodologies to public programs and policies (Stephan, 1935).

Technological advances in statistics and computer processing coupled with advocacy of program assessments generated by a more objective process turned this possibility into a reality. However, while emphasis was placed almost entirely on quantitative methods during the late 1960s and early 1970s, there appears to be a broadening of the methodological dimension to encompass systematic qualitative methods as well. This new emphasis on descriptive methods may be a response, in part, to increasing doubts expressed by public officials concerning the utility of quantitative data alone for decision purposes (Rein & White, 1977).

In terms of the program focus dimension, the early literature tended to emphasize impact studies more than process studies. Again, the historical roots of program evaluation may account for this: Impact studies resemble the formal hypothesis testing models of psychologists and sociologists involved early in promoting formal evaluation of public programs. That is, ideally, impact studies test linkages between independent program variables and dependent outcome variables as well as account for intervening influences. Yet, as practitioners and scholars in other disciplines engaged in program evaluation and articulated broader needs, the initial clarity in terms of restricting program evaluation to impact studies soon vanished. Now it appears that many practitioners and scholars alike have expanded their perspectives of program evaluation to include both process and impact studies.

The distinction made between formative and summative evaluation has, perhaps, been the most enduring. The timing of an evaluation effort in the program cycle dimension created by Scriven acknowledged

early that different types of broad program choices warranted different types of research strategies. Although there seems to be some treatment of both perspectives in recent texts and articles, the literatures devoted to formative and summative evaluation still tend to be distinct. Here, formative evaluation is often discussed and researched under the rubric of policy and needs analysis rather than program evaluation.

Although conceptualizations often differ in terms of the three dimensions cited, throughout the evolution of the program evaluation concept nearly all scholars and practioners agree that program evaluations provide informational assessments of public programs--whether they are systematic quantitative or qualitative assessments, whether they provide information about program implementation or program impact, or whether they generate information for planning or implementation/outcome program decisions. That is, program evaluation efforts produce information assessing various aspects of public programs which are relevant to program considerations and choices. While some scholars discuss at length this informational aspect of program evaluation (eg., Alkin, 1972, Bennett & Lumsdaine, 1975b; Cohen, 1972; Freeman, 1975), few have explored systematically the implications of program evaluation as information. The approach taken in this dissertation builds upon the informational aspects of the program evaluation process.

Program Evaluation as Information

The perspective taken here views program evaluation as a systematic process for producing information which may influence various program choices and strategies.

Definition 2.1 - Program Evaluation

Program evaluation is a process making some claim or attempt to gather program information according to scientific research principles. The program evaluation process produces information within the action context for potential use by individuals in the action context.

For clarity, the outcome of this systematic process is termed a "program evaluation product".

Definition 2.2 - Program Evaluation Product

A program evaluation product is the package of information supplied by the program evaluation process.

The definitions of program evaluation and a program evaluation product are consistent with many of the perspectives offered in the literature and reviewed earlier. However, unlike many definitions found, the definition of program evaluation presented is not restricted in terms of the three dimensions often used to differentiate program evaluation efforts from other types of information generated for program decisions: (1) Methods used, (2) focus on program, and (3) timing in a program's cycle. Thus, impact studies, formative studies, quantitative or qualitative studies, needs assessments, process studies, and the like are encompassed by this broad definition of program evaluation.

By defining program evaluation broadly as a systematic process that generates information concerning a public program, the dissertation applies to and builds on many of the perspectives, empirical findings, and conventional wisdom found in the evaluation literature. Unlike many previous conceptualizations, however, the perspective taken here emphasizes and builds upon program evaluation products as information--program information generated by a formal and systematic evaluation process. Besides stressing the informational aspects of program evaluation, the action context is another key element in Definition 2.1. This action context, explored in the following section, serves as the environment for conducting program evaluation efforts as well as provides the setting for program choices.

THE ACTION CONTEXT:
CONSTRAINTS ON BASIC RESEARCH &
SETTING FOR PROGRAM CHOICES

The environment for conducting an evaluation effort has always been a topic of concern among practitioners and scholars. Weiss (1972a) used the term "action context" to refer to the highly political setting of program evaluation. The action context is a key component of the definition of program evaluation (Definition 2.1) in two major respects. First, the action context serves to differentiate program evaluation efforts from basic social science research endeavors. Secondly, the action context provides the setting for program choices and decisions--the context for the demand, production and consumption of information bearing on program choices. The action context of program evaluation is defined more formally below.

Definition 2.3 - Action Context

The action context is the environment for planning, implementing and assessing a public program.
--

The action context of a public program is highly political: It is an environment composed of divergent political views characterized by bargaining and compromise among various individuals; it is an environment where time and money resources are often limited.

Constraints on Basic Research. Including the action context in the definition of program evaluation is important in terms of distinguishing program evaluation efforts from more basic scientific research endeavors. One of the more enduring issues in the literature concerns whether or not the program evaluation process is isomorphic to the basic social science research process. Coleman (1972) makes, perhaps, the strongest arguments for separating applied (evaluation) research from more basic (academic) research. He argues that from their philosophical bases to their implementation, applied research and basic research are fairly distinct: Applied research is intended as a guide to program action and bridges the academic discipline and policy concerns while basic research is intended to test and to develop theories within the academic discipline. While some may take exception to the complete polarization of applied and basic research as argued by Coleman, many agree that the organizational settings do indeed differ (eg., Weiss, 1972c; Rossi, 1969; Mann, 1969): The setting for conducting program evaluations often creates a number of obstacles for executing rigorous systematic research that attempts to assess a public program. Thus, it is often argued, program evaluation products may be less valid and reliable than the results obtained from more basic social science research endeavors.

The action context often places political and organizational constraints on gathering program assessments according to the tenets of basic social research. For example, political considerations often prohibit employing one of the most powerful research designs

for assessing a public program--the classical experimental design. Evaluators cannot generally withhold program benefits to potential program clients in order to compare them to clients who receive program benefits. Demands from various individuals in the action context and involved in the program cycle may also constrain the execution of rigorous social science research. These demands may range from diverting evaluation efforts from sensitive program issues to tempering program evaluation findings to fit individual and program biases.

The action context also places resource constraints on program evaluation efforts in terms of funding and time. Perhaps time constraints affect conducting evaluations in a rigorous manner the most. In general, time for producing information is limited severely when compared to much academic research. The demands for program information often are tied to budget hearings, impending program decisions, and the like. Here, program evaluation efforts must often settle for easily obtainable data that may not be reliable and valid measures, rely on less elaborate designs, and/or lack the time to analyze the data thoroughly for alternative explanations.

The constraints imposed by the action context of program evaluation may affect the final validity and reliability of the program evaluation product and, in turn, limit the program inferences that can be drawn from the information. Although many of the same problems are shared in more basic research, the action context tends to exacerbate problems of following sound social science

research procedures when conducting a program evaluation. By including the action context as an element in the definition of program evaluation, program evaluation efforts can be differentiated from basic research efforts and the political context of program evaluation is emphasized. Besides differentiating program evaluation from basic research, the action context is also an important element in terms of providing a setting for program decisions and choices.

A Program's Planning-Implementation-Assessment Cycle

The perspective of program evaluation taken here emphasizes program evaluation products as information--program information generated by a formal systematic process. Thus, it is important to provide an explicit framework for examining the demand, supply, and consumption of this type of program information. For this purpose, a simplified model of a program's cycle is constructed in order to depict program choices upon which program information may have some bearing. Figure 2-2, displayed on the following page, diagrams the simplified model of a program's planning-implementation-assessment cycle and its surrounding action context proposed.

A program's planning-implementation-assessment cycle is initiated when demands for action are made and governmental agents design a program to address these demands. At the beginning of this cycle, alternative plans for action are considered and developed (planning phase). A program, then, is established and implemented to meet these policy objectives (implementation phase).

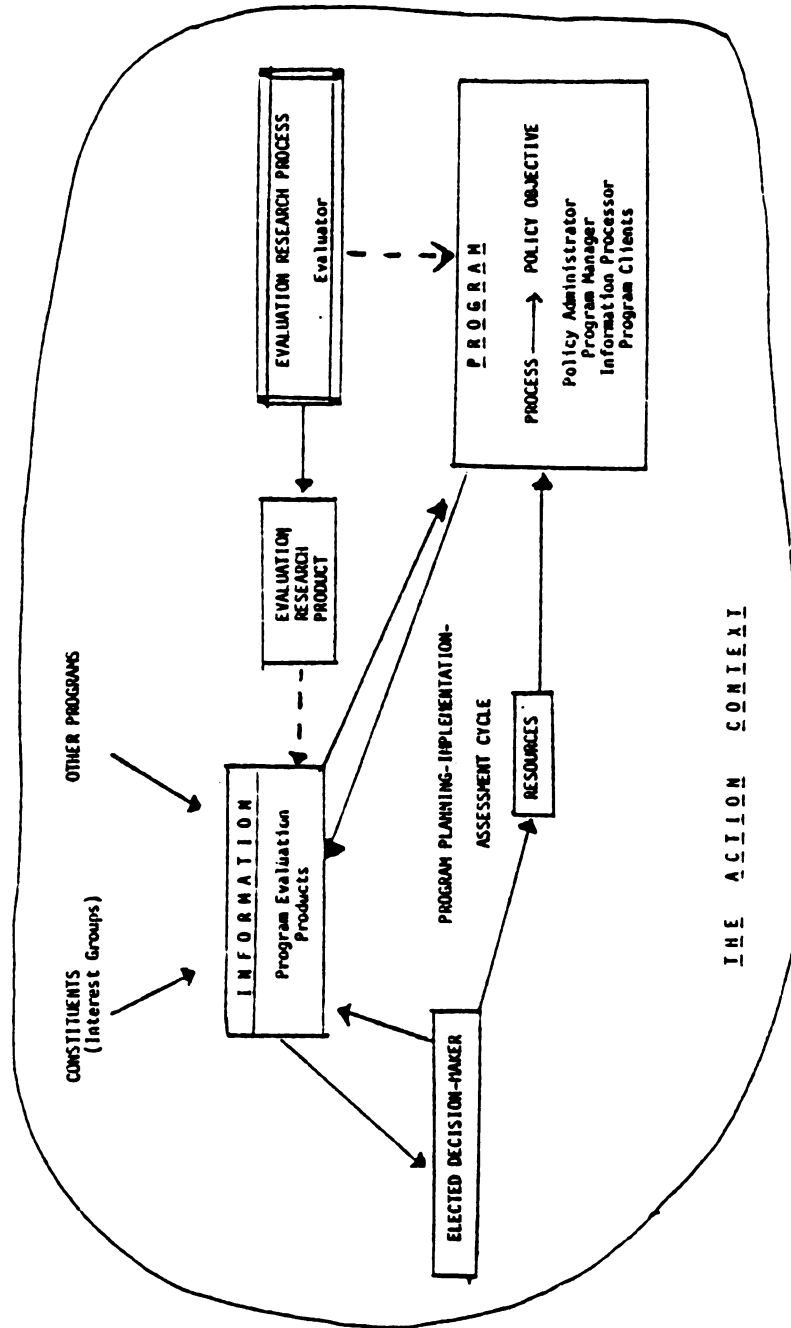


Figure 2-2. The Program Planning-Implementation-Assessment Cycle and The Action Context

The impact of the program in obtaining policy objectives is examined as well as the program strategies and plans (assessment phase). However, not all programs receive formal assessments as indicated by the broken lines linking the program evaluation process to the planning-implementation-assessment cycle. As new policy issues arise, old programs hit new snags, or resources become more scarce, the program cycle returns to the planning phase, continues through the implementation phase, and again repeats the assessment phase. Thus, the program cycle is a continuous one.

Throughout the program cycle, governmental agents confront and create a number of program choices: institute, continue, modify, or terminate a program strategy. For analytical purposes, major types of governmental agents involved directly throughout this program cycle are distinguished: elected decision-makers, policy administrators, and program managers. If any of these individuals have a staff member who summarizes information for him, the individual is termed an "information processor". Information, whether generated by the formal program evaluation process or generated by less formal procedures such as eclectic observations, relying on past experiences and the like, is consumed by these governmental agents when making their program choices or recommendations. The section which follows discusses the assumptions made about the behavior of these individuals.

THE RATIONAL CHOICE APPROACH

By making some plausible rational choice assumptions about individuals' behavior and their state of knowledge, the proposed model of a program's cycle becomes dynamic. In addition, the rational choice perspective gives information in general and program evaluation products in particular a central role in the program cycle. Much of the reasoning presented here is based on Randall Bartlett's Economic Foundations of Political Power (1973). Three key assumptions are made about an individual's behavior in the program cycle and its surrounding action as well as an individual's state of knowledge concerning program strategies and outcomes.

Assumption 2.1 - The Rational Choice Assumptions

- (1) Self Interest--an individual is motivated primarily by his own self-interest.
- (2) Rational Strategies--an individual chooses strategies consistent with this self-interest (goal).
- (3) Uncertainty--an individual does not possess perfect knowledge concerning program outcomes; there is always a degree of uncertainty involved concerning program options or strategies and subsequent program outcomes.

Behavioral Assumptions

The first two assumptions made apply to the behavior of individuals in the action context and program cycle. The first assumption listed posits the motivation of an individual. However,

it is posited further that the goals of individuals found in the program cycle and its action context varies. The definition of each type of individual and the assumptions made concerning self-interest are listed on the following page.

While the goals of each type of actor differs, the second rational choice assumption made is applicable to all individuals. That is, an individual, such as a decision-maker or a program client, chooses strategies that maintain or enhance his given self-interest. Since each actor is motivated by different goals and each attempts to maintain or enhance his self-interest, conflicts often emerge. In turn, these conflicts often spill over to the program evaluation process.

The Uncertainty Assumption

The third rational choice assumption, uncertainty, makes information critical in the program cycle and for the dynamics of the program cycle. It is assumed that only probabilities can be estimated for the consequences of a given program strategy or choice [$0 < (P) < 1$]. Often when employing a rational choice approach, it is assumed that the consequences associated with each strategy or choice are known--the perfect knowledge assumption. If the assumption of perfect knowledge concerning the outcomes of program choices was made, information would not be needed or processed by governmental agents. That is, there would be no demand, supply, or consumption of program information--either program evaluation products or information produced by less formal techniques.

Assumptions 2.2 - Rational Actors in The Action Context

Governmental Agents - individuals who occupy public sector positions.

An Elected Decision-Maker occupies elected office and attempts to maintain/enhance re-election chances.

A Policy-Administrator occupies a top-level appointed position in a bureaucracy and attempts to maintain/enhance the favor of the decision-maker who appoints him.

A Program Manager is a career civil servant whose primary goal is to maintain/enhance his position by promoting program security.

An Information Processor is an individual who summarizes information for an elected decision-maker, policy administrator, or program manager. He attempts to maintain/enhance the favor of his superior.

Producers of Information - individuals who attempt to subsidize the knowledge of governmental agents concerning program strategies.

A Program Client is an individual who receives program benefits whose primary goal is maintaining/enhancing program benefits.

A Constituent is an individual who elects decision-makers and attempts to maintain/enhance benefits derived from tax expenditures.

A Governmental Agent may produce information for other governmental agents.

An Evaluator is an individual who formally assesses program operations and impact. There are two types of evaluators that share the same self-interest: maintaining their jobs within the evaluation unit. However, the secondary goals for these evaluators differs.

A Practitioner is an evaluator who reflects a policy-orientation and possesses policy skills. Besides maintaining his position, he also attempts to maintain/enhance his professional reputation among the policy community.

A Social Scientist is an evaluator who reflects a scientific orientation and possesses methodological skills. Besides maintaining his position, he also attempts to maintain/enhance his reputation among the social science community.

Governmental agents would not demand program information or sponsor evaluation research if perfect knowledge is assumed. Given that the outcomes associated with various program choices are known, a governmental agent would not need additional information or waste scarce resources by sponsoring evaluation research. Consequently, there would be no supply of information under the assumption of perfect knowledge. It would waste efforts of individuals to produce program information since program outcomes are known. Likewise, under the perfect knowledge assumption, there would be no consumption of information by governmental agents. A governmental agent faced with a program choice would know the outcomes associated with each strategy. Under the first two rational choice assumptions, then, a governmental agent would select the course of action which best maintains or enhances his self-interest.

However, by introducing uncertainty, a more plausible assumption than perfect information, into the program choice procedure, information becomes a critical factor: There is now risk involved in making program decisions which may damage an individual's self-interest and information may reduce or increase this risk. ^{1/} When uncertainty is introduced, there is a demand, supply, and consumption of program information which may alter the probabilities associated with various program strategies and outcomes.

^{1/} Bartlett (1973) provides a much fuller discussion on the implications of the uncertainty assumption.

Given the uncertainty assumption, demands for information are made. An individual may demand program information in order to reduce his own uncertainty concerning a program or to alter the probability assessments held by other individuals. Since the outcomes of program are characterized by uncertainty, information which bears on a program is supplied by many individuals, other governmental agents as well as program clients and the like, attempting to influence program choices and strategies. The producers of information, pursuing their self-interest, will supply information which may alter or reinforce the probabilities associated with various program choices and strategies.

There is also consumption of program information by various governmental agents throughout the program cycle when the uncertainty assumption is made. Thus, the program planning-implementation-assessment cycle provides the setting for consuming program evaluation products. Given that a governmental agent attempts to make program choices or recommendations which maintain or enhance his self-interest (under the first two rational choice assumptions made), then a governmental agent may consume program evaluation products in order to reduce uncertainty or reinforce prior calculations concerning program choice alternatives. While some program strategies may have a high level of probability attached to program outcomes, particularly old and established programs, other program strategies, particularly new ones, may have highly uncertain outcomes. However, since the knowledge concerning

program strategy outcomes is probabilistic rather than deterministic, a program evaluation product may alter or reinforce probabilities attached to program strategies.

CONCLUSION

The remainder of the dissertation builds on the informational aspects of program evaluation, uses the model of a program's cycle and its action context as a clarification device, and relies on a rational choice perspective to examine the evaluation enterprise. By approaching the evaluation enterprise based on this framework and explicit assumptions, the often fragmented literature can be integrated in a systematic manner as well as focus research to new areas. In order to proceed systematically, the analysis of the evaluation enterprise is broken into three major aspects: Demand, supply, and consumption aspects.

The three chapters which follow apply this framework and the assumptions posited to the conditions and motives underlying the demand and sponsorship of a program evaluation (demand aspects), to the factors which may affect the total supply of program assessments with which program evaluation products compete (supply aspects), and to the factors which may contribute to use of the final program evaluation product (consumption aspects). Each of these chapters begins with a review of the relevant literature, examines the aspect from a rational choice perspective, and develops testable propositions. In order to construct a model of the evaluation enterprise, the major variables associated with each aspect are identified and discussed in the second section of the chapter. Figure 2-3 displays the linkages among the demand,

supply, and consumption aspect variables that will be identified and discussed in the next three chapters. Chapter Six, continuing the rational choice approach and building on the propositions and models developed, explores the linkages among these variables. This model of the evaluation enterprise provides a framework for identifying, organizing, and analyzing variables which contribute to the characteristics of information produced by an evaluation effort--characteristics which, in turn, may affect the amount of influence a program evaluation product makes in altering individual's uncertainty about program strategies and outcomes.

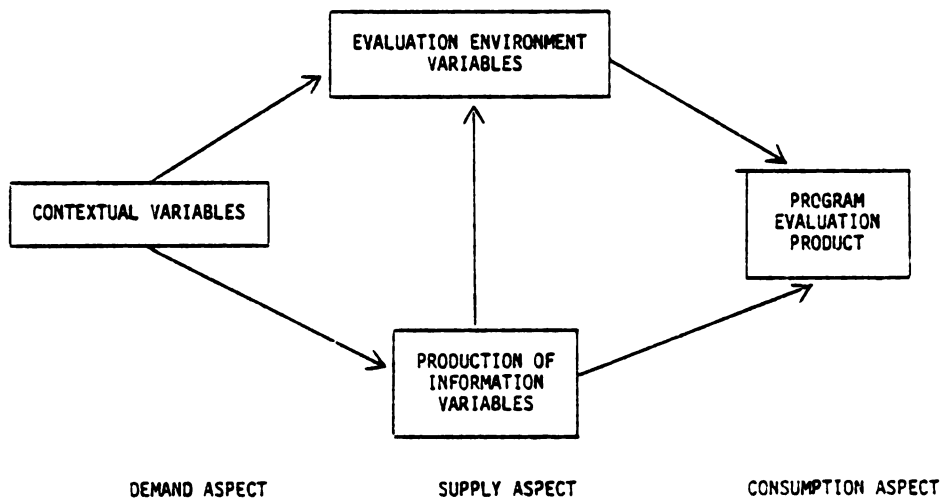


Figure 2-3. The Evaluation Enterprise:
Categories of Variables

CHAPTER THREE

THE DEMAND ASPECTS OF THE EVALUATION ENTERPRISE

The evaluation enterprise begins when demands made for formal assessments of a public program are acted upon and a governmental agent sponsors a program evaluation. Yet unlike many issues and topics in the evaluation literature, there is minimal discussion given to these demand aspects. The first section of this Chapter analyzes the motives and conditions that stimulate the demand for and evaluation as well as affect the decision to sponsor a formal evaluation effort from a rational choice perspective. This section begins by constructing a typology of motives based on a rational choice assumptions. While the literature often enumerates a number of possible motives for demanding and sponsoring an evaluation, this typology organizes motives into two broad categories: This typology is based on the degree of uncertainty an individual holds about a program and the purpose for demanding the evaluation effort.

While many individuals may demand a formal evaluation of a program, not all demands are acted upon by a governmental agent. Here, rational choice reasoning suggests that a governmental agent sponsors a formal program evaluation when the potential benefits in terms of altering uncertainty calculations concerning

a program outweigh production and decision costs detrimental to his self-interest (Proposition 3.1). In addition to analyzing the motives underlying the demand and sponsorship of a program evaluation, the program conditions which stimulate a program evaluation are considered as well. Here, it is suggested that programs that are new and controversial, require relatively large expenditures, experience a crisis, and/or unsuccessfully defended by the program manager tend to stimulate formal evaluations of public programs (Proposition 3.2 and Proposition 3.3).

After examining the motives stimulating the demand for an evaluation, the cost and benefit calculation made by a governmental agent when deciding to sponsor the effort, and the conditions that may affect the decision to evaluate a program, the demand aspect variables in the evaluation enterprise model are identified and discussed.

EXAMINING DEMAND ASPECTS FROM A RATIONAL CHOICE PERSPECTIVE

The demand aspects of the evaluation enterprise is an area often neglected in the program evaluation literature. While some issues and topics, such as appropriate methodology or the political context of the evaluation process, have received repeated discussion, the variables underlying the demand for a formal evaluation have received comparatively little attention in the literature. One area, however, which is the topic of some scattered comments concerns the motives that underlie demands for a formal program evaluation.

Motives and States of Uncertainty

Often, the early program evaluation literature suggested that individuals requested formal assessments of public programs based on scientific methodology "to improve decision-making" (eg., Weiss, 1972c, Hatry et al, 1973). Program evaluation, it was thought, injects relatively unbiased as well as valid information which could help clarify program choices. Suchman, however, hinted at some motives for demanding an evaluation that departed from this early conventional wisdom (Suchman, 1972:84). He suggested that professional groups and public outcry could also stimulate a formal evaluation of a public program. In this case, program evaluations were not necessarily undertaken "to improve decision-making" but rather to pacify disgruntled individuals or to justify abandonment of an unpopular program strategy. More recently,

Rein and White mention yet other reasons for requesting or sponsoring a formal evaluation of a public program. They suggest that government officials request and sponsor evaluations as an instrument of power which serve to contain or to police a public program (Rein & White, 1977). Thus, different motives for requesting a formal evaluation have been suggested in the literature.

However, there has been little effort to integrate these differing motives, from improving program choices to policing a program, in a systematic way. Perhaps Down's approach comes closest as an integrative mechanism when he suggests that government officials seek information to reduce the costliness of making a mistake. This is similar to the perspective taken here: A program evaluation effort provides information and this information is critical to governmental agents who do not have perfect knowledge about program strategies and outcomes. However, Downs, who discussed issues involved with giving economic advice, fails to link rational information seeking to the demands for a formal evaluation (Downs, 1965).

By taking the rational choice perspective outlined in the previous Chapter, the divergent motives for demanding and for sponsoring a formal evaluation of a program can be organized in a more systematic manner. Based on this rational choice framework, it is reasoned that demands for information concerning a public program stem from the self-interest of an individual and his state of uncertainty about program operations and alternatives. That is, an actor in the program cycle or its action context may request a formal evaluation in order to produce information which

may alter uncertainty about a program strategy when he feels that his self-interest may be threatened. Any number of individuals in the program planning-implementation cycle and its action context may request a formal evaluation of a program: An elected decision-maker, policy administrator, program manager, program client, interest group, or constituents. However, sponsors of a formal program evaluation are generally governmental agents. ^{1/}

For analytical purposes, the motives of these various individuals underlying the demand for an evaluation can be grouped into two broad categories: (1) Dissatisfaction and/or puzzlement concerning program strategies and outcomes, and (2) generating positive/negative evidence concerning a program strategy. The degree of individual uncertainty involved seems to be linked to this distinction. Since the outcomes associated with various program choices and strategies cannot be known with absolute certainty, program information can alter the probability assessments made: Some actors may demand information primarily to reduce their own uncertainty (puzzlement/dissatisfaction) or demand information primarily to alter the uncertainty calculations of other actors (generate supporting/negative program evidence). That is, if the individual is highly uncertain about the

^{1/} While it is possible that interest groups can sponsor a formal program evaluation effort, the bulk of formal evaluations seem to be sponsored by governmental agents. This analysis is restricted to government-sponsored program evaluation efforts.

the consequences of alternative program strategies and choices, dissatisfaction or puzzlement motivates the demand for program information. On the other hand, if an individual feels rather certain about program consequences, generating positive (or negative) evidence motivates the demand for program information. Figure 3-1 summarizes this proposed linkage between motive and degree of individual program uncertainty underlying the demand for a formal evaluation effort.

MOTIVES		
LEVEL OF UNCERTAINTY	DISSATISFACTION/ PUZZLEMENT	GENERATING POSITIVE/NEGATIVE EVIDENCE
LEVEL OF UNCERTAINTY (PROGRAM OUTCOMES)	The individual feels that program outcomes are highly uncertain.	The individual feels that program outcomes are fairly certain.
PURPOSE OF DEMAND	Reduce own uncertainty concerning outcomes of program choices	Alter the uncertainty calculations of other individuals making program choices.

Figure 3-1. Typology of Motives

Brief examples of each type of motive follow.

Dissatisfaction and Puzzlement. The following examples demonstrate how dissatisfaction or puzzlement with a program may stimulate demand for information and reduce the level of uncertainty for the individual requesting the information. A number of factors, such as a widely publicized program failure or advocates of a competing program strategy, may foster dissatisfaction with a public program. For example, a newspaper may print a series exposing the lack of placement for Job Corps trainees. A constituent, upon reading this series, feels uncertain about program effects and

doubts the expenditure of his tax dollars for the program. Thus, the constituent demands an evaluation to assess current and alternative job placement strategies.

Puzzlement concerning program outcomes may also stimulate demand for program information. For example, a program may be so new or so complicated that the program effects are highly uncertain. Faced with this situation, an elected decision-maker, concerned with maintaining election support, may demand an evaluation in order to reduce his own uncertainty about program strategies since it is difficult to estimate the probable outcomes of program choices.

Generating Positive or Negative Program Evidence. Unlike the previous examples, the following cases demonstrate demanding information as a means of generating evidence for or against a public program. A program manager, for example, may feel quite certain that the current program strategy he is implementing achieves effectively its policy goal. However, he senses that a decision-maker may entertain doubts about the program and is considering alternative program courses. In this situation, the program manager demands information which substantiates program successes in order to reduce the uncertainty of the elected-decision-maker--not his own uncertainty about the program.

An actor may also demand an evaluation in order to generate negative evidence against a program that he feels is not justified. For example, a program client may feel fairly certain that the program strategy is not succeeding in meeting its policy objectives. In this case, the program client demands information

which may tend to substantiate these claims in order to influence the probability assessments held by other actors.

Organization of The Literature. The typology of motives, displayed in Figure 3-1, serves to integrate the seemingly divergent motives offered in the literature based on the degree of program certainty held by the individual requesting a formal evaluation of a public program. The rationale claiming an evaluation is needed in order to improve decision-making falls within the dissatisfaction/puzzlement category. Here, there is a high degree of uncertainty about the program strategy and an individual seeks information to reduce this uncertainty. Many of the other rationales offered, such as sponsoring evaluations as instruments of power or placating an enraged constituency, fall within the category of generating program evidence. Here, the individual, feeling rather confident about program consequences, requests an evaluation in order to generate information in an attempt to alter the uncertainty calculations of others.

The Decision to Sponsor A Formal Program Evaluation

Not all of the demands for program evaluation are translated into action by a governmental agent. When considering the demands for a formal evaluation, a governmental agent weighs potential costs and benefits of sponsoring such an evaluation. A program evaluation, which makes an attempt to gather program information according to scientific procedures, incurs a number of direct production costs. Formal evaluations may seem to be rather costly endeavors in terms of time and money. For example, executing systematic research

in the action context may entail more time than is allotted in order to bear on a pending program decision. Thus, less formal program assessments, such as collecting opinions about a program from program personnel in a rather unsystematic manner may take less time than a formal evaluation.

There are administrative costs incurred when sponsoring a formal evaluation: Reports must be filed, a program manager or policy administrator must devote some time to managing the evaluation effort, routines of agency personnel are often interrupted, and the like. If agency personnel are deployed to conduct the formal evaluation, opportunity costs can also be incurred.

Monetary costs are involved when sponsoring a program evaluation effort as well. Often, additional personnel must be hired to collect and to analyze the information. In addition, funds must be available for computer processing, printing final evaluation products, and the like. In order to meet these monetary expenses, additional funds often must be allocated or existing budgetary funds shifted from other program commitments.

Time and money costs may be important factors when a governmental agent decides whether or not to translate demands for an evaluation into actual sponsorship. However, the cost in terms of a possible program decision detrimental to a governmental agent's self-interest may be the primary factor affecting his decision to sponsor an evaluation effort. The expected benefit of a formal program evaluation is information that can alter the uncertainty calculations of individuals about a public program. Unlike many

less formal information gathering techniques, it is often perceived that a program evaluation produces objective and accurate information. ^{2/} Therefore, a governmental agent may sponsor an evaluation in an effort to generate information which contributes to maintaining or enhancing his self-interest; without this information, a governmental agent may damage his self-interest by making a detrimental program choice or recommendation.

Proposition 3.1 posits the conditions for sponsoring a formal program evaluation.

Proposition 3.1

A governmental agent sponsors a program evaluation when the potential benefits in terms of altering the uncertainty calculations concerning a program outweigh the production costs of the evaluation and the costs of making a program choice detrimental to his self-interest.

When deciding whether to sponsor an evaluation, a governmental agent may view the potential benefit as information that reduces his own program uncertainty or information that may alter other individual's program uncertainty. A couple of examples help to clarify Proposition 3.1 further.

^{2/} Whether a formal program evaluation actually yields information that is more objective and accurate than program information produced by less formal methods has never been determined empirically. While this is a debatable point, there seems to be the widely held belief that the scientific basis of program evaluation tends to produce relatively unbiased and accurate information.

Altering One's Own Program Uncertainty. In this first example, an elected decision-maker is faced with the choice of expanding the experimental Health Maintenance Organizations (HMOs) to a nationwide health program. However, he is puzzled about the effectiveness of HMOs in delivering services and is highly uncertain about the actual costs involved in implementing this health policy nationally. In this case, the elected decision-maker calculates that if he supports broadening HMOs and they are not cost-effective, his constituents, who feel that the benefits derived from their tax expenditures are minimal, may not support his re-election. Thus, there is a high degree of risk that the elected decision-maker's self-interest, re-election chances, may be damaged. It seems that the benefits derived from sponsoring a formal program evaluation outweigh the actual costs of conducting the evaluation and the potential costs which may be incurred if he makes a wrong choice concerning this health policy: The decision-maker sponsors evaluation research, then, to reduce his own program uncertainty.

Altering Other's Program Uncertainty. This second example demonstrates sponsoring a formal program evaluation in an effort to alter the program uncertainty calculations of other actors in the program cycle and action context. In this case, a policy administrator for the Education Department is recommending that a vocational education program be terminated in the next fiscal year. A program manager, whose self-interest rests with the security of the program, is threatened by this situation: There are substantial costs to the program manager's security implied by the choice to terminate this program. In this situation, the

program manager feels that the vocational education program is fairly successful and believes that the program is needed. Thus, the program manager sponsors a formal evaluation effort in order to generate information which will increase the certainty calculations of the policy administrator that the program is successful: The potential benefits derived from the evaluation product, which the program manager uses to alter the policy administrator's probability assessments, outweigh the costs of conducting the evaluation effort and the costs to the program manager if the program is terminated. While the motives and the considerations involved when deciding to sponsor an evaluation are important demand aspects of evaluation, it is also important to examine the types of programs that tend to receive formal evaluative scrutiny.

Programs Subject to Formal Evaluation Efforts

Characteristics of programs may affect the decision to sponsor a program evaluation. Although there are some remarks made in the literature, few have attempted to speculate about program conditions leading to the demands for and sponsorship of systematic program evaluations. The following two propositions identify characteristics of programs which may affect whether or not a program stimulates demands for an evaluation and a governmental agent's decision to fund such an effort.

Proposition 3.2

A formal program evaluation is not sponsored when:

- (a) Programs are old and established and create little controversy;
- (b) the program does not require large expenditures of money; and/or
- (c) program managers and clients defend successfully the program.

Proposition 3.3

A formal program evaluation is sponsored when:

- (a) The program is new and controversial;
- (b) the program requires large expenditures;
- (c) the program experiences a notable crisis; and/or
- (d) program managers and program clients are unsuccessful in defending the program.

It is hypothesized that one of the major program characteristics affecting the decision to sponsor an evaluation is the age and degree of uncertainty associated with a program. If programs are old and established, there is often a conventional wisdom attached to program outcomes and choices (Weiss, 1972c). That is, there is often widespread agreement (a low degree of uncertainty) and little controversy concerning program operations and effects of an old and established program. Thus, funding an evaluation that substantiates this widespread agreement would waste resources and provide little benefit to a governmental agent. On the other

hand, if a program is relatively new, the consequences associated with program choices throughout the program cycle are often highly unpredictable and controversial. Thus, it is more likely that a governmental agent would sponsor an evaluation endeavor in order to reduce the risks of a program choice detrimental to his self-interest.

The amount of funds appropriated to a program is another factor which may affect the decision to sponsor an evaluation. Often it is the programs which require large expenditures of tax dollars which receive the scrutiny of constituents and interest groups. Thus, a decision-maker, whose re-election chances rest on providing his constituents benefits from their tax dollars, is more likely to sponsor a program evaluation when large amounts of public funds are spent on a program. The benefits derived from the information produced by a program evaluation outweigh the costs of a wrong program choice for those programs which receive relatively large expenditures. However, when program expenditures are relatively small, there seems to be less of an incentive to sponsor an evaluation, all other things being equal.

A program crisis may spur a program evaluation as well, particularly if the crisis is well-publicized. When the program crisis threatens the self-interest of a governmental agent, evaluation research may be sponsored.

The success of efforts by program managers and clients to defend program strategies and operations may also affect the demand for and sponsorship of a program evaluation. If the program

manager and clients supply program information which defends successfully their program interests, then the costs of producing an evaluation would not outweigh the potential benefits of the information produced. On the other hand, if program managers and clients fail to gain support for their program based on less formal techniques for generating information, then evaluation efforts may be sponsored.

Summary

Many demand aspects of program evaluation receive little systematic attention in the literature. Yet, formal program evaluations are not conducted until demands for formal assessments are acted upon by governmental agents. By applying a rational choice perspective, the divergent motives for demanding and for sponsoring a formal evaluation discussed in the literature can be organized in a more systematic manner: An individual's degree of uncertainty about a program may be related to whether the individual seeks program information to reduce his own uncertainty (dissatisfaction/puzzlement) or to reduce the uncertainty of other individuals involved with a program choice (generating positive/negative evidence).

The program evaluation literature provides little guidance with respect to the factors involved when a governmental agent decides to sponsor a program evaluation effort. Likewise, few scholars speculate about program conditions that may lead to formal evaluative scrutiny. In this section, rational choice reasoning was applied to the decision to sponsor an evaluation and program

characteristics affecting this decision identified. Proposition 3.1 directs future research efforts towards examining the costs and benefits of sponsoring an evaluation. Research directed at assessing the relative importance of production costs and decision costs would enhance our understanding concerning sponsorship. Propositions 3.2 and 3.3 suggest research comparing characteristics of programs that become targets of program evaluation efforts. Empirical investigation of program characteristics that increase the chances of a formal evaluation effort would provide insight in an area of demand where little is known.

This section examined one aspect of the evaluation enterprise from a rational choice perspective--the demand aspect. However, besides examining each aspect of the evaluation enterprise from this perspective in an effort to organize the literature and to generate testable propositions, another major task of the dissertation is to identify important variables associated with each aspect. The following section, then, identifies and describes important variables involved with the demand for an evaluation that can be linked with supply and consumption aspect variables in a proposed model of the evaluation enterprise.

THE CONTEXTUAL VARIABLES

The first group of variables identified for purposes of constructing a model of the evaluation enterprise are demand aspect variables. These contextual variables, culled from the literature, influence and provide the setting for conducting a formal program evaluation effort; they capture the demand for a formal evaluation as well as important characteristics of the program under scrutiny. Two variable subsets are identified in this contextual grouping: demand and organizational variables. The demand for a formal evaluation may come from actors in the program cycle or its action context. The sponsor of the evaluation, form of the demand, and the type of program decision implied by the evaluation effort comprise the demand variables. Organizational variables consist of such factors as the type of program agency, program manager characteristics, and the program funding-sponsorship mix.

The Demand Variables

The evaluation process begins with the demand for and sponsorship of a formal program evaluation. These variables capture the beginning of the evaluation enterprise. Three important variables comprise the decision to evaluate a particular public program which, in turn, may constrain the supply aspects of the evaluation enterprise: (1) The sponsor of the evaluation,

(2) the form of the evaluation demand, and (3) the program decision implied by the demand for a formal program evaluation effort.

The Evaluation Sponsor. In order to conduct a formal program evaluation, access to program operations and data is generally needed. That is, evaluators generally need access to clients, agency records, budgets, program personnel and the like in order to assess the program. Thus, governmental agents tend to sponsor a program evaluation since they possess the authority and legitimacy needed to obtain access to public programs. There are a number of potential sponsors of a program evaluation at both the federal and state levels of government--elected decision-makers, policy administrators, and program managers.

Demand Variable - Evaluation Sponsor

A governmental agent generally sponsors and funds a formal program evaluation effort. The sponsor can be either a federal or a state governmental agent--a elected decision-maker, policy administrator, or program manager.

The sponsor of an evaluation--whether a federal or a state governmental agent--is one important demand variable that may affect other aspects of the evaluation enterprise. Whether an elected decision-maker, a policy administrator, or a program manager acts upon demands for an evaluation and sponsors the effort may bear directly on supply aspects such as the evaluation environment and utilization aspects such as the credibility of the final evaluation product.

Type of Program Decision Implied. When an evaluation effort is sponsored, there is often some type of program choice pending or implied. The program issues and questions addressed by a formal evaluation effort may imply different degrees of threat to program operations. Based on discussions found in the literature, five major purposes for undertaking an evaluation can be identified and arranged from least threatening to most threatening for program operations.

Demand Variable - Type of Program Decision Implied

1. Perfunctory Review
2. Modify Existing Program Strategies
3. Improve Program Management
4. Choose among Competing Program Strategies
5. Continue/Terminate The Program Strategy

The purpose of an evaluation may be to review routinely program operations and/or effectiveness. Here, the program evaluation effort may become an annual ritual that began during initial program operation or systematic evaluations performed routinely for budgetary purposes. While puzzlement and a high degree of uncertainty may have motivated the initial sponsorship of a formal evaluation, a governmental agent may sponsor a perfunctory review to reinforce the growing confidence about the program.

On the other hand, a governmental agent may sponsor an evaluation in an effort to modify the existing program strategy or to improve program management. Here, the motives underlying sponsorship

could be either to reduce one's own uncertainty or, feeling rather confident about current operations, to generate information to alter the uncertainty calculations of others.

The decision to sponsor a program evaluation may imply a choice among competing program strategies or imply possible program termination. Once again, the motive underlying sponsorship could be either to reduce one's own uncertainty or to generate information that can be used to alter other governmental agents' uncertainty calculations about the program. These types of decisions implied by an evaluation effort, compare competing program strategies and possible program termination, often pose a severe threat to current operations.

The type of decision implied by an evaluation is an important variable that may affect the environment for conducting an evaluation. Like other demand variables, it is relatively fixed prior to conducting a program evaluation. Another demand variable, that also is related to the amount of tension or cooperation present during an evaluation effort is the form of the demand.

Form of The Demand. Program evaluation efforts may be sponsored in the form of a requirement or a request by a governmental agent.

Demand Variable - Form of The Demand

A governmental agent may require or request a formal program evaluation.

The initial uncertainty or puzzlement about a program may stimulate requiring a periodic evaluation. A governmental agent may sponsor the evaluation by requiring a program through legislative statute. Many state and federal governmental agents may also require systematic evaluations of a program as part of the budgetary cycle (eg., a zero-based budgetary system). Required program evaluations, whether found in legislative enactments or budgetary procedures, are scheduled at rather predictable intervals.

Requested program evaluations, on the other hand, may catch program personnel by surprise. In this case, a governmental agent may sponsor an evaluation due to a publicized program crisis, emerging dissatisfaction or puzzlement with a program, bolstering support for the program, and the like. A requested evaluation, unlike a required one, may disrupt program operations and build tension in the evaluation environment.

The Organizational Variables

A public program provides the site for a program evaluation endeavor. Three types of organizational variables are identified based on a number of discussions found in the literature: (1) Type of program agency, (2) program funding-sponsorship mix, and (3) program manager characteristics. Organizational variables, like the demand variables, often affect the managerial support and commitment given to program evaluation activities; they also provide points for program manager resistance to evaluation efforts.

Type of Program Agency. In order to conduct an evaluation effort successfully, program manager cooperation is often essential since he generally controls access to necessary program data or client/staff interviews. One important variable that may affect this cooperation is the type of agency housing the program and its incentives and sanctions for a program manager's behavior.

Organizational Variable - Type of Program Agency

Operating Agency - concerned primarily with delivering services.

Research Agency - concerned primarily with broad policy issues rather than services.

The distinction between operating and research agencies is based on the empirical work of Bernstein and Freeman (1975). They found that operating agencies, concerned primarily with delivering client services, were often less receptive to evaluation efforts than more research-oriented agencies.

Program Funding-Sponsorship Mix. The amount of congruency between the evaluation sponsor and the source of program funds is another organizational variable that may also affect the amount of program manager cooperation.

Organizational Variable - Program Funding-Sponsorship Mix

Congruent Mix - the sponsor of the evaluation also controls the funding of the program.

Incongruent Mix - the sponsor of the evaluation does not control the funding of the program.

A congruent mix between evaluation sponsorship and program funding occurs when (1) a federal governmental agent sponsors an evaluation of a state program funded primarily with federal monies, or (2) a state governmental agent sponsors an evaluation of a state program funded primarily with state dollars. An incongruent mix between evaluation sponsorship and program funding occurs when a governmental agent sponsors an evaluation of a program funded primarily a different level of government. A program manager is likely to be more responsive to the evaluation effort when the sponsorship and program funding are congruent.

The two organizational variables discussed this far focus on characteristics of a program's organization which may affect the incentives of a program manager to cooperate with or resist evaluation efforts. The third organizational variable considered, program manager characteristics, identifies a trait of a program manager integral to implementing a formal program evaluation.

Program Manager Characteristics. While a program manager attempts to maintain his job and program security, the political situation in which a program manager exists and his orientation towards the program in general may affect his behavior. Campbell (1969) characterizes these orientations as trapped or experimental administrators.

Organizational Variable - Program Manager Characteristics

Trapped Program Manager - a staunch program advocate whose political situation does not allow him to risk failure.

Experimental Program Manager - more committed to the importance of the problem the program addresses than to the certainty of the program solution whose political situation allows him to take risks.

The experimental program manager, who is willing to take program risks, is more likely to be receptive to an evaluation effort than the trapped program manager, who views his security linked directly with the current program operations.

CONCLUSION

The program evaluation process begins when demands are made for formal assessments of a public program and a governmental agent, upon weighing the costs and potential benefits involved, sponsors the evaluation effort. In general, it is hypothesized, programs with fairly uncertain outcomes or require large expenditures of funds tend to be prone to evaluation efforts. Based on discussions found throughout the evaluation literature, a number of contextual variables involved with the demand aspects are identified. Figure 3-2, displayed on the following page, depicts the demand aspect variables in the proposed model of the evaluation enterprise.

The contextual variables, both demand and organizational, provide the impetus and the setting for conducting the program evaluation; they tend to affect the environment and set constraints for implementing the evaluation endeavor. The next Chapter turns to the supply aspect of the evaluation enterprise. Chapter Four first examines the supply aspect from a rational choice perspective and then identifies the important supply aspect variables in order to continue to build a model of the evaluation enterprise.

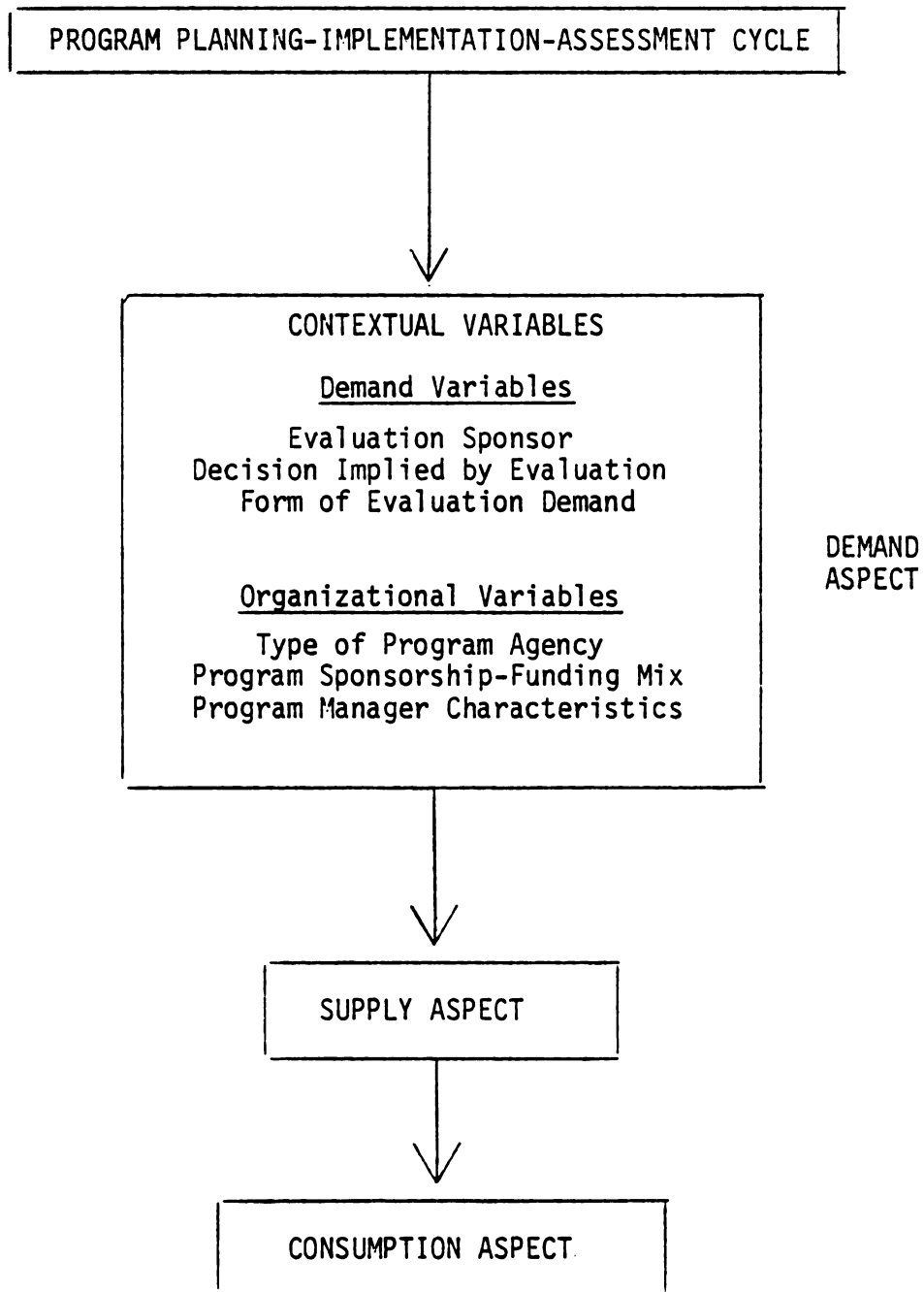


Figure 3-2. The Evaluation Enterprise:
Demand Aspect Variables

CHAPTER FOUR

THE SUPPLY ASPECTS OF THE EVALUATION ENTERPRISE

A number of the traditional issues and concerns discussed in the program evaluation literature comprise what is termed here as the supply aspects of the evaluation enterprise: Evaluation methodologies, evaluator characteristics, the type of evaluation unit, and the evaluation environment. That is, after the decision has been made to sponsor an evaluation, resources are allocated and the program evaluation effort is implemented. While there is a plethora of literature discussing these supply aspects, it tends to ignore the supply of other types of program information that may compete with a program evaluation product in terms of influencing program strategies and choices. Yet it seems important to identify the potential sources and supply of competing packages of information in order to increase our understanding of the use of program evaluation products.

The first section of this Chapter focuses on the potential producers of program information, the conditions which affect the total supply of competing information packages, and characteristics of the information that may affect its amount of influence in terms of altering program uncertainty calculations. Based on the model of the program cycle and its action context, a number of

potential producers of program information can be identified. Many of these individuals have the skills and program familiarity necessary to produce program assessments relying on informal methods such as past experience, casual observation, and the like.

By applying rational choice reasoning, it is posited that these individuals produce information in an effort to influence a program choice when their self-interest is threatened. When making the decision to produce the information, the individual weighs the potential damage to his self-interest as well as the costs involved in generating the information (Proposition 4.1). In turn, the total supply of information generated depends, in part, on the number of individuals whose self-interest may be threatened. Here, situations and conditions that may threaten a number of individuals and generate a relatively large supply of program information include programs with highly uncertain outcomes, a threatening type of program choice pending, and/or the sponsorship of a formal program evaluation effort (Proposition 4.2).

The analysis of potential producers and the total supply of program information generated underscores that program evaluation products must compete with other information packages in terms of altering uncertainty calculations made by governmental agents. Thus, it is posited, a rational producer attempts to produce information which, if acted upon by a governmental agent, will maintain or enhance his self-interest (selected facts). In addition,

however, the rational producer also attempts to make his information package competitive for use by providing accurate information (technical quality) and including policy relevant facts and data (useability potential).

The supply aspect variables in the evaluation enterprise model are presented and discussed in the second section. The variables suggested in the literature are organized into two major groupings: (1) Production of information variables--resources and institutional arrangement and (2) the evaluation environment.

EXAMINING SUPPLY ASPECTS FROM A RATIONAL CHOICE PERSPECTIVE

Unlike the demand aspects of program evaluation, some topics and issues concerning the supply aspects of the evaluation enterprise have received considerable attention. Some of the traditional areas of concern include: (1) Evaluation methodologies, (2) evaluator characteristics, (3) types of evaluation unit, and (4) the evaluation environment.

The State of The Evaluation Literature

Evaluation Methodologies. Many books and articles focus on the appropriate methodologies and statistical issues involved when conducting a formal assessment of a public program. Most of these texts and articles stress quantitative methodologies and problems of application in the action context (eg., Rossi et al, 1979; Bernstein, 1976, Nachmias, 1979). However, as noted in the first chapter, works outlining systematic qualitative strategies and methods for conducting program evaluations are also beginning to appear in the literature (eg., Cook & Reichardt, 1979, Merton, 1979; Patton, 1980).

Evaluator Characteristics. Another area receiving considerable discussion concerns the skills and training necessary for conducting a formal program assessment. Here scholars reflect not only on the statistical skills needed to execute an evaluation but also stress the need for evaluator knowledge of policy and program problems (eg., Ball & Anderson, 1978; Floden & Weiner, 1978; Freeman, 1975).

The Type of Evaluation Unit. A recurring topic in the literature, closely aligned with the issue of evaluator skills, concerns the type of evaluation unit employed to conduct the program evaluation. Here, scholars writing early in the field distinguished between evaluation organizations attached directly to program operations (internal evaluation unit) and evaluation organizations detached from ongoing program concerns (external evaluation unit). They suggest that the type of evaluation unit may have different consequences for the conduct of a formal evaluation: An internal evaluation unit tends to generate less anxiety for a program manager than does an external evaluation unit (Suchman, 1972; Weiss, 1972b).

The Evaluation Environment. Perhaps the supply aspect receiving the most comment and scrutiny by scholars concerns the evaluation environment. Here, a body of literature suggests that differences in orientations and organizational incentives often lead to tension between evaluators and program personnel. In turn, tensions between an evaluator and the program manager create problems when attempting to execute scientific procedures assessing a public program (eg., Aronson & Sherwood, 1967; Weiss & Rein, 1972; Stake, 1967; Weiss, 1973).

Thus, a large amount of the evaluation literature discusses and examines a number of issues involved with conducting a formal evaluation effort: Methodological guidelines for conducting an evaluation, evaluator training and skills necessary for executing an evaluation, different arrangements used when conducting an evaluation, and sources and consequences of tension

during the evaluation effort. Yet, the literature discussing supply aspects seems to pay relatively little attention to other program assessments generated by less formal methodologies. However, as suggested in the section that follows, there are important implications, especially for utilization of program evaluation products, of competing supplies of program assessments.

Producers of Program Information

The methods employed to produce information assessing a program are not limited only to the formal program evaluation process. While a program evaluation relies on systematic scientific principles to gather information,^{1/} there are less formal ways to generate information assessing a program as well. Program assessments may be produced by such methods as relying on past experience, making intuitive judgements about the program, and the like. Using the framework and rational choice assumptions posited earlier, the potential producers of program assessments and their decision to produce information can be examined.

Given that program outcomes associated with various strategies can never be known with absolute certainty, information can potentially alter the program choices of governmental agents. Thus, an individual may produce information bearing on a program strategy or choice in an effort to subsidize the knowledge

^{1/} This definition of program evaluation was developed and discussed in Chapter Two.

of a governmental agent and influence his behavior. A couple of terms need clarification before turning to examine the factors involved with the decision to produce program information.

An individual who generates information in an attempt to subsidize the knowledge of a governmental agent concerning program strategies is termed a "producer" of information. Any of the individuals in the program cycle or its surrounding action context may produce information for possible consumption by governmental agents facing program choices: policy administrators, elected decision-makers, program managers, program clients, constituents, and interest groups. Many of these individuals possess the skills and familiarity with programs necessary to produce information based on their past experience, intuition, or casual observation. Thus, an evaluator, an individual sponsored by a governmental agent who assesses a public program using on scientific methods, is only one of many potential producers of program information. Two primary factors seem to be involved with an individual's decision to produce information: (1) His self-interest and (2) the costs of producing the information.

Self-Interest. It is assumed that individuals in the program cycle and its action context pursue rational strategies that maintain or enhance their self-interest. While the self-interest of different types of individuals varies, ^{2/} all types of

^{2/} See the assumptions listed (Assumption 2.2) in Chapter Two for a listing of the self-interest posited for different types of individuals.

individuals attempt to pursue strategies that do not damage their goals or may help to advance their goals. An individual, such as a program client or an elected decision-maker, is a potential producer when he feels that his self-interest may be threatened by a program strategy or he sees the opportunity to advance his self-interest given a pending program choice.

However, not all program strategies and choices involving an individual's goals lead to the production of program information. Like a governmental agent's decision to sponsor a formal evaluation, a potential producer considers the costs of producing the information along with the potential benefits gained by influencing a program strategy or choice.

Production Costs. When a producer generates information by informal techniques such as past experience, casual observation, and the like, there are a number of monetary and time costs involved. While a producer relying on one of these methods may not need to spend additional funds to collect program data, funds may be needed to package the information (i.e., typing, duplicating, and disseminating a report or a memo). Besides considering the monetary costs involved, a potential producer also weighs the costs of time spent. Although information based on past experiences or intuition may not entail additional time to collect new information, an individual must generally allocate some time for organizing and preparing the information. These costs in terms of time may be increased if the individual needs to make additional observations, such as talking with program

personnel, in order to generate program assessments. Disseminating the information may also consume time for an individual or his staff. Here, an individual may need to testify at hearings, prepare mailing lists, talk informally with other governmental agents, and like in order to distribute the information.

When estimating the monetary and time costs involved with producing information, an individual also considers the costs involved with the lost opportunities. Often, funds and time must be diverted from another activity which also may bear on an individual's self-interest. Thus, an individual must decide if the potential benefits derived from spending resources to produce information offset the potential damage to his goals by neglecting another area. For example, a program manager may be faced with the decision to allocate money and time for producing information or for improving direct client services. In this case, then, the program manager must decide which strategy, producing information or improving services, best maintains or enhances his job and program security.

Given the previous discussion concerning self-interest and production costs, the following proposition suggests the logic for an individual's decision to supply program information.

Proposition 4.1

An individual produces program information when he feels that his self-interest will be affected by a program choice and the potential benefits derived from producing the information outweigh the costs involved in producing the information.

A couple of brief examples illustrate the dynamics of an individual's decision to produce information.

The Decision to Produce Program Information. An individual in the action context may view a pending program choice as detrimental to his self-interest and produce information in an effort to alter the uncertainty calculations of a governmental agent. For example, an elected decision-maker, when considering the annual budget for welfare expenditures, is faced with the choice of continued governmental funding of Medicaid abortions. A number of his constituents form an interest group which opposes government financed abortions. They base their opposition on the use of tax dollars to provide benefits which they deem as undesirable. They feel that their self-interest, maintaining or enhancing the benefits derived from tax expenditures, is involved with the decision to continue or to terminate abortion funding. They determine that the potential benefits derived (influencing the decision-maker to terminate funding) outweigh the costs of producing information. Thus, the interest group collects signatures of other constituents opposed to the abortion program and clips newspaper articles citing management abuses in abortion clinics. The interest group then forwards the information collected to the elected decision-maker.

An individual in the program cycle may supply information as well. For example, the manager of the Medicaid abortion program, who wants to maintain or enhance his program's security, sees his self-interest involved with the pending abortion decision. When deciding whether or not to produce information that bears on this

program choice, he considers the production and opportunity costs involved. In this case, the program manager determines that the potential benefit of supplying information which may lead to program continuation outweighs the production and opportunity costs involved. Thus, the program manager diverts staff resources to prepare a report based on their experiences with the program in an effort to promote program continuation.

The Total Supply of Program Information

The model of the program planning-implementation-assessment cycle and its action context identify a number of potential producers of information. Thus, evaluators are just one of many potential producers of information assessing programs. Since the rational choice perspective concerning the behavior of these individuals suggests that individuals generate information in an attempt to accomplish their goals, the following proposition posits that self-interest plays a critical role in the total supply of information produced assessing a public program.

Proposition 4.2

The supply of program information depends, in part, on the number of actors whose self-interest may be affected by a program choice and the potential benefits outweigh production costs.

The total supply of program information with which a program evaluation product must compete may be affected by the type of program choice pending, the amount of program uncertainty, and the sponsorship of an evaluation effort itself.

Programs with highly uncertain outcomes, such as a program recently implemented, may stimulate many individuals in the program cycle and the action context to supply information. Since a governmental agent, in this case, entertains doubts about a program strategy, individuals may see the opportunity to influence program choices that enhance their self-interest. However, the potential benefit in terms of influencing choices may be less when governmental agents are fairly certain about program strategies and outcomes. Thus, older and established programs may not tend to stimulate a large number of individuals to produce information.

Another condition that may affect the total supply of information is the type of program choice facing a governmental agent(s). Here, one could expect a number of individuals to supply information: program clients, a program manager, an elected decision-maker, and constituents. On the other hand, less threatening types of program choices, such as possible changes in management or minor program adjustments, would probably stimulate fewer individuals to produce information. Here, the potential benefit of influencing a program choice may not justify the costs incurred to produce the information.

Perhaps the sponsorship of a formal program evaluation effort creates another situation that may threaten the self-interest of a number of individuals as well. For example, the conduct of a formal evaluation may signal increasing doubts about a program and individuals, perceiving a threat to their self-interest if program changes are made, may produce information in an attempt to subsidize the knowledge of governmental agents. Individuals may

also be uncertain about the potential findings of an evaluation effort. Thus, in an effort to counter any information generated by a formal program evaluation that may damage his self-interest, an individual may decide to supply information bearing on the program as well.

Situations and conditions that threaten a number of individual's self-interest such as the amount of program uncertainty present, the type of pending program choice, and the sponsorship of a formal program evaluation effort, often lead to a large supply of program information. That is, when there is a high degree of program uncertainty, a threatening type of program choice pending, and/or a formal evaluation effort sponsored, it seems that many individuals in the program cycle and its action context are likely to produce information in an effort to maintain or enhance their goals. Thus, a program evaluation product, a package of information produced by a formal and systematic evaluation process, must generally compete with other packages of information supplied by various individuals in the program cycle and action context. The section which follows explores three major characteristics of the information supplied, whether produced by the formal program evaluation process or by less formal techniques, that may affect its competition with other information supplied for use by a governmental agent.

The Nature of Program Information Supplied

Individuals in the action context and the program cycle, it is reasoned, produce information in an attempt to influence choices made by governmental agents. Thus, when packaging program information, a rational producer includes information which if acted upon by

a governmental agent will preserve or enhance his self-interest (selected facts). Yet, the rational producer is aware that the package of information supplied must also be accurate (technical quality) and relevant to the policy choices (useability potential) in order to influence or alter the uncertainty calculations of a governmental agent. Thus, there are three major characteristics of information produced affecting its potential influence on governmental agents facing program choices: selected facts, technical quality and useability potential.

Selected Facts. One major characteristic scrutinized by a governmental agent when utilizing a package of information is the nature of its findings. The content of the information produced depends, in part, on the self-interest of the individual of producing the information.

Assumption 4.1 - Selected Facts

<p>An individual who produces information tends to select facts which maintain or enhance his self-interest.</p>
--

Given that the self-interest of actors in the program cycle and the action context varies, so do the packages of information in terms of the information they contain. Examples of selected facts were presented in the hypothetical Medicaid abortion choice described earlier: The interest group, deeming continued funding detrimental to the benefits they derive from tax expenditures, tend to select facts that criticize the current abortion program--management abuses, medical complications resulting, and the like;

the program manager, deeming loss of funding as detrimental to his job and program security, tends to select facts that demonstrate program successes--providing services to clients that otherwise could not afford abortions, increases in the number of clients served, and the like. Even though the producer supplies information that reflects his self-interest, the rational producer must also be concerned with the accuracy of this information, even though it may be restricted to one narrow aspect and reflect a partial view of program operations and strategies.

Technical Quality. Another major characteristic affecting the influence a package of information may make on program choices is technical quality. Technical quality refers to the reliability and validity of the information (Bernstein & Freeman, 1975; Weiss & Bucuvalas, 1978; Minnesota Systems Research, Inc., 1975). If the information facilitates sound and accurate inferences concerning a program, then a producer has a better chance of protecting or enhancing his goals. However, when the package of information is not credible in terms of its validity and reliability, often it will be discounted by a governmental agent and not help to alter the probability calculations of a governmental agent. Thus, the producer may minimize his chances of influencing the choice and protecting or enhancing his self-interest.

Useability Potential. The third major characteristic, useability potential, refers to the policy relevance of the information. If a package of information includes policy relevant facts and data, then the potential influence of a producer may be greater on a program choice which affects his goals.

One important element of useability potential concerns information that bears on a pending program choice. If the program choice involves making decisions about the efficiency of the current program strategy, then a rational actor, attempting to influence the decision, will supply information that describes and/or assess program implementation. However, if the program choice involves making decisions about the effectiveness of a current program strategy, then a rational producer will attempt to supply information that describes and/or assess program effectiveness. In some situations, the program choice may involve both the efficiency and effectiveness aspects of a program strategy. Thus, a rational producer may attempt to supply a package of information to a governmental agent which assesses both program efficiency and impact. ^{3/}

Another important useability potential element is the feasibility of the program strategies suggested by the information. Since governmental agents make program choices under conditions of scarcity, a package of information which takes into account budgetary constraints and considerations may possess greater useability potential than one which ignores these considerations. If the program strategies suggested are not feasible in terms of budgetary considerations or policy variables amenable to manipulation

^{3/} In terms of program evaluation products, the literature tends to distinguish among these types of studies as process, impact and comprehensive studies (see Chapter Three).

by governmental agents, then the chances for influencing a program choice may be diminished--a program choice may be made which threatens the goals of the producer.

Given that a governmental agent is often supplied with many packages of information, these three characteristics may affect, in part the relative weight given to the competing information received. Since information is produced in an effort to influence program choices, a rational producer tends to select facts which tend to reflect his self-interest. Yet, given that the potential supply of program information may be quite large, a rational producer attempts to make his information package competitive. Two characteristics affecting the competitiveness of the information are technical quality and useability potential. That is, the rational producer also attempts to provide information that is relatively accurate and is relevant to policy concerns and choices at hand.

Summary

While the supply aspect of the evaluation enterprise is one area of traditional concern generating a plethora of literature, it tends to ignore the supply of program assessments produced by less formal and systematic methods. Yet, there seem to be interesting implications in terms of the competitive environment for use of program evaluation products when the supply of program assessments are analyzed explicitly. By applying rational choice assumptions to individuals in the program cycle and action context, a number of potential producers of information can be identified.

These individuals, it is suggested, produce information when the potential benefits derived from generating information that may influence a program choice outweigh the production costs involved. (Proposition 4.1). In turn, the total supply of program information produced depends, in part, on the number of individuals whose self-interest may be affected (Proposition 4.2). Situations and conditions which may threaten a number of individuals' self-interest tend to generate a large supply of program information--a high degree of program uncertainty present, a threatening type of program choice pending, and/or a formal evaluation effort sponsored. This analysis of potential producers and the total supply of program information generated has important implications often treated unsystematically in the literature: Program evaluation products are not the only source of information bearing on a program choice or strategy; generally, a program evaluation product must compete with program assessments produced by less formal methods for use by governmental agents.

Three important characteristics of information produced that may affect its utilization by governmental agents are identified: selected facts, technical quality, and useability potential. Since information is produced in an attempt to influence program choices, a rational producer tends to select facts which reflect his self-interest. Yet, given that the potential supply of program information may be quite large, a rational producer must also attempt to make his information package competitive. Two major characteristics affecting the competitiveness of the information are technical quality and useability potential. That is, the rational producer

also attempts to produce information that is relatively accurate and is relevant to policy issues and choices at hand.

It seems useful to explore the supply aspects of program assessments in order to identify explicitly the sources and packages of information produced which may compete with program evaluation products. The reasoning and propositions offered point to some interesting new areas of research--types of actors supplying program assessments, conditions affecting the total supply of information produced, and comparative analysis of various information package characteristics affecting utilization.

TOWARDS A MODEL
THE EVALUATION ENTERPRISE:
THE SUPPLY VARIABLES

A major task of this dissertation is to isolate important variables within each aspect in order to construct a model of the evaluation enterprise. As noted in the beginning of this Chapter, the evaluation literature often addresses a number of topics and issues related to the supply aspects of the evaluation literature: Evaluation methodologies, evaluator characteristics, types of evaluation units, and the evaluation environment. From this literature, a number of important variables can be identified. In an effort to construct a model of the evaluation enterprise which links together the demand, supply, and consumption aspects, the supply variables are grouped into two major categories: Production of information variables and the evaluation environment.

Production of Information Variables

Production of information variables generate the final program evaluation product and are grouped into two major categories: Resource and institutional arrangement variables. Time and funding are resource variables while the structure of the evaluation unit and evaluator characteristics comprise the institutional arrangement variables.

Resource Variables: Time and Funds

Program evaluation efforts require expenditures of time and funds in order to produce an evaluation product. The two resources, consumed while implementing a formal program evaluation effort, are allocated by the evaluation sponsor.

Time Allocated for An Evaluation Effort. The evaluation literature often points to time constraints as one of the factors impeding rigorous adherence to scientific methodology when executing a formal evaluation effort (eg., Pressman, 1975; Freeman, 1975; Weiss, 1973; Coleman, 1972). Timing the release of information to coincide with pending program choices is another important factor ultimately affecting the use of the final evaluation product.

Resource Variable - Time

The period allocated for conducting an evaluation effort and completing the final program evaluation product.

While time is often cited as an important variable affecting the choice of research strategies, there has been little empirical work in this area. The conventional wisdom holds that impact studies require a longer allocation of time to complete than do process studies. Some suggest that certain elaborate research strategies, such as a social experiment design, require large expenditures of time as well as funds (Rossi et al, 1979).

Funds Allocated for An Evaluation Effort. The second resource variable, funds allocated to execute a formal evaluation, depends, in part, on the amount a governmental agent is willing and able to

spend in an effort to potentially alter uncertainty calculations concerning program choices.

Resource Variable - Funds Allocated

Funding refers to the amount of money spent to implement a program evaluation effort.

Like the time resource variable, the amount of funds allocated also places some constraints on research strategies: The scope of the evaluation effort, the types of data collection efforts possible, the hiring of personnel to conduct the evaluation, and the like. While the cost of formal evaluation efforts has been of increasing concern to public officials, there are few breakdowns of the actual costs involved for different types of evaluation efforts and the impact of cost constraints on research strategies. Besides the two resource variables, there are important institutional arrangement variables classified as production of informational variables.

Institutional Arrangement Variables: Type of Unit and Evaluator Skills

Institutional arrangement refers to the structure of the evaluation unit and the training and orientations of evaluators who conduct the research. After the decision to evaluate a specific program is made, evaluators operating within an evaluation unit implement the program evaluation effort.

Evaluation Unit Structure. Four types of evaluation units, based on discussions found in the literature, are identified for purposes of constructing a model of the evaluation enterprise.

Institutional Arrangement Variable - Type of Evaluation Unit

Internal Evaluation Unit - located in a governmental agency and attached directly to an operating program.

Government Evaluation Unit - located in a governmental agency and organizationally separate from an operating program.

Entrepreneurial Evaluation Unit - a profit-making firm located outside of government.

Academic Evaluation Unit - located outside of a governmental agency and attached to a university.

The evaluation unit, like any organization, provides sanctions and incentives for the behavior of evaluators. These various types of units differ in terms of structuring reward systems reflecting a program orientation and/or a scientific research orientation. ^{4/}

Evaluator Skills. The evaluator and his skills and orientation is another variable integral to the supply aspects of the evaluation enterprise. Here, the literature cites the need for evaluators with methodological and policy training (eg., Weiss, 1973; Coleman, 1972; Pressman, 1975; Freeman, 1975). For analytical purposes, evaluators with different skills and orientations are dichotomized into "ideal types": Social scientists and practitioners.

^{4/} The various orientations and reward structures posited for each type of evaluation unit are presented more fully in Chapter Six.

Institutional Arrangement Variable- Evaluator Skills

An evaluator is an individual who assesses formally program operations and/or impacts.

A practitioner possesses policy skills and reflects a program/policy orientation.

A social scientist possesses methodological skills and reflects a scientific research orientation.

An evaluator classified as a social scientist develops skills that emphasize the technical quality of the program evaluation effort--developing valid and reliable measures, collecting data in a systematic manner, and the like. In addition, this type of evaluator embraces a number of scientific norms--publication in journals, peer critique and review, and autonomy to formulate and conduct research of interest (Bernstein and Freeman, 1975).

Evaluators termed as practitioners, on the other hand, develop skills that emphasize policy and program considerations when conducting the evaluation effort--releasing information by program deadlines, considering budgetary constraints, offering specific program recommendations, and the like. This type of evaluator follows a different set of norms than social scientists--maintaining client relationships, suggesting innovative policy strategies, meeting program deadlines, and the like.

The institutional arrangement variables, type of evaluation unit and evaluator skills and orientations, are central to producing the final evaluation product. While the choice of an evaluation unit is a function of contextual variables, the reward structure

of the evaluation unit and the evaluator's skills affect two important evaluation product characteristics--technical quality and useability potential.

The Evaluation Environment

The environment is a variable often cited as critical to the successful implementation of a formal program evaluation effort (eg., R. Weiss & Rein, 1972; Weiss, 1975; Rodman & Kolodny, 1964; Ward & Kassebaum, 1972).

The Evaluation Environment

The evaluation environment is a short-hand term for the amount of tension encountered between evaluators and program managers during the conduct of a program evaluation.

This relationship between an evaluator and a program manager may be characterized by varying degrees of cooperation or tension. Tension or cooperation between an evaluator may be determined, in part, by a number of the contextual variables such as the demands made by the sponsor, the type of program decision implied by the evaluation effort, program manager characteristics, and the like. In addition, production of information variables, such as the resource constraints and the amount of autonomy evaluators enjoy from a program manager may also contribute to the milieu of the program evaluation effort. In turn, the amount of tension or cooperation present during an evaluation effort constrains or enhances research strategies which affect the characteristics of the final evaluation product.

CONCLUSION

After the decision has been made to sponsor an evaluation, attention turns to implementing the program evaluation effort. The issues and topics associated with this supply aspect of the evaluation enterprise have been a traditional concern in the program evaluation literature. Based on this literature, variables central to implementing a program evaluation effort can be identified: time and funding resources, the evaluation unit, evaluator skills and orientations, and the evaluation environment. Figure 4-1, depicted on the following page, displays these supply aspect variables and their relationship to demand and consumption aspects in the model of the evaluation enterprise.

The supply variables, both production of information and the evaluation environment, produce the final evaluation product--the major consumption aspect variable. The next Chapter turns to the utilization of this final evaluation product. Chapter Five first examines the consumption aspect of the evaluation enterprise from a rational choice perspective. Following this discussion, the consumption aspect variable, the evaluation product, is explored further.

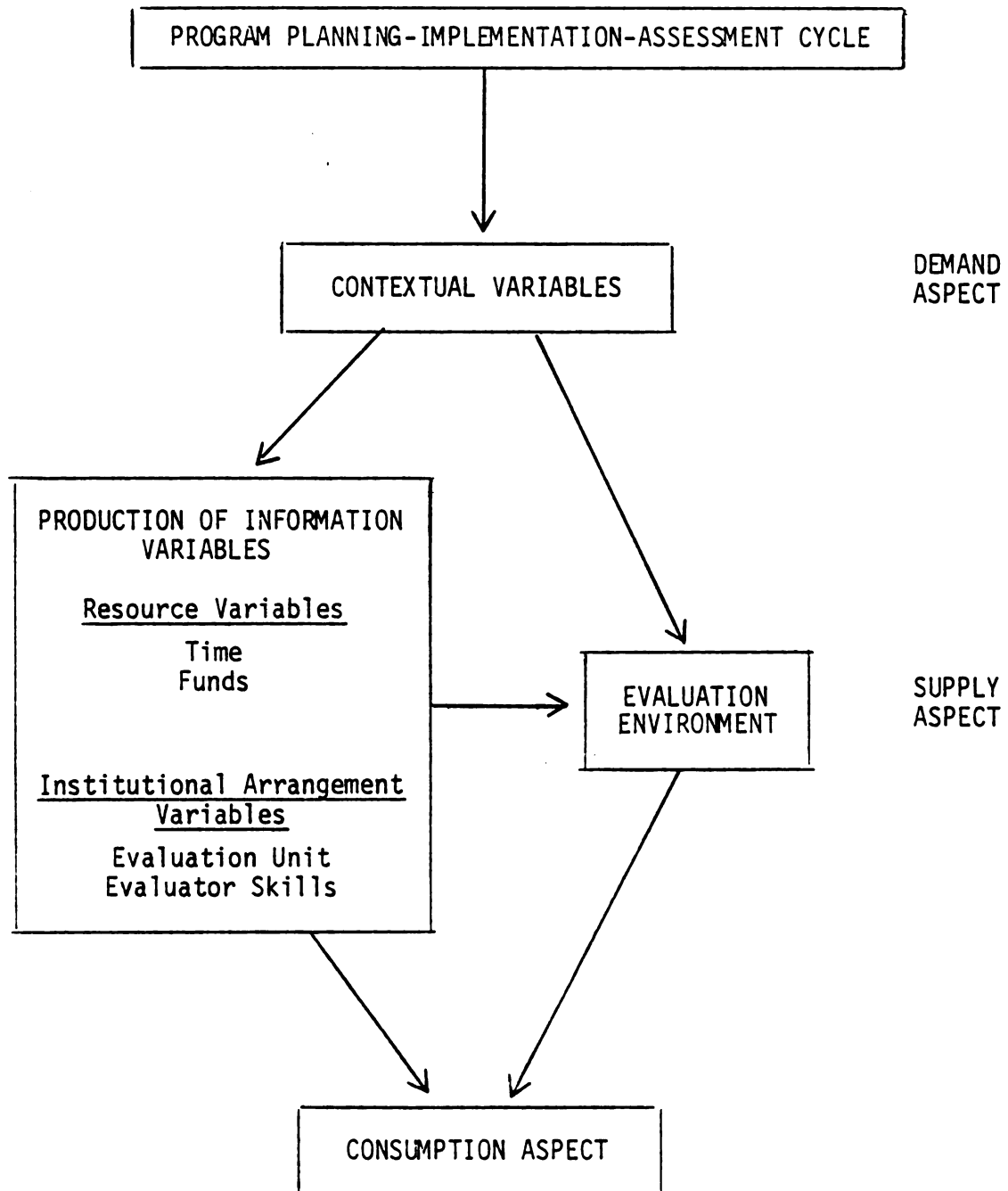


Figure 4-1. The Evaluation Enterprise:
Supply Aspect Variables

CHAPTER FIVE

CONSUMPTION ASPECTS OF THE EVALUATION ENTERPRISE

Utilization of program evaluation products is a topic of increasing interest in the evaluation literature and among government officials. Unlike the demand aspects of the evaluation enterprise which still remains relatively unexamined, scholars have conducted recently a number of empirical studies assessing the utilization of various program evaluation products. While supply aspects of the evaluation enterprise, such as appropriate methodologies and tensions in the evaluation environment have been traditionally the object of scholarly speculation and scrutiny, current activity by evaluation scholars tends to focus on utilization aspects. Much of this recent work focussing on utilization has been empirical in nature yet proceeding in a rather serendipitous manner. Thus, when turning to the literature, one is often confronted with ambiguity and an eclectic collection of empirical findings concerning the use or nonuse of program evaluation products.

The State of The Literature

In general, the concept of utilization is often beset with confusion and change in the program evaluation literature. Much of this confusion and change, it seems, stems from a lack of a well-developed framework. Far too often, utilization remains a

primitive term in the literature. Early scholars who did provide a definition characterized utilization as observable changes made in programs based on an evaluation study (Davis & Salasin, 1975; Weiss, 1972c; Floden & Weiner, 1978). Floden & Weiner (1978) label this perspective as the "decisionistic approach"--program evaluation products are used by elected decision-makers and program managers to make rational changes in programs.

A number of scholars comment on the desirable uses of evaluations by governmental agents: improving program operations; terminating, modifying or changing program operations; and/or setting rational standards for action. A number of secondary uses of evaluation products are often deemed as undesirable by these scholars embracing the decisionistic approach: settling disputes; justifying decisions already reached; deferring a program decision; and/or using evaluation studies as an instrument of power (Freeman, 1975; Suchman, 1972; Weiss, 1972c; Downs, 1965).

While the decisionistic view of utilization and its normative bias tends to dominate the literature, some scholars writing in the early 1970s defined utilization broadly in terms of influence in making changes or consideration of findings by public officials (Riecken, 1972). A growing awareness that few program evaluation products fostered an observable change in a program and empirical studies investigating the role of formal evaluations in program decisions beginning in the mid-1970s contributed to broadening the conceptualization of utilization. Cohen and Garrett (1975), when discussing educational evaluations, suggest that formal

evaluations are influential in changing programs when information criticizing program strategies accumulate; program changes are made only when the weight of a number of program evaluation products suggest changing program operations and strategies.

In 1974, Caplan and his associates interviewed 203 upper-level executives in federal agencies in order to assess the use of social science information in policy decisions. Relying on a broad approach, they defined utilization in terms of consideration by a federal executive; no program changes based on the information needed to be made in order for information to be considered used. They found that social science information, including formal program evaluation efforts, was used like "news" in augmenting less formal information received (Caplan et al, 1975).

Carol Weiss, an influential evaluation scholar, is also shifting to a broader definition of utilization over the past couple of years based on her Columbia University studies. Although she notes that the prevailing concept still stresses application of specific research conclusions to discrete decisional choices (the decisionistic approach), she also feels that utilization must be viewed as sensitizing public officials to findings and information generated by formal program evaluations (Weiss, 1977). She characterizes this conceptualization as the "enlightenment function" of program evaluation products. Thus, Weiss implies that consideration of formal evaluation results, not only detecting observable program changes constitutes another aspect of utilization. Other scholars are also adopting this broader view of utilization in recent years (eg., Patton, 1978; J. Weiss, 1976). Thus, the

perspective towards utilization has been broadened from a restrictive view of detecting observable program changes made based on an evaluation product to include a broader influencing function of formal evaluations.

Much of this shift stems from a number of empirical works (eg., Patton, 1975; Weiss, 1977; Caplan et al, 1975). While Patton attempts to provide an ex post facto framework for this broader definition of utilization based on organizational theorists, Weiss and Caplan offer no theoretical underpinnings for conceptualizing utilization in terms of an enlightenment function. Thus, most efforts to date focus on justifying empirical findings rather than proceeding from a set of assumptions and, in turn, deriving explanations and predictions of a governmental agent's use of a program evaluation product. However, it seems that the rational choice reasoning and assumptions may offer a framework for integrating the often divergent empirical findings and conceptualizations found concerning the utilization of a program evaluation product.

Plan of The Chapter

The first section of this Chapter examines the consumption aspects of the evaluation enterprise from a rational choice perspective. Based on rational choice reasoning, a model of a governmental agent's decision calculus and discounting procedure is proposed. This model suggests that when processing information, a governmental agent considers the source of the information, the policy relevance of the information, the reliability and

validity of the information, and interprets the information with reference to his perceptual biases. After performing this discounting procedure, a governmental agent then estimates the probable outcomes associated with various program strategies and selects that program option that best maintains or enhances his self-interest. A discussion concerning the implications of the program decision calculus model in terms of explaining past empirical findings, decision costs, and processing costs follows.

The second major section of this Chapter turns to the major consumption aspect variable identified in the model of the evaluation enterprise. Here, important characteristics of a program evaluation product, linked to the program decision calculus model, are identified and discussed: The producer, the nature/implications of the findings, useability potential, and technical quality.

EXAMINING THE CONSUMPTION ASPECT
FROM A RATIONAL CHOICE PERSPECTIVE

After a governmental agent acts upon demands for a formal assessment of a program, the resulting evaluation effort produces a package of information which assesses program implementation and/or impact (a program evaluation product). Attention then shifts from the supply aspects of the evaluation enterprise to the consumption aspects. Use of this program evaluation product is approached from the choice perspective developed in Chapter Two. The proposed model of the program planning-implementation-assessment cycle provides the setting for generating and making a number of program choices: institute, continue, modify, or terminate program strategies.^{1/} The three rational choice assumptions made apply to utilization of program evaluation products as well as to the demand and supply aspects of the evaluation enterprise. Thus, the outcomes associated with alternative choices and strategies are characterized by various degrees of uncertainty. Given that a governmental agent attempts to make program choices and recommendations that maintain or enhance his self-interest, a governmental agent may consume a

^{1/} While actors in the action context can also utilize the information produced, this Chapter focuses more specifically on choices and recommendations facing governmental agents in the program cycle.

program evaluation product in order to reduce uncertainty or to reinforce prior calculations concerning various program choice alternatives. To discern further the role a program evaluation product may play in these program choices or recommendations, a model of a governmental agent's decision calculus is proposed.

The Program Decision Calculus Model

The proposed model of a governmental agent's program decision calculus relies on the three rational choice assumptions made throughout the dissertation: Self-interest, rational strategies, and uncertainty. A governmental agent's decision calculus, displayed in figure 5-1, consists of two steps: Estimating the probable outcomes associated with various program strategies and selecting a program choice or strategy.

$p(\text{OUTCOME}) = f [\text{Information, Discount Factor}]$ $\text{PROGRAM CHOICE} = f [p(\text{OUTCOME}), \text{Self-Interest}]$

Figure 5-1. A Program Decision Calculus Model

In brief, the program decision calculus of a governmental agent operates as follows. When faced with a program choice or making a program recommendation, a governmental agent uses program evaluation products (or information produced by less formal methods) to estimate the probable outcomes associated with different program strategies. While processing this information, a governmental applies a discount factor. Upon creating a matrix

of probable outcomes associated with different program strategies, such as instituting a new strategy or continuing the current strategy, a governmental agent then selects the program strategy which best maintains or enhances his self-interest. Before turning to a discussion of the discounting procedure, the implication of the program decision calculus model concerning use of program evaluation products needs emphasis: If information generated by the program evaluation process enters a governmental agent's decision calculus, then a program evaluation product is considered used.

Definition 5.1 - Utilization

Utilization of an evaluation product is defined as information generated by the formal program evaluation process entering a governmental agent's decision calculus.

- (a) Decisionistic - observable changes or recommendations made based on evaluation research findings.
- (c) Enlightenment - thinking or ideas affected by evaluation research findings.

Discounting A Program Evaluation Product

Integral to the proposed program decision calculus model is the discount factor. ^{2/} The uncertainty assumption implies that a program evaluation product (or information produced by less

^{2/} The concept of a discount factor was introduced by Bartlett (1973). However, the discount factor presented here extends the concept to include additional elements.

formal methods) does not relay "perfect knowledge" about the outcomes associated with various program strategies and options. Thus, a governmental agent must scrutinize characteristics of information and weigh the information accordingly. During his decision calculus, then, a governmental agent applies a discounting procedure to characteristics of a program evaluation product. Four elements comprise the discount factor: The source of the information, the interpretive bias of the governmental agent, the policy relevance of the information given the choices pending, and the reliability and validity of the information.

Definition 5.2 - Discount Factor

<p>DISCOUNT FACTOR = f [Source of Information, Interpretive Bias, Policy Relevance, Reliability & Validity]</p>

A brief discussion of each element follows.

Source of Information. A governmental agent may increase or decrease his probable outcome assessment of a program strategy based on the source producing the information. A governmental agent may weigh his past experience with the information producer or consider the reputation of the producer when processing the information. Since the rational choice assumptions apply to all individuals in the action context, producers of information tend to select facts which may enhance their self-interest. Hence, a

governmental agent attempts to detect the self-interest of the producer and attaches probabilities to program choices accordingly.

Interpretive Bias. The interpretive bias of a governmental agent is another element in the discount factor. A governmental agent holds a number of biases and attitudes when processing program information. If the information is counter-intuitive or contrary to a governmental agent's beliefs, then the information may be highly discounted. On the other hand, information which reinforces his beliefs and attitudes may tend to receive a higher probability assessment attached to program outcomes.

Policy Relevance. A governmental agent also assesses the policy relevance of the information when estimating probable program outcomes. This element of the discount factor evaluates information in terms of variables a governmental agent can manipulate as well as the relevance of the information for the decision at hand.

Reliability and Validity. The reliability and validity of the information is another consideration as well. This element of the discount factor focuses on the accuracy of the information. If the information facilitates sound inferences concerning program outcomes, then the information tends to be weighed more heavily. The implications of the program decision calculus model concerning the consumption aspects of the evaluation enterprise are examined next. After this discussion, which identifies propositions concerning utilization, the characteristics of an evaluation product and their linkage to the discounting procedure will be explored.

The Implications of The Program Decision Calculus Model

Explaining Past Empirical Findings. The program decision calculus model sheds some insight into and provides an explanation for why so few program evaluation products contribute to identifiable program changes. In addition, this model lends some support to broadening the perspective of utilization to include an enlightenment function of program evaluation products. Based on the rational choice approach taken and subsequent model developed, utilization of a program evaluation product is defined as information generated entering a governmental agent's decision calculus. This perspective of utilization accommodates both the decisionistic and enlightenment approaches found in the literature.

The decisionistic perspective defines utilization in terms of observable program changes or recommendations made based on program evaluation findings. The program decision calculus model accommodates this perspective as follows: If the information in a program evaluation product is not discounted highly in the first step of the calculation and the probability outcomes calculated in the second step is compatible with a governmental agent's self-interest, then one would expect an observable program change or recommendation for change based on a program evaluation product. Yet, the previous Chapter concerning the supply aspects noted that there are often a number of individuals who produce information in an effort to influence a program choice.

Proposition 5.1

When a governmental agent confronts a program choice, there are a number of individuals producing information attempting to influence his program choice.

Proposition 5.1 implies that a program evaluation product must compete with other information produced by less formal methods. Thus, a program evaluation product must either be weighed more heavily or be consistent with the bulk of other information processed by a governmental agent in order to effect observable program changes. The lack of observable influence on program choices and strategies has troubled both evaluation scholars and practioners alike (eg., U.S. Senate Committee on Human Resources, 1977; Rein & White, 1977).

However, the program decision calculus model may shed some insight into why relatively few program evaluation products contribute directly to observable program changes and recommendations: A program evaluation product must compete with other information packages and be discounted accordingly. If a program evaluation product survives this competition and discounting scrutiny and fairly high probable outcomes are given to a program strategy based primarily on the evaluation, then the assessment must be compatible with a governmental agent's self-interest. Thus, it is hardly surprising that few program evaluation products can be linked to discrete and observable program changes or recommendations.

The decision calculus model also provides some underpinning for expanding the concept of utilization to include "enlightenment". This perspective of utilization defines use as influencing a governmental agent's thinking or ideas concerning program strategies and options. Unlike many previous efforts, this model and its discounting procedure provide a framework for determining the amount of influence a program evaluation product may make on a program choice or recommendation. In addition, this proposed model directs studies of enlightenment use to explicitly examine the amount of influence a program evaluation product makes to a program recommendation vis-a-vis other information processed during the decision calculus process.

Utilization and Decision Costs. From the perspective of utilization taken here, the potential influence of a program evaluation product depends, in part, on the information entering the decision calculus process.

Proposition 5.2

In order to enter a governmental agent's decision calculus, the information must be available when it is needed.

Thus, a program evaluation product has a greater chance to influence a program choice, decisionistic or enlightenment, if the timing of the information coincides with a pending decision. This proposition directs research efforts to investigate and to isolate production parameters which affect timing the final results to meet decisional needs. Besides releasing information when it is needed by govern-

mental agents, the search and processing costs involved with a program evaluation product may also affect its competitiveness vis-a-vis other types of program information.

Search Costs. Consumers of information, like the suppliers of information, incur costs when searching for and processing information. Based on a rational choice perspective, it seems that a governmental agent tends to seek information when the benefits derived from the information outweigh the costs of making a program choice detrimental to his self-interest. Acquiring information is not a costless enterprise. Search costs include time and money expended when attempting to gather information which bears on a program choice. Thus, procedures used for disseminating a program evaluation product seems to be an important issue for utilization.

Proposition 5.3

If the search costs for a governmental agent are reduced, then an evaluation product has a greater chance of entering a governmental agent's decision calculus.

While there is some discussion of dissemination practices in the evaluation literature, Proposition 5.3 directs research efforts towards ascertaining the search costs involved with different types of dissemination practices such as briefings, conferences, memoranda, and the like. Guba (1972) noted early that little attention was given to dissemination methods by evaluators or their sponsors. More recently, empirical studies suggest that informal dissemination techniques tend to distribute evaluation

findings for consideration by government officials. For example, Caplan and his associates (1975) found that 80 percent of the federal executives interviewed cited the news media as one primary source of social science information in general. Another study of practitioners in mental health found that many findings were spread by word-of-mouth, particularly at professional conferences (Brown et al, 1978). Thus, it does not seem sufficient to merely send a program evaluation product to a governmental agent and assume that it enters his decision calculus.

Since program managers, elected decision-makers, and policy administrators have many competing demands on their time, they often rely on information processors to reduce their search and processing costs. Therefore, when disseminating program evaluation products, whether through formal or informal channels, it seems advantageous to identify these information processors. By ensuring that staff analysts receive program evaluation products as well as their superiors, one increases the likelihood that the information will be used--either in a decisionistic or an enlightenment sense.

Processing Costs. Processing costs, like search costs, involve expenditures of time. A governmental agent must read or assimilate the information provided by a program evaluation effort during the first step of his decision calculus. However, governmental agents vary in their capacity to process and to assimilate highly technical information.

Proposition 5.4

If the processing costs of a governmental agent are reduced, then a program evaluation product has a greater chance of entering a governmental agent's decision calculus.

Some governmental agents may have a specialized policy staff, information processors, to interpret information--especially methodologically sophisticated information--while others may lack this resource. Thus, one factor which may affect the amount of potential influence an evaluation product has in the program decision calculus is the format of the information. Format issues which may affect processing costs include the amount of technical jargon used in the report, the presence or absence of an executive summary, and the like.

Summary

Applying a rational choice perspective and constructing a model of a governmental agent's decision calculus based on these assumptions offers a framework for integrating past empirical findings and points to fruitful areas of further research. The model of the decision calculus encompasses seemingly disparate conceptualizations of use--decisionistic and enlightenment. In addition, this model and its discounting procedure provide a framework for determining the relative amount of influence a program evaluation product makes vis-a-vis other types of program assessments. By approaching utilization from a rational choice perspective, research is directed at determining the decision costs, search costs, and processing costs involved.

Utilization of program evaluation products has emerged as a topic of increasing interest to practioners and as an area subject to more empirical assessment by scholars. The framework offered here provides some clarity and organization concerning the consumption aspects of the evaluation enterprise. The section which follows identifies and discusses the important consumption aspect variable in the model of the evaluation enterprise: The final program evaluation product.

THE CONSUMPTION ASPECT VARIABLE:
THE PROGRAM EVALUATION PRODUCT

The major consumption aspect variable identified for purposes of building a model of the evaluation enterprise is the program evaluation product. During his program decision calculus process, a governmental agent applies a discount factor to the information received. For analytical purposes, the four elements of this discount factor correspond to characteristics of the information scrutinized: (1) The producer, (2) the implication/nature of the findings, (3) useability potential, and (4) technical quality. Figure 5-2 lists each discount factor element and its correspondence to a major characteristic of a program evaluation product.

DISCOUNT ELEMENT	PROGRAM EVALUATION PRODUCT CHARACTERISTIC
Source of Information	Producer
Interpretive Bias	Nature/Implications of Findings
Policy Relevance	Useability Potential
Reliability & Validity	Technical Quality

Figure 5-2. The Discount Element and The
Corresponding Program Evaluation
Product Characteristic

The characteristics which receive scrutiny and subsequent weighing or discounting by a governmental agent are culled, in part,

from discussions and empirical findings found in the program evaluation literature. A discussion of the discounting procedure and each characteristic of a program evaluation product receiving scrutiny follows.

The Producer of A Program Evaluation Product

During the discounting process, a governmental agent considers the source which produced a program evaluation product.

Definition 5.3 - Producer of Information

A producer of information is the individual who generates a program evaluation product (or less formal types of information).

The probabilities attached to program outcomes may be reinforced, increased, or decreased according to a governmental agent's assessment of the source that produces the information. Two factors may affect this assessment: (1) The past experience with and/or the reputation of the producer, and (2) the bias of the producer of a program evaluation product.

Whether or not a governmental agent applies a high discount to a program evaluation product depends, in part, on his individual experiences with or predispositions concerning the producer. In effect, a governmental agent judges the credibility of the producer and his product, then discounts the information accordingly. There is some empirical evidence which suggests that the status of and evaluator and the past experience with an evaluator affects the weight given the final product. For example, Brown and his associates (1978) found the the final evaluation findings were

more acceptable when the title "researcher" rather than "evaluator" is attached to the principle investigator conducting evaluation research. In addition, Carter (1971) notes that if a manager's past experience with an evaluator is positive, then his evaluation product seems more acceptable.

A governmental agent also attempts to detect the bias of the producer of a program evaluation product and discount the information accordingly. Given that different amounts of uncertainty prevail when a governmental agent is facing a program choice, then information is a potential source of influence. Thus, an individual produces information in attempt to influence a program choice or strategy and tends to selectively present facts which may enhance or maintain their goals. Therefore, a governmental agent attempts to discern this potential bias and to detect omitted or exaggerated facts when engaging in the discounting process.

There appears to be a peculiar bias in program evaluation products. With respect to program impact studies, scholars note that there is a preponderance of program evaluation products which are either inconclusive or negative regarding program impact. If one assumes that an evaluator's self-interest is maintaining or enhancing his position as an evaluator, then it seems that an evaluator has a vested interest in these inconclusive or negative findings. That is, if a program evaluation is positive about program impact or reinforces a high level of certainty about the outcomes of program strategies, it suggests that there was no need for an evaluation in the first place. However, inconclusive

or negative findings suggest a need for future evaluations--the lifeblood for an evaluator (Cook & Gruder, 1978).

The Nature or Implications of Findings

The interpretive bias of a governmental agent enters the discounting procedure as well. Every governmental agent has individual preferences, predispositions, ideologies, and the like. When processing information, a governmental agent may selectively perceive facts and interpret information provided based on these preferences and biases. Thus, it is possible for a governmental agent to either misperceive the information or discount/weigh heavily the information because of this perceptual bias. The characteristic of a program evaluation product which triggers this interpretive bias is the nature (or implications) of the program evaluation product findings.

Definition 5.4 - The Nature (Implications) of Findings

The nature of the findings refers to the selected facts presented in a program evaluation product. These facts may tend to support, negate, or be inconclusive about program strategies and outcomes.

The findings of a program evaluation product can be categorized as either reinforcing, negating, or being inconclusive concerning the outcomes associated with program strategies. Thus, a program evaluation product which seems counter-intuitive or contrary to a governmental agent's beliefs or prior calculations may be highly discounted. On the other hand, a program evaluation product which reinforces his prior beliefs or calculations may receive higher probability assessments attached to program strategies.

Psychological characteristics of consumers of program evaluation products have been cited in the literature as a factor which may affect utilization. For example, reactions to negative findings vary given the interpretive bias of a particular governmental agent. Carter found that if findings are opposed to a governmental agent's expectations or pose a threat to his security, then the program evaluation product is generally ignored when making a program choice or recommendation (Carter, 1971). Other empirical studies have substantiated this as well (Caplan et al, 1975; Weiss, 1975). Yet enemies of a program may seize upon negative findings in order to justify program cuts (House, 1974). In this case, the implications of a program evaluation product coincide with a governmental agent's interpretive bias and the information may not be discounted as highly.

Useability Potential Characteristics

A governmental agent also considers the policy relevance of a program evaluation product when estimating probable outcomes and weights the information accordingly during his discounting procedure. Thus, the useability potential of a program evaluation product is scrutinized. The definition that follows also lists specific components of useability potential.

Definition 5.4 - Useability Potential

Useability potential refers to characteristics of a program evaluation product which facilitates a program choice (policy relevance).

1. Addressing pertinent policy and program issues;
2. Releasing information when it is needed by governmental agents;
3. Offering fiscally conservative program options;
4. Stating conclusive findings;
5. Offering recommendations for action and presenting information in an interpretable format.

Although the useability potential characteristic and its components are based on a number of discussions found in the literature (eg., Pressman, 1975; Patton, 1978; Weiss, 1975), few have linked the components together into one characteristic. ^{3/}

Addressing Pertinent Policy and Program Issues. In order for a program evaluation product to enter a governmental agent's decision calculus and not be greatly discounted, it must address programmatic questions and problems. Thus, the useability potential of the information tends to increase if it addresses policy relevant questions. Different types of governmental agents may desire specific types of policy relevant information: A program manager may not discount heavily an implementation study or performance measures in terms of its useability potential. Yet,

^{3/} Weiss & Bucuvalas (1978) found a similar factor when performing a factor analysis of their data.

it seems that there are some general characteristics of program evaluation products which increase or decrease its useability potential regardless of the type of governmental agent consuming the information. While there may be a number of variables affecting program processes and impact, governmental agents need information concerning variables that they can manipulate or change (such as increasing program efforts to effect change rather than raising income of program clients). Thus, a program evaluation product's useability potential hinges in part on addressing questions of interest to governmental agents and studying variables amenable to change. However, many program evaluation products fail to imply politically feasible guidelines for action (eg., Hatry et al, 1976; Caplan et al, 1975).

Releasing Information When Needed. The timing of a program evaluation product with respect to the program cycle is another component of useability potential. If a program evaluation product is to contribute to an observable program change or affect a governmental agent's program choices or recommendations, it must be available when the program decision is pending (eg., Davis and Salasin, 1975; Suchman, 1972; Stake, 1967).

Offering Fiscally Conservative Program Options. The budgetary expenditures implied by an evaluation effort is another component of useability potential. Resources in the action context are generally scarce. Thus, a program evaluation product which suggests fiscally conservative program options may tend to have greater useability potential. Often, however, the budgetary implications are not attached to the research findings.

The fiscal implications of program evaluation products affect both decisionistic and enlightenment aspects of utilization.

Stating Conclusive Findings. Program evaluation products may vary in terms of the degree of certainty attached to the research findings. If the final product makes no conclusive statements, then a governmental agent may have increasing doubts about the program or may find the information is of no help in making a decision. As noted earlier, it may often be in the self-interest of an evaluator to keep findings vague.

Format and Recommendations for Action. The format of a program evaluation product refers to the style and the readability of a product. Thus, an evaluation product which is readily interpretable is more likely to enter the decision calculus and be weighed more heavily than information which is fairly uninterpretable. Here, the conventional wisdom and empirical studies suggest that excessive "scientific jargon" erects barriers which may prevent consideration by governmental agents (eg., Brown et al, 1978; Pressman, 1975; Suchman, 1972). Program evaluation products which contain an executive summary, another format element, may also be given more weight during the discounting procedure.

Recommendations for program action also tend to decrease the processing costs for governmental agents. However, this component of useability potential often seems to be lacking in evaluation products. The literature suggests that the scientific norms held by many evaluators often prevent including recommendations in the final product (eg., Riecken, 1972; Weiss, 1972a).

Technical Quality Characteristics

A governmental agent also considers the reliability and validity of program evaluation products during his discounting process. This element of the discount factor focuses on the technical quality characteristics of a program evaluation product. If the information facilitates sound inferences concerning program outcomes, then a program evaluation product is weighed more heavily in terms of its reliability and validity.

Definition 5.6 - Technical Quality

Technical quality refers to the reliability and validity characteristics of a program evaluation product.

1. Use of a theoretical framework;
2. Measurement reliability and validity;
3. Systematic research design and data collection;
4. Appropriate data analysis techniques and methods;
5. Maintaining objectivity and submitting findings to peer review.

These five elements of technical quality bear on the reliability and the validity characteristics of the final program evaluation product.

Reliability. The reliability of the information refers to the consistency of the measurement. Reliability is defined more specifically on the following page.

Definition 5.7 - Reliability

Reliability refers to whether (1) repeated use of the same indicators and/or (2) use of different indicators for the same concept will provide consistent values (scores) for the phenomenon being measured.

Based on Legee & Francis, 1974

If the measuring process is unreliable, random error results. These random errors occur from idiosyncratic, accidental, or unpredictable responses from the phenomenon being measured or from the measuring instrument itself. If the measuring process is not reliable, the explanatory power (explained variation) of the statistical analysis tends to be reduced.

Validity. Validity is the other major component of technical quality that is scrutinized by a governmental agent who wants to make accurate predictions concerning program outcomes and strategies. Validity, in a broad sense, refers to measuring what one purports to measure. Cook and Campbell (1979) distinguish between four types of validity concerns important for program evaluation efforts. The following definitions are based on their distinctions.

Definition 5.8 - Validity

Validity, in general, refers to actually measuring what one purports to measure.

Internal Validity - measuring the influence of program variables (or other variables of interest) rather than some other confounding factors.

External Validity - concerned with the generalizability and applicability of findings to other times, programs, and clients.

Statistical Conclusion Validity - concerned with the soundness of conclusions drawn about the program on the basis of statistical evidence.

Construct Validity - concerned with measuring the conceptual with the operational definitions and measures.

Based on Cook & Campbell, 1979.

Threats to validity during the measurement and analysis processes raise questions concerning the accuracy of the inferences which can be made based on the program evaluation findings. Systematic error and biased estimates result when threats to any of the four types of validity listed are not controlled adequately.

Both reliability and validity, then, affect the soundness and accuracy of inferences that may be drawn by a governmental agent. If the information is not reliable and valid, a governmental agent can erroneously calculate the probabilities associated with different program strategies. The reliance on reliability and validity as key issues concerning technical quality are also justified since program evaluation relies on scientific research principles.

While the language used to describe the activities often differs, the prescribed conduct of a program evaluation effort parallels the scientific method. Figure 5-3 summarizes the steps of these two research methods and lists the corresponding validity and reliability issues raised during the various stages of evaluation research. ^{4/}

The five specific criteria listed in the definition of technical quality reflect scientific components which are used to judge the validity and reliability of program evaluation products. These five criteria are based, in part, on other conceptualizations and empirical findings discussed in the program evaluation literature (Bernstein & Freeman, 1975; Minnesota Systems Research, Inc., 1975; Patton, 1978; Weiss & Bucuvalus, 1978). They also reflect scientific criteria found in discussions of philosophy of science and methodology (eg., McGaw & Watson, 1976; Leege & Francis, 1974; Cook & Campbell, 1979). Appendix C contains a discussion of each criterion in terms of the validity and reliability issues raised when conducting a program evaluation effort.

^{4/} The steps in the program evaluation process are based on a number of texts (eg., Weiss, 1972a; Franklin & Thrasher, 1977). The scientific method is summarized from a number of works (eg., Babbie, 1973; Leege & Francis, 1974; McGaw & Watson, 1976).

SCIENTIFIC METHOD	EVALUATION RESEARCH METHOD	VALIDITY ISSUE	RELIABILITY ISSUE
(1) Problem Formation, Hypothesis Formation (Derived from Theory)	(1) Identify Program Goals	External Internal	
(2) Concept Formation & Measures Developed	(2) Develop Measures	Construct	***
(3) Design Research & Collect Data	(3) Design Evaluation & Collect Data	External Internal Construct	***
(4) Analyze Data	(4) Analyze Data	Statistical Conclusion	***
(5) Present Findings	(5) Report Findings		

Figure 5-3. Scientific and Evaluation Research Methods :
Validity and Reliability Issues

Concern with Technical Quality. The technical quality of program evaluation products and the affect of this characteristic on utilization has received considerable attention recently in the literature. The lack of technical quality is often cited as a reason for dismissing program evaluation findings. Caplan and his associates (1975) found that the "shoddy" quality of program evaluations was cited by federal executives as the major reason for discrediting any social science information that they received. Likewise, Weiss and Bucuvalas (1978) found that technical quality correlates positively with the use of a program evaluation product.

In general, program evaluation scholars are pessimistic concerning the technical quality of program evaluation products. Impressionistic judgments concerning the quality of evaluations made during the early 1970s concluded that the final products were unsatisfactory, poor, or mediocre at best (Scriven, 1972; Weiss, 1972c). Empirical research has substantiated these early impressions. An early assessment of 181 evaluations of social programs found that 80 percent relied on a pre/post-test research design, only 20 percent employed randomization procedures, and all but a few studies were based on a small number of cases (Mann, 1965). More recent empirical studies assessing the methodological quality of evaluation products reinforce Mann's pessimistic findings.

Bernstein and Freeman (1975) analyzed 236 evaluations sponsored by the federal government which were conducted by academic and entrepreneurial firms. They concluded that on all six of their

scales measuring technical quality, such as sampling and measurement adequacy, only 13 percent were deemed as adequate. In 1974, Minnesota Systems Research, Inc., evaluated 110 studies sponsored by the Department of Health, Education, and Welfare. Upon assessing 50 evaluations and 60 more basic research studies on over 200 attributes, they concluded that less than 10 percent were free of competing explanations regarding their findings or implications.

CONCLUSION

The evaluation enterprise begins when demands for information concerning a program are made and a governmental agent sponsors a formal program evaluation. After this decision is made, a program evaluation effort is implemented in order to produce a package of information for possible consumption by governmental agents facing program choices and decisions. This Chapter, based on a rational choice perspective, developed a model of a governmental agent's decision calculus that links the discounting procedure to four characteristics of the information produced: The producer, the nature of the findings, useability potential, and technical quality.

Utilization of program evaluation products, although always an important topic in the literature, has received increased empirical attention in recent years. The four characteristics identified as important in influencing the amount of impact the information may make when a governmental agent uses a program evaluation product are based, in part, on the discussions and empirical studies offered in the program evaluation literature. Figure 5-4, displayed on the following page, depicts the program evaluation product, the major consumption aspect variable, and its relationship to supply and demand aspect variables.

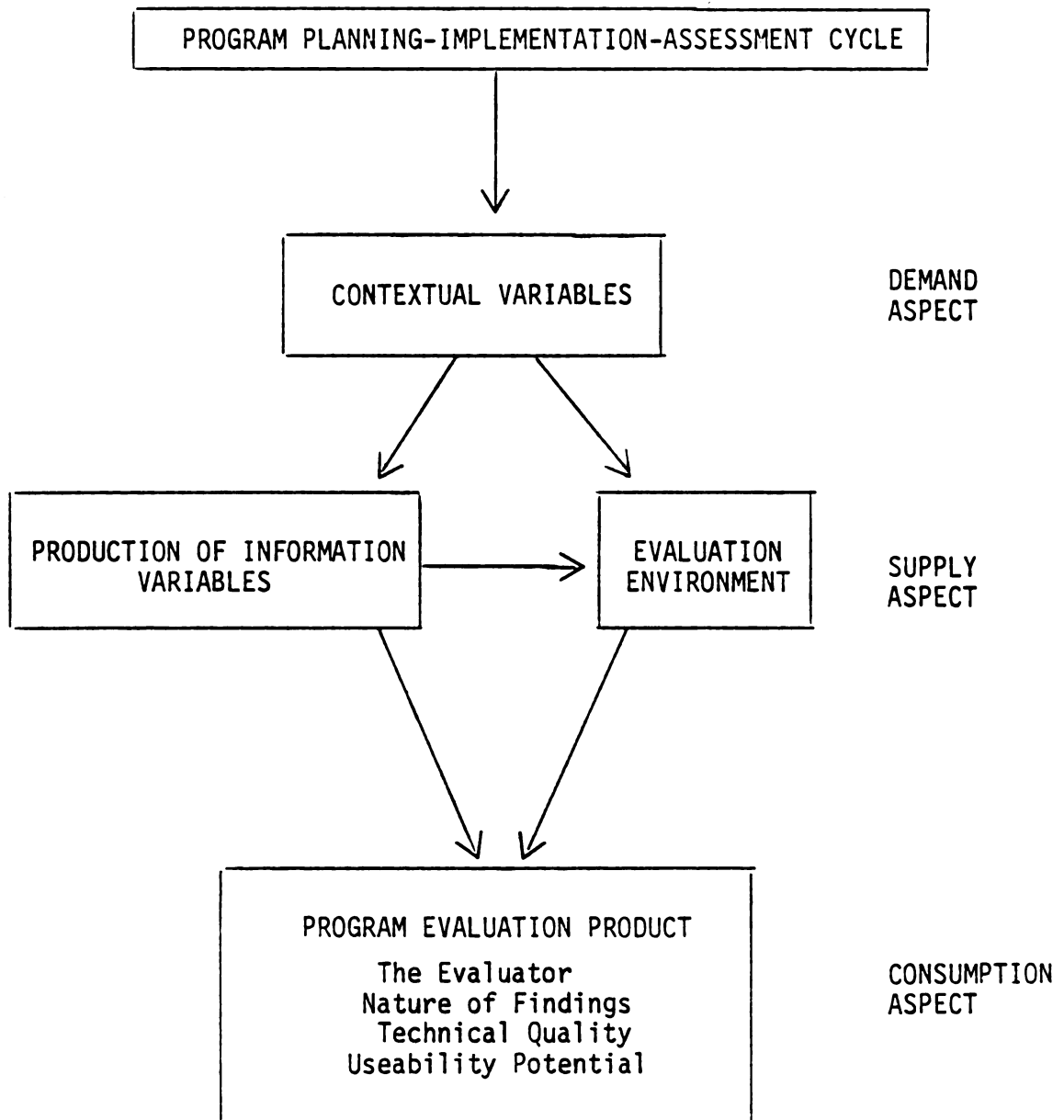


Figure 5-4. The Evaluation Enterprise:
The Consumption Aspect Variable

Each aspect of the evaluation enterprise has been examined based on and extending the broad framework of the program cycle and the rational choice assumptions made about individuals in the program cycle and its action context. In addition, important variables associated with each aspect have been culled from the program evaluation literature in an effort to build a model of the evaluation enterprise which identifies the conditions and motives instigating the evaluation effort (demand aspect), the production of information given resource and environmental constraints (supply aspect), and the characteristics of the final product which may affect its influence on a governmental agent's program choices and recommendations (consumption aspect). The next Chapter describes the full model of the evaluation enterprise, explores possible linkages between technical quality and useability characteristics of the final product and supply aspect variables, and examines possible relationships between the contextual variables and the choice of an evaluation unit, allocation of resources, and the evaluation environment which generate the program evaluation product.

CHAPTER SIX

A MODEL OF THE EVALUATION ENTERPRISE

A rational choice perspective has been taken throughout the dissertation in order to examine the demand, supply, and consumption aspects of the evaluation enterprise. The emphasis in the three previous chapters was on analyzing each individual aspect: The demand aspect which is often neglected in the literature, the supply aspect which has been a traditional concern to evaluation scholars, and the consumption aspect which is receiving increasing attention by scholars and practitioners alike. Some interesting insights and propositions develop by applying a rational choice framework to each aspect. Yet, one of the major shortcomings in the field of program evaluation is the lack of integration among the demand, supply, and consumption aspects of the evaluation enterprise. In an effort to address this perceived shortcoming, then, important variables associated with each aspect were identified and discussed briefly in the previous chapters. This Chapter links these variables in a proposed model of the evaluation enterprise by relying on the reasoning offered in earlier chapters and suggests a number of testable propositions for future research.

THE EVALUATION ENTERPRISE:
A PROPOSED MODEL

The evaluation enterprise begins when demands for a formal assessment of a public program arising from individuals in the program planning-implementation-assessment cycle or its action context and a governmental agent decides to sponsor a program evaluation effort. Program evaluation, as defined here, consists of applying scientific principles to address policy and program questions--an attempt to collect program information systematically and minimize potential bias. In brief, an evaluation effort begins by identifying program goals and issues that the evaluation sponsor wants addressed. Next, measures for the variables of interest are developed, a research strategy designed, and collection of data begins. Upon successfully collecting the data, an analysis is performed and the final information is packaged for possible consumption in the program cycle. While this production method may seem quite straightforward, conflicting demands, incentives, and resources often impinge on and constrain an evaluator's efforts.

The model proposed in this Chapter identifies the sources of conflicts and constraints for a program evaluation effort and examines their impact on characteristics of the final product which affect utilization. Figure 6-1 depicts the proposed model of the evaluation enterprise which links the demand, supply, and

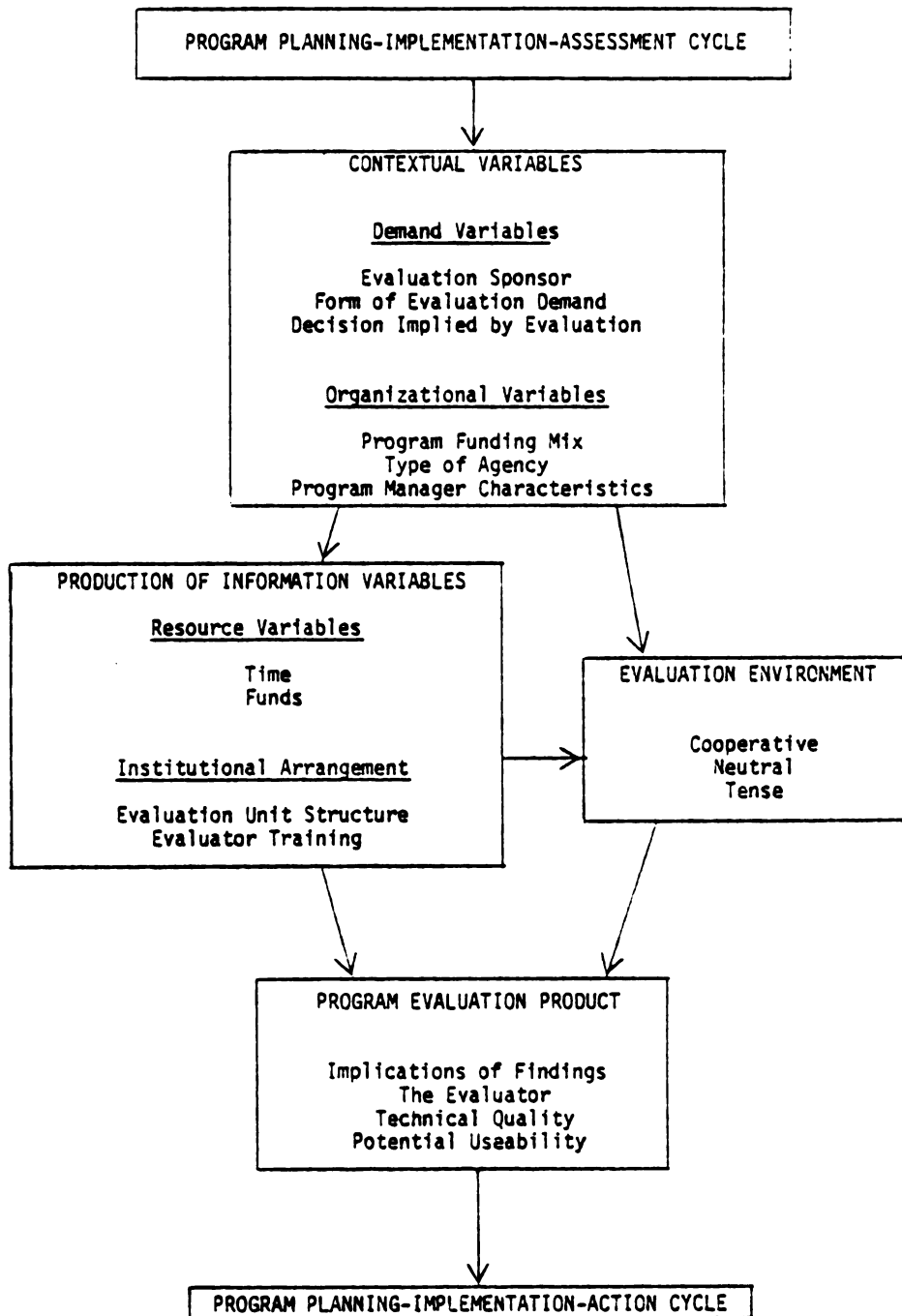


Figure 6-1. A Model of The Evaluation Enterprise

consumption aspect variables identified and discussed in the three previous chapters.

A Description of The Proposed Model

The contextual variables reflect the demand aspects of the evaluation enterprise. The demand variables (evaluation sponsor, form of the demand, and the type of decision implied by an evaluation effort) all are involved with the decision to evaluate a public program. The organizational variables (type of program agency, program manager characteristics, and funding mix for the program) provide the setting for the evaluation effort. The demand and organizational variables are grouped as contextual variables since they are relatively fixed prior to conducting a program evaluation.

After the decision to evaluate has been made and a program becomes the organizational target for the evaluation effort, attention turns to the supply aspects in the proposed model. Here, there are two important groupings of supply variables: Production of information variables and the evaluation environment. As noted by the arrows in Figure 6.1, the contextual variables affect production of information variables (resources and institutional arrangement) as well as the evaluation environment. The production of information variables generate the final program evaluation product as well as influence another supply aspect variable--the evaluation environment. The evaluation environment refers to the milieu for actually conducting the program evaluation.

SUPPLY ASPECT VARIABLES AND
TECHNICAL QUALITY AND USEABILITY POTENTIAL
CHARACTERISTICS

This section explores the linkages between supply variables identified in the evaluation enterprise model (resources, institutional arrangement, and evaluation environment) and two major characteristics of the final evaluation product which may affect its use by governmental agents--technical quality and useability potential. First, evaluation product characteristics scrutinized by a governmental agent are reviewed briefly. This discussion suggests that the implications of the findings and the evaluator characteristics scrutinized may be a direct effect of the demand aspect variables while technical quality and useability potential characteristics seem to be more directly linked to variables identified with the supply aspect of the evaluation enterprise. Based on this reasoning and in an effort to explore the joint effects of the supply variables, the analysis is restricted to technical quality and useability potential characteristics. After analyzing each supply aspect variable's potential affect on these characteristics, a matrix of various technical quality and useability mixtures based on joint combinations of resources, institutional arrangements, and evaluation environment conditions is proposed.

The outcome of the program evaluation effort is a program evaluation product. Both major supply aspect groupings, the production of information variables and the evaluation environment, may affect the characteristics of the final product as depicted by the arrows in Figure 6.1. With the program evaluation product variable, attention shifts to the consumption aspects of the evaluation enterprise. When using a program evaluation product, it was suggested in Chapter Five, a governmental agent scrutinizes four characteristics and weights it accordingly: The producer (the evaluator), implication of findings, and useability potential. These characteristics affect, in part, the amount of influence the information generated by a formal program evaluation effort may make on a governmental agent's program recommendations and choices. The remainder of this Chapter focuses on how the variables in the evaluation enterprise model may contribute to characteristics of the final product which contribute to utilization. Analysis first focuses on linkages between supply aspect variables and technical quality and useability potential characteristics. Following this discussion, possible linkages between contextual variables and the supply aspect variables (choice of an evaluation unit, resources allocated, and the evaluation environment) are examined.

Characteristics Affecting Use

A governmental agent scrutinizes four characteristics of an evaluation product during his decision calculus and weights the information accordingly. Two of these characteristics, the nature of the findings and the producer of the information, seem to be largely determined by the demand aspect variables in the model. That is, the evaluation sponsor selects the producer (evaluator) during the initial stages of an evaluation effort. The dynamics of the choice of an evaluator may also be related to the nature of the findings produced in terms of evaluator responsiveness (autonomy) to evaluation sponsor demands. Technical quality and useability potential characteristics, on the other hand, seem linked more directly to supply aspect variables in the model. That is, the final mixture of technical quality and useability potential is a function of the skills and orientations of the evaluator given various resource and environmental constraints.

Technical quality, as defined earlier, is comprised of attributes that reflect the validity and reliability of the information. Thus, a program evaluation product possesses high technical quality if it is implemented systematically, minimizes bias through a peer review process, and uses appropriate means to address the policy questions raised. The useability potential of a program evaluation product is comprised of attributes that are policy relevant. That is, a program evaluation product is potentially useable if it addresses policy questions, the findings are conclusive,

information is released when it is needed, budgetary implications tend to be fiscally conservative, and the final report contains recommendations for action.

Institutional Arrangements: Affects on Technical Quality and Useability Potential Characteristics

Many scholars in the evaluation literature have addressed the issue of how evaluator skills and an evaluator's organizational setting affect the technical quality and useability potential of the final product produced. These discussions and empirical works often suggest that technical quality is a function of methodological skills while useability potential is a function of the policy and program skills possessed by the evaluator. However, the skills and orientation of the evaluator are not the only important factor; the evaluation organization and its reward structure also shape the research strategies of evaluators and, in turn, impact on the final product's technical quality and useability potential (eg., Bernstein & Freeman, 1975; Pressman, 1975; Suchman, 1972; Weiss, 1972b; Floden & Weiner, 1978; Freeman, 1975). This literature, however, tends to focus mainly on the technical quality characteristics of an evaluation product. In addition, analysis is generally restricted to examining one variable, such as the type of evaluation unit, and its affect on technical quality characteristics. In an effort to extend analysis to incorporate additional supply aspect variables and examine their joint affect on both technical quality and useability potential characteristics, the reasoning developed in

earlier chapters is continued. This section begins by stating the assumptions about evaluator skills and orientations made, next examines organizational incentives that shape evaluation strategies, and propositions concerning mixtures of technical quality and useability potential one would expect given different types of institutional arrangements are then posited.

Evaluator Skills and Training. Evaluators, like all actors in the program cycle and action context, are rational and pursue strategies that maintain or enhance their self-interest. Here it is posited, building on distinctions made in the literature, that an evaluator is motivated by two goals when conducting a program evaluation: maintaining his professional reputation and maintaining his job in the evaluation unit.^{2/} For analytical purposes, two types of evaluators possessing different skills and orientations are distinguished: social scientists and practitioners. These "ideal types" of evaluators possess different types of training and skills (methodological/program) as well as orientations to different professional communities (scientific/policy).

It is assumed here (Assumption 6.1) that evaluators termed social scientists receive the methodological and statistical training conducive to producing an evaluation product which tends to exhibit relatively high technical quality, skills which enhance technical quality characteristics such as

^{2/} See Chapter Four (pages 100 to 102) for a discussion of this distinction and relevant literature.

Assumption 6.1

The type of training and orientation that an evaluator possesses affects, in part, his research during an evaluation effort.

- (a) A social scientist possesses methodological skills and an orientation towards the scientific community and tends to pursue research strategies that enhance technical quality characteristics.
- (b) A practitioner possesses policy skills and an orientation towards the program community and tends to pursue research strategies that enhance useability potential characteristics.

concept measurement, design, data collection, and data analytic methods. When acquiring these technical skills, the social scientist's training also instills a number of scientific norms: publication in scientific journals, peer critique and review, and autonomy to formulate and conduct research of interest. Thus, the social scientist tends to embrace the norms of the professional social science community.

Practitioners, on the other hand, develop policy and program skills that tend to enhance useability potential characteristics such as identification of variables manipulable by policy-makers, knowledge of political and budgetary constraints, and formulating policy recommendations and alternatives. Practitioners generally embrace a different set of norms than do social scientists. For example, they tend to stress innovative policy strategies suggested, political practicality, meeting program cycle deadlines, and the like.

Based on these differences in training and norms, social scientists and practitioners tend to look at different communities for professional approval: A social scientist attempts to maintain his reputation within academic and scientific circles while a practitioner guards his professional reputation within bureaucratic and policy circles. Thus, a social scientist tends to pursue evaluation strategies that will maintain or enhance his reputation with the scientific community such as focussing on a program aspect that can be fit into a larger theoretical framework, collecting the data in a manner that increases its generalizability, employing appropriate and sophisticated methodological techniques, and completing a study which can be potentially published. In turn, the technical quality of the final product may be enhanced.

In contrast, the practitioner tends to pursue evaluation research strategies that maintain or enhance his reputation with the policy and program community such as writing the report in a manner that summarizes the findings, completing the evaluation effort when scheduled, collecting information concerning costs of various program strategies, and offering innovative policy recommendations for action. In turn, the useability potential of the final product may be enhanced.

Although the evaluator, whether a social scientist or a practitioner, possesses skills and tends to pursue evaluation strategies that protect or enhance his professional reputation, the primary self-interest of an evaluator is to secure or enhance his position in the evaluation unit.

When the incentives or sanctions of the evaluation unit conflict with an evaluator's professional goals, he is faced with two choices--either leaving the evaluation unit or modifying his research strategies to accommodate evaluation unit goals. Therefore, an evaluator's research strategies to enhance his professional reputation are subjugated, in large part, to evaluation unit demands and constraints. It is important, then, to analyze organizational incentives and sanctions which shape an evaluator's strategies when conducting a formal program evaluation effort.

Type of Evaluation Unit: Organizational Incentives and Sanctions. While the evaluation literature has often noted differences among types of evaluation units, organizational incentives, and the technical quality and useability potential of the final product generated, the linkages among them have not been integrated systematically. By applying rational choice assumptions to the administrator of the evaluation unit, organizational incentives and sanctions perhaps can be linked more systematically to various types of evaluation units: internal, governmental, entrepreneurial, and academic.

It is suggested here that an evaluation unit's reward structure depends, in large part, on the individual who administers the unit. Like an evaluator, it is assumed that the evaluation unit administrator pursues strategies that maintain or enhance his self-interest--protecting his job and professional reputation. It seems important, then, to identify the individual who can threaten the evaluation unit by withdrawing support as well as the

professional orientation of the administrator. The assumption listed on the following page (Assumption 6.2) lists the primary individual who may threaten an administrator's job security, the major professional community with which the administrator identifies and guards his professional reputation, and the resulting types of research strategies stressed in different types of evaluation units. These assumptions are based, in part, on a number of discussions and empirical work found in the program evaluation literature.

The notion that the major audience of an evaluation unit shapes organizational incentives and sanctions was suggested by Bernstein and Freeman (1975). The classification of specific audiences addressed is based on a number of other works found in the literature. A number of scholars suggest that internal evaluation units tend to respond to the program manager and be oriented towards the policy community (eg., Weiss, 1972b; Suchman, 1972; Pressman, 1975). The assumptions made concerning academic and entrepreneurial evaluation units are based largely on Bernstein and Freeman's empirical work. However, Coleman (1972) and Williams (1972) also lend support for the assumptions made concerning academic units. In terms of the governmental evaluation unit, the assumptions made are more speculative in nature, based primarily on casual observation. The resulting organizational reward structure based on these assumptions and its impact on shaping evaluator research strategies is explored further.

Assumption 6.2

An evaluation unit administrator pursues strategies that maintain or enhance his job security and his professional reputation. In turn, the organizational sanctions and the orientation or the evaluation unit are dependent, in part on the response of the evaluation unit administrator to demands by those who can threaten the security of the unit. In addition, the organizational sanctions and orientation are dependent, in part, on the professional community (major audience) with which the administrator identifies.

Type of Unit	Individual Threatening Security	Major Audience (Professional Community)	Organizational Incentives/Sanctions
Internal	Program Manager	Policy Community	Stress Useability Potential
Governmental	Policy Administrator	Policy/Scientific Community	Stress Useability Potential/ Technical Quality
Entrepreneurial	Evaluation Sponsor	Policy Community	Stress Useability Potential/ Technical Quality
Academic	University Official	Scientific Community	Stress Technical Quality

Given the assumptions posited concerning evaluator skills and orientations (Assumption 6.1) and organizational incentives (Assumption 6.2), the relationship between various institutional arrangements and technical quality and useability potential characteristics of the final product can be examined. Institutional arrangement refers to the incentives and staffing patterns of the evaluation unit. The four types of institutional arrangements

Definition 6.1 - Institutional Arrangements

An Internal Evaluation Unit, located within an operating program, stresses policy-oriented research strategies and is staffed mainly by evaluators with practitioner skills and orientations.

A Governmental Evaluation Unit, located in an agency that is not directly attached to an ongoing program, stresses both policy and scientific-oriented research strategies and is staffed by a mix of evaluators with practitioner and social science skills and orientations.

An Entrepreneurial Evaluation Unit, located outside of government, stresses both policy and scientific-oriented research strategies and is staffed by a mix of evaluators with practitioner and social science skills and orientations.

An Academic Evaluation Unit, located within a university, stresses scientific-oriented research strategies and is staffed mainly by evaluators with social science skills and orientations.

vary in terms of social scientist/practitioner mixtures, organizational incentives, and in turn, the research strategies pursued which tend to affect the technical quality and useability potential of the final product.

Internal Evaluation Unit. A program evaluation unit attached directly to an ongoing program is one type of institutional arrangement commonly employed for conducting a program evaluation. The proximity of the evaluation unit to the operating program results in a milieu that tends to reward policy aspects of the final product. Given the proximity of the unit to the program, the evaluation administrator attempts to maintain or enhance his position which can be threatened by the program manager as well as protect his reputation within policy circles. In terms of the skills and orientations of the evaluators, the literature suggests that evaluators classified as "practitioners" tend to staff internal evaluation units (eg., Weiss, 1977; Suchman, 1972). Although some evaluators with social science skills and training may be part of the staff, in order to maintain their positions they tend to pursue research strategies which are rewarded. Thus, the internal evaluation unit tends to produce a final evaluation product that emphasizes useability potential characteristics such as policy pertinence, timing considerations, conclusive findings, action recommendations, and options that are feasible within current budgetary constraints. The lack of social science skills and orientations in the evaluation unit, coupled with organizational incentives that do not necessarily emphasize methodological rigor but tend to stress completing the evaluation on time and the like, tends to result in a final product that often seems to lack technical quality.

There is some support found in the literature concerning the types of skills, organizational norms, and the generally low technical quality yet perhaps policy pertinent information produced by an internal evaluation unit (eg., Suchman, 1972; Weiss, 1977; Pressman, 1975; Aronson & Sherwood, 1965). However, the skills, norms and incentives associated with a governmental unit are more speculative in nature.

Governmental Evaluation Unit. While the evaluation literature tends to associate an internal evaluation unit with a final product that tends to suffer in terms of technical quality yet the proximity to the program tends to enhance useability potential, there has been little speculation or examination of what is classified here as governmental evaluation units. Yet increasingly, it seems, that agencies are establishing evaluation units separated from any particular program and often afforded a high degree of organizational autonomy from program manager's demands and pressures. While admittedly speculative and based on casual observations, it seems that many governmental evaluation units tend to resemble academic evaluation units in terms of stressing research norms yet their governmental base also stresses policy community norms. It is posited here that evaluation unit administrator shares both policy and research community norms. In addition, his position can be threatened by a policy administrator who may tend to stress policy relevant aspects of the program evaluation effort-- demands for releasing information when it is needed, addressing pertinent policy issues, formulating feasible recommendations,

and for suggesting alternatives within budgetary constraints. Yet, the evaluation unit administrator is also concerned with maintaining the integrity of the evaluation unit and insuring its continued existence by providing information that is reliable and valid. This results in organizational incentives that may shift in terms of rewarding policy oriented and social science oriented research strategies.

In terms of the staffing patterns, the governmental evaluation unit is staffed by both social scientists and practitioners. Although there may be some evaluators with practitioner orientations and skills, it is posited that more social science oriented evaluators tend to staff this type of unit. While the evaluation unit may reward research strategies that enhance technical quality, there are demands placed on evaluator strategies that may enhance useability potential as well. Thus, an evaluator, whether a social scientist or a practitioner, in an effort to maintain his job within the organization, must design research strategies accordingly. The governmental evaluation unit, given the organizational norms and evaluator skills, tend to produce a final product that exhibits more technical quality than one produced by an internal unit; they also, however, tend to produce evaluation products that are sensitive to the issues and demands of the policy community.

Entrepreneurial Evaluation Unit. Another institutional arrangement employed for conducting a program evaluation effort is the entrepreneurial firm. This type of evaluation unit is a

profit-making organization that contracts with evaluation sponsors to conduct an evaluation. Since an entrepreneurial firm is a profit-making venture, the evaluation unit administrator is concerned with maintaining sound relationships with evaluation sponsors and cultivating future evaluation clients; he also tends to look to the policy and program community for professional approval. Thus, resultant organizational norms dictate adaptation to demands made by evaluation sponsors (Bernstein & Freeman, 1975). The emphasis on technical quality and useability potential research strategies shifts depending on the sponsor's demands outlined in the formal contract. Often, however, the criteria contributing to useability potential such as time deadlines are easier to specify in a formal contract than are demands for technical quality such as concept measurement. In part, the entrepreneurial evaluation unit's continued existence and prestige rests on delivering a final product that contains a minimum level of validity and reliability (technical quality). However, in order to ensure future contracts, the firm must also deliver a final product that facilitates decision-making within time constraints (useability potential).

The entrepreneurial firm is staffed with a mixture of social scientists and practitioners. In order to minimize organizational costs and maximize profits, however, there is some suggestion that the entrepreneurial evaluation unit tends to assign a small number of well-trained social scientists to conduct the evaluation research. Although social scientists are expected to choose evaluation research strategies that protect their professional

reputation among other social scientists, their positions in the unit may impinge on those strategies. In order to maintain his position, a social scientist must choose evaluation strategies that meet the demands of evaluation sponsors and enhance the firm's profits such as minimizing costs by using methods already developed even though technical quality could be improved by designing new methods and measures.

Practitioners staff entrepreneurial evaluation units as well. These individuals, in contrast to social scientists, tend to pursue strategies that contribute to the useability potential components of the final product. However, they must often pursue some research strategies that emphasize technical quality components in order to maintain the firm's reputation as well as meet demands of evaluation sponsors.

The entrepreneurial evaluation unit, given the organizational incentives and evaluator skills, tend to produce a final product that may exhibit more technical quality than one produced by an internal unit and perhaps less technical quality than a governmental unit depending on the skills of the social scientists; since useability characteristics may be easier to specify than technical quality in formal contracts, the final product may tend to possess greater useability potential than technical quality.

There is some support in the literature concerning the types of evaluator skills and organizational norms posited here. In addition, the empirical work of Bernstein and Freeman (1975)

as well as the conventional wisdom offered in the literature suggests that entrepreneurial firms tend to produce evaluation products that often sacrifice methodological rigor in order to meet client demands for releasing the information when needed (Bernstein & Freeman, 1975).

Academic Evaluation Unit. A program evaluation unit or an evaluator based in a university is an institutional arrangement discussed frequently in the literature (eg., Coleman, 1972; Bernstein & Freeman, 1975; Williams, 1972). Here, the evaluation unit administrator is oriented towards the social science community and is responsive to university officials sanctions and incentives. This results in organizational incentives and sanctions that reflect social science norms such as publication in scholarly journals and the like. Evaluators with social science skills and orientations tend to staff the academic evaluation unit while evaluators with practitioner training and skills are generally absent.

An evaluator's social science training and norms are consistent with their organizational environment. In order to retain his job in the university and enhance his professional reputation in the social science community, the evaluator pursues research strategies that tend to enhance technical quality. In order to publish findings in a scholarly journal and retain their university jobs as well as protect their professional reputations, the social scientist focuses attention on linking the evaluation research to a theoretical framework, measuring concepts in a

valid and reliable manner, collecting data systematically, and analyzing the data using appropriate methods. Unlike many evaluators found in other types of evaluation units, the academic evaluator is likely to submit research findings for criticism by the larger scientific community.

While the academic evaluation unit, composed primarily of social scientists, rewards research strategies that stress technical quality, incentives stressing useability potential aspects of the evaluation effort are often absent. The training of social scientists, strong methodological training, often does not sensitize them to policy concerns. Furthermore, organizational incentives and professional goals do not tend to encourage policy relevance at the expense of technical quality. Therefore, when faced with a research decision that requires a choice that impairs either potential useability or technical quality, the academic evaluator is likely to follow a strategy that improves technical quality.

The academic evaluation unit, given the organizational incentives and cluster of evaluator skills and orientations, then, tends to produce an evaluation product that exhibits high technical quality yet may suffer in terms of useability potential characteristics. There is some support found in the literature concerning the types of skills, organizational norms, and the resulting evaluation product generated by an academic unit. This literature suggests that academic evaluation units tend to produce high quality evaluation products that often neglect policy

relevance issues, timing considerations, and the like (eg., Weiss, 1977; Coleman, 1972; Williams, 1972).

Institutional Arrangements and Mixtures of Technical Quality and Useability Potential Characteristics in The Final Product

By applying rational choice reasoning, making explicit assumptions, and building on the discussions found in the literature, the different types of institutional arrangements and the technical quality and useability potential mixtures of the final product can be examined. Proposition 6.1, based on the previous discussion, suggests mixtures of technical quality and useability potential characteristics that may tend to be associated with different institutional arrangements given that evaluators pursue rational strategies. These posited mixtures associated with various types of institutional arrangements are based on assumptions concerning the research strategies an evaluator ideally pursues given the organizational incentives and sanctions of his evaluation unit.

There are, undoubtedly, exceptions to the mixtures of technical quality and useability potential found in the evaluation products generated by the institutional arrangements as suggested by Proposition 6.1. That is, there may be entrepreneurial firms that stress technical quality aspects and tend to be staffed primarily by evaluators with social science skills and orientations, some academic evaluation units may tend to be client and policy oriented and tend to reward evaluation strategies that lead to higher technical quality, and the like. However, Proposition 6.1

Proposition 6.1

Different institutional arrangements, organizational incentives and evaluator skills and orientations, affect the research strategies an evaluator pursues and, in turn, results in different mixtures of technical quality and useability potential characteristics found in the final evaluation product.

- (a) An internal evaluation unit tends to produce a final product that exhibits useability potential but often lacks a high degree of technical quality.
- (b) A governmental evaluation unit tends to produce a final product that exhibits greater technical quality than one produced by an internal evaluation unit; the final product also exhibits a fairly high degree of useability potential.
- (c) An entrepreneurial evaluation unit tends to produce a final product that has greater useability potential than exhibits technical quality.
- (d) An academic evaluation unit tends to produce an evaluation product that exhibits high technical quality but often lacks useability potential.

provides a beginning point to examine systematically the relationship between characteristics of final evaluation products and various types of institutional arrangements used to implement the evaluation effort. Here, attention is focused on identifying and comparing evaluator skills and norms and organizational reward structures of various types of evaluation units cited in the literature. The reasoning and proposition offered suggest extending analyses of types of evaluation units and product characteristic linkages to include comparative analysis of internal and governmental evaluation units as well as consider the

Previous empirical work in this area focussed mainly on academic and entrepreneurial evaluation units and their affect on technical quality characteristics (eg., Bernstein & Freeman, 1975).

In terms of examining the relationship between an institutional arrangement used to conduct an evaluation effort and product characteristics, it is suggested here that technical quality tends to be enhanced when an evaluator possesses sound methodological skills, an orientation towards the scientific community, and his organizational setting rewards research strategies which focus on the reliability and validity of the information produced; useability potential tends to be enhanced when an evaluator possesses sound policy skills, an orientation towards the program community, and his organizational setting rewards research strategies which focus on the policy relevance of the information produced (see Figure 6-2). That is, a rational evaluator pursues research strategies that enhance his self-interest in terms of his professional reputation as well as his job security. Yet, whether an evaluator can pursue these strategies is also shaped by resource and environmental constraints.

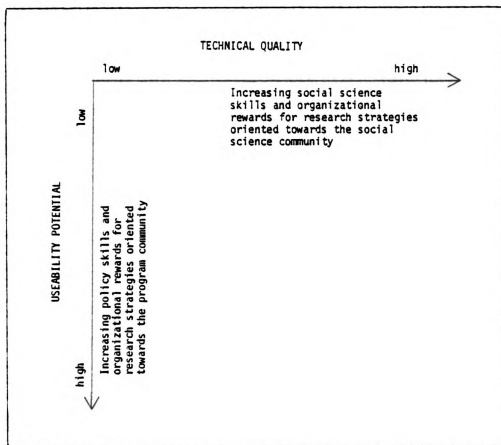


Figure 6-2. Institutional Arrangement:
Technical Quality and
Useability Potential
Characteristics

Resource Variables--Constraints on Rational Strategies

While the type of evaluation unit conducting an evaluation effort may be a central variable affecting technical quality and useability potential, the literature suggests that time and funding constraints may also affect these characteristics. Although these resource variables identified in the evaluation enterprise model may affect both technical quality and useability potential, analysis is restricted here to examining the possible impact of adequate funding on research strategies affecting technical quality and adequate time given on research strategies affecting useability potential. ^{3/}

During the demand aspects of the evaluation enterprise, funds are allocated to implement the effort. The adequacy of funds allocated given the evaluation research task may serve as a constraint on an evaluator, often a social scientist, pursuing research strategies that tend to enhance technical quality. If funds are adequate, the evaluator may be able to develop new measures perhaps more appropriate to the research effort. Yet, if the funding level is less adequate, the evaluator may have to rely on measures already developed although they may be less reliable and

^{3/} Analysis is restricted for the sake of brevity and simplicity. The purpose here is to illustrate how the model and earlier reasoning may be extended to link the variables in a systematic manner.

In addition, when funding is adequate, an evaluator may be able to draw a larger sample of clients or programs for study which, in turn, enhances the generalizability of the findings. Thus, in terms of data collection strategies and ability to process the information, the funding level may impinge, in part, on the rational strategies an evaluator attempts to pursue in order to maintain his position and/or enhance his reputation in the academic community.

The time allocated for an evaluation endeavor also may constrain research strategies an evaluator pursues. Here, the adequacy of time allocated for the research task at hand may affect one important useability potential characteristic--releasing the information when it is scheduled and needed in the program cycle. ^{4/} Here, if the time allotted by the evaluation sponsor is not adequate in terms of finishing the product and injecting the information into the program cycle, the useability potential in terms of influencing a pending decision is diminished. As time becomes more adequate, however, the evaluator, particularly a practitioner, can pursue strategies that may enhance the useability potential of the final product. Besides timing information to program cycle decisions, the practitioner can explore additional policy recommendations, gather information concerning budgetary implications, and the like in a more thorough

^{4/} Time constraints may also affect technical quality characteristics such as examining data for alternative explanations, the scope of the effort, and the like. For purposes here, however, only the linkage between time and useability potential is explored.

manner. Figure 6-3 diagrams one potential affect of funding and time resources on evaluation strategies that may affect technical quality and useability potential characteristics.

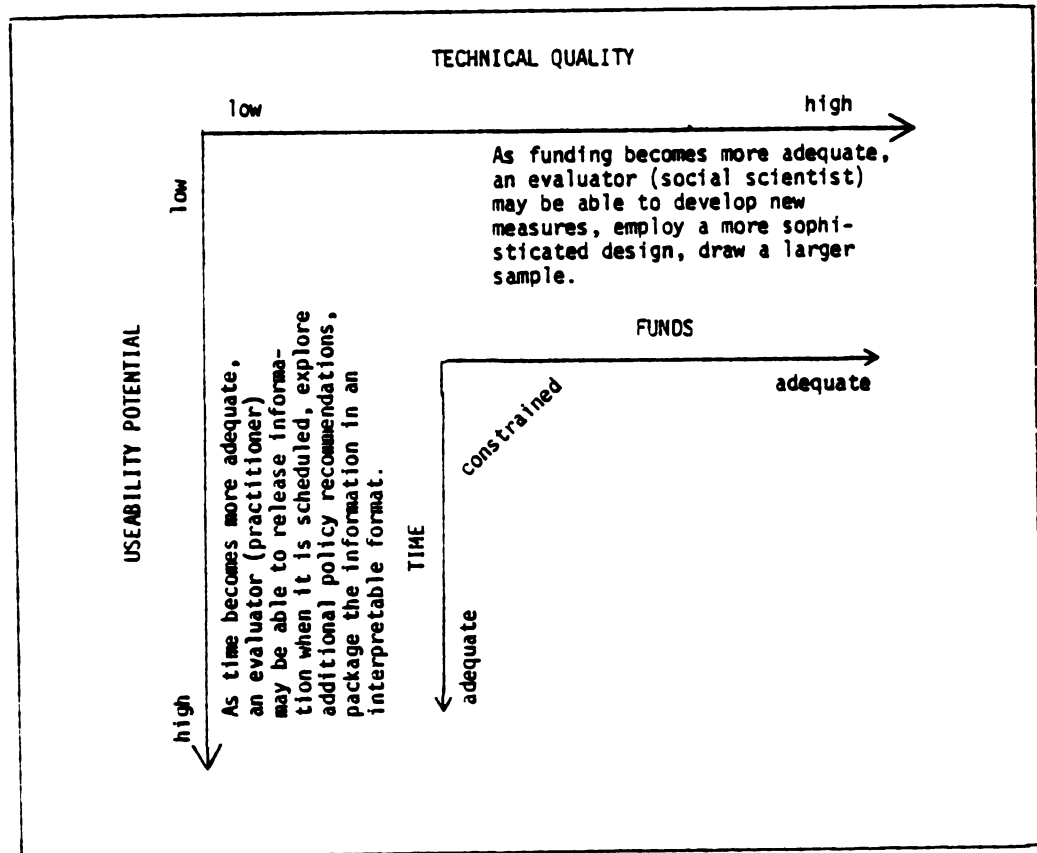


Figure 6-3. Resource Variables:
Technical Quality and
Useability Potential
Characteristics

What is suggested here is that as funding becomes more adequate, an evaluator, particularly a social scientist can better pursue strategies that increase the technical quality of the final product; as time becomes more adequate, the practitioner can better pursue research strategies that enhance the useability potential of the final product.

While the affect of time and fund constraints posited is admittedly partial and speculative, it provides one initial point for investigating resource constraints on strategies that a rational evaluator attempts to pursue. When turning to the constraints imposed by the evaluation environment, there is more guidance in the literature concerning environmental affects on evaluation strategies.

The Evaluation Environment--Constraints on Rational Strategies

The environment for implementing a program evaluation effort may also constrain or enhance an evaluator's research strategies. The evaluation environment, as defined here, refers to the amount of tension or cooperation encountered during the conduct of the evaluation effort. A program manager's cooperation during an evaluation effort is often essential in order for an evaluator to pursue strategies that may enhance technical quality and useability potential characteristics.

A program manager often controls access to program data and to clients, a necessary condition for producing an evaluation product that possesses some degree of validity and reliability.

In a tense evaluation environment, a program manager employs strategies such as obstructing evaluator access to this data. Even though an evaluator may possess skills and an orientation that should produce a high quality evaluation, unsuccessful attempts to collect data will not allow him to pursue research strategies that may enhance the full amount of technical quality possible.

Tension between an evaluator and a program manager may also lead to a final evaluation product lacking useability potential when compared to an evaluation executed under more favorable circumstances. In a tense environment, a program manager may attempt to obscure policy issues and to substitute tangential policy issues for more important ones; he may also withhold cooperation in drafting feasible recommendations for action. An evaluator often needs guidance from program personnel in order to frame and to package the evaluation product in a useable format. Thus, an evaluator, particularly one with social science training and norms, is susceptible to a program manager's sabotage efforts.

While tension between a program manager and an evaluator may constrain severely rational evaluation strategies, cooperation allows an evaluator to pursue research strategies that may enhance technical quality and useability potential characteristics given his skills, professional orientation, and organizational setting. A cooperative evaluation atmosphere results in a program manager sharing program insights with an evaluator and affording evaluator access to clients and to data. Although the actual

amount of technical quality found in the final product is dependent, in part, on an evaluator's methodological skills, cooperation allows an evaluator to pursue research strategies that may enhance technical quality. When a program manager is not threatened by an evaluation, he is less likely to thwart data collection efforts, to restrict findings and an evaluator is freer to disseminate findings for wide critique and review.

Useability potential is enhanced as well by a cooperative evaluation environment. As cooperation increases, a program manager is more likely to direct an evaluator to important policy concerns, to offer suggestions for politically feasible recommendations, and to aid evaluators in drafting the final report in a policy readable format. Thus, as cooperation increases, an evaluator may pursue research strategies that tend to enhance the useability potential of the final product. Figure 6-4 diagrams the potential affect the degree of tension in the evaluation environment may make on evaluation strategies which contribute to technical quality and useability potential characteristics.

Mixtures of Technical Quality and Useability Potential Characteristics

The model of the evaluation enterprise identifies important supply aspect variables while the rational choice assumptions made link these variables to the possible mixtures of technical quality and useability potential found in the final evaluation product. In order to examine the joint effects of the institutional

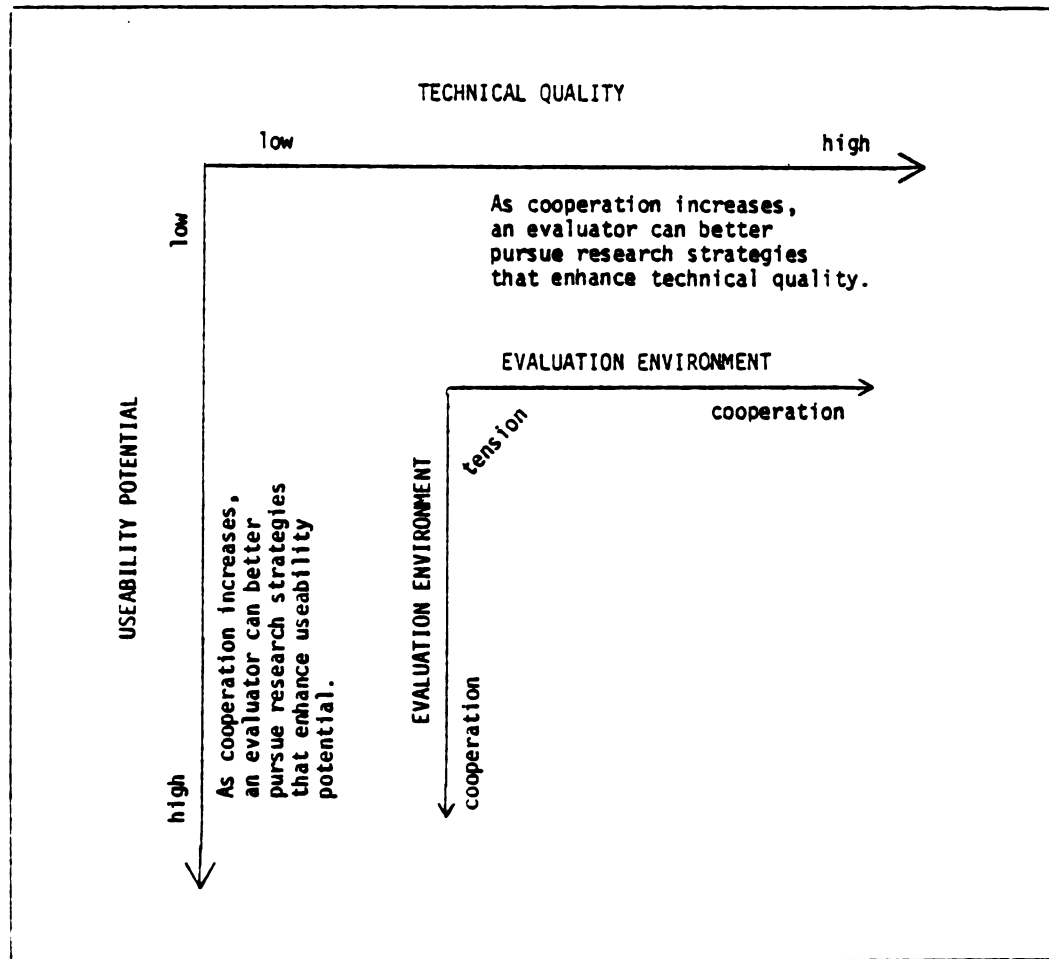


Figure 6-4. The Evaluation Environment:
Technical Quality and
Useability Potential
Characteristics

arrangement, resource constraints, and the evaluation environment on technical quality and useability potential characteristics, each supply aspect variable and its possible linkage to product characteristics was examined. These characteristics are perhaps in large part, of evaluator skills and orientations as well as the organizational incentives under which the evaluator implements the evaluation effort (see Figure 6-2).

An evaluator defined as a social scientist possesses methodological skills and attempts to pursue evaluation research strategies that maintain or enhance his reputation in the scientific community. A practitioner, on the other hand, possesses policy skills and attempts to maintain or enhance his professional reputation in the program community. While the strategies a social scientist attempts to pursue tends to enhance technical quality and the strategies a practitioner attempts to pursue tends to enhance useability potential characteristics, organizational incentives and sanctions may shape or impinge on these strategies. Here, the professional orientation of the evaluation unit administrator coupled with his concern in terms of maintaining his position in the evaluation unit may shape evaluation unit reward structures.

The evaluation research strategies an evaluator attempts to pursue are not only shaped by his professional orientation and organizational setting, resource and environmental constraints also determine, in part, the behavior of the evaluator. Here, it is suggested that as funding becomes more adequate, a social scientist is able to pursue strategies such as sampling larger

populations and developing new measures which, in turn, may increase the technical quality of the final product (see Figure 6-3).

The evaluation environment may also shape and impinge on an evaluator's research strategies. Here, it is suggested that a cooperative evaluation environment enables an evaluator to pursue strategies that may enhance both technical quality and useability potential, given the skills, orientation, and organizational setting of the evaluator; a tense evaluation environment, however, tends to constrain evaluation research strategies and may result in lower levels of technical quality and useability potential in the final product (see Figure 6-4). Based on this reasoning, a matrix of useability and technical quality mixtures expected given different institutional arrangements, resource constraints, and evaluation environment conditions is proposed (Proposition 6.2).

This proposed matrix is offered in an effort to extend previous work and direct research efforts towards exploring the joint affects of supply aspect variables on technical quality and useability potential--two characteristics which may affect the amount of influence a program evaluation product makes during a governmental agent's decision calculus. The remainder of this Chapter turns to explore linkages between the demand and supply aspect variables in the evaluation enterprise model. Here possible relationships between various contextual variables and the choice of an evaluation unit, the allocation of resources and the evaluation environment are suggested.

Proposition 6.2

TECHNICAL QUALITY

	low	medium	high
low	<p>INSTITUTIONAL ARRANGEMENT Internal Entrepreneurial</p> <p>RESOURCES Funds - Severe Constraints Time - Severe Constraints</p> <p>EVALUATION ENVIRONMENT Extreme Tension</p>	<p>INSTITUTIONAL ARRANGEMENT Entrepreneurial Governmental</p> <p>RESOURCES Funds - Some Constraints Time - Severe Constraints</p> <p>EVALUATION ENVIRONMENT Tension</p>	<p>INSTITUTIONAL ARRANGEMENT Academic</p> <p>RESOURCES Funds - Adequate Time - Severe Constraints</p> <p>EVALUATION ENVIRONMENT Some Tension</p>
medium	<p>INSTITUTIONAL ARRANGEMENT Internal Entrepreneurial</p> <p>RESOURCES Funds - Severe Constraints Time - Some Constraints</p> <p>EVALUATION ENVIRONMENT Some Tension</p>	<p>INSTITUTIONAL ARRANGEMENT Entrepreneurial Governmental Academic</p> <p>RESOURCES Funds - Some Constraints Time - Some Constraints</p> <p>EVALUATION ENVIRONMENT Some Cooperation</p>	<p>INSTITUTIONAL ARRANGEMENT Governmental Academic</p> <p>RESOURCES Funds - Adequate Time - Some Constraints</p> <p>EVALUATION ENVIRONMENT Cooperation</p>
high	<p>INSTITUTIONAL ARRANGEMENT Entrepreneurial Internal</p> <p>RESOURCES Funds - Severe Constraints Time - Adequate</p> <p>EVALUATION ENVIRONMENT Some Cooperation</p>	<p>INSTITUTIONAL ARRANGEMENT Governmental Entrepreneurial</p> <p>RESOURCES Funds - Some Constraints Time - Adequate</p> <p>EVALUATION ENVIRONMENT Cooperation</p>	<p>INSTITUTIONAL ARRANGEMENT Governmental</p> <p>RESOURCES Funds - Adequate Time - Adequate</p> <p>EVALUATION ENVIRONMENT High Cooperation</p>

U S P O
S E A B I L I T Y
T E N T I A L

DEMAND ASPECT LINKAGES TO
THE ALLOCATION OF RESOURCES,
CHOICE OF AN EVALUATION UNIT,
AND THE EVALUATION ENVIRONMENT

The contextual variables are linked to the implementation of an evaluation effort in terms of the allocation of resources and the selection of an evaluator to conduct the program evaluation effort. In addition, the contextual variables affect another important supply aspect variable--the environment for conducting the evaluation effort. This section relies on and extends the reasoning offered in previous chapters in order to explore some possible linkages between contextual variables and the allocation of resources, the choice of an evaluation unit, and the evaluation environment.

Allocation of Resources

During the decision to evaluate a public program, the evaluation sponsor allocates two major resources needed for conducting the effort: Funds and time. The amount of funds and time allocated, given the scope of the evaluation problem addressed, serve as constraints when producing the final evaluation product.

Funds. When making the decision to evaluate a public program, a governmental agent weighs the costs and potential benefits involved in producing the information. ^{5/} One of the costs

^{5/} Chapter Three discusses the decision to sponsor an evaluation in more detail.

considered involves the amount of funds needed to conduct the evaluation effort; the major potential benefit considered by the governmental agent is program information that can reduce his own uncertainty or alter the uncertainty calculations of other actors. Thus, the amount of funds allotted may depend, in part, on the amount the governmental agent is able and willing to spend in an effort to alter or to reinforce uncertainty calculations.

The degree of uncertainty about a program strategy may be an important factor related to the adequacy of funds allocated for a formal program evaluation effort. When there is a high degree of uncertainty present, the potential influence of information may be greater than when there appear to be few doubts about a program strategy. That is, an evaluation sponsor may be more willing to spend the funds needed to conduct adequately the proposed evaluation effort when (1) he experiences a high degree of uncertainty about the program strategy and/or (2) he feels other actors involved with program decisions are highly uncertain about program strategies. One demand variable that may reflect this amount of program uncertainty in the proposed model is the type of decision implied by the evaluation effort. Figure 6.5 depicts this demand variable and the degree of program uncertainty attached to various program choices pending or implied by the program evaluation effort.

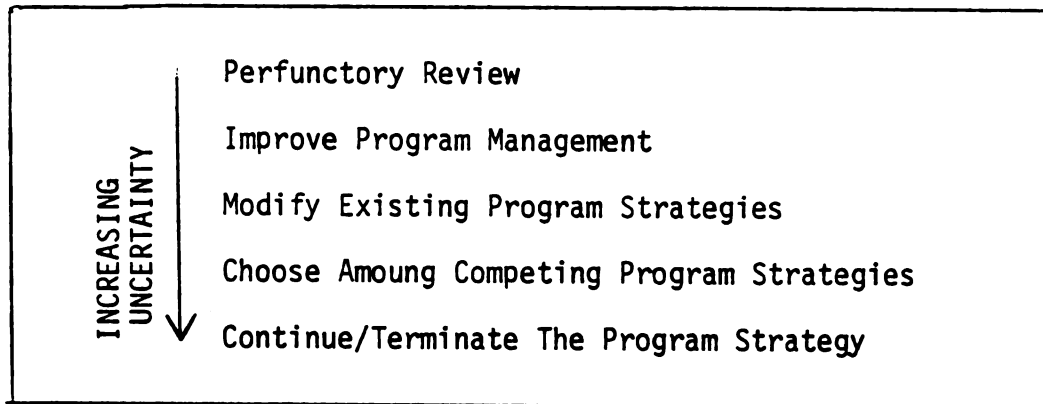


Figure 6-5. Type of Program Decision Implied:
Increasing Uncertainty

An evaluation effort that implies a routine check of program operations reflects the least amount of uncertainty. Here, the evaluation effort is based more on the traditions established by budgetary procedures or enabling legislation than on emerging doubts about the program (eg., Floden & Weiner, 1978; Rein & White, 1977). In this situation, the evaluation sponsor may allocate minimal funds for the effort since the potential benefit in terms of altering uncertainty calculations is relatively small. There is more doubt and uncertainty about a program evaluation strategy when an evaluation is sponsored in order to modify current program operations or to improve program management. An evaluation sponsored in an effort to generate program information bearing on a decision to choose among competing program strategies or to continue/terminate the current strategy often signals a high degree of uncertainty and doubt present (eg., Suchman, 1972;

Weiss, 1975). In this situation, information may have more potential influence on a governmental agent's program choice. Thus, the evaluation sponsor, in an effort to influence these choices, may be more receptive to funding requests for the evaluation effort.

Proposition 6.3 suggests that when program uncertainty is not at a particularly high level, as indicated by a routine type of program review, an evaluation sponsor is less likely to be responsive to requests for money on the basis that the funding may be inadequate to address the evaluation problem at hand; conversely, the greater the program uncertainty implied by the decision to evaluate, such as a possible termination decision, an evaluation sponsor may be more likely to be responsive to funding requests and allocate adequate funds for the task.

Proposition 6.3

The more the program decision underlying the demand for and sponsorship of an evaluation implies uncertainty about a program, the more likely the evaluation sponsor will allocate or be responsive to demands for adequate funds.

This proposition emphasizes the relative nature of funding. That is, rather than examine the absolute funds allocated, attention is directed towards examining funds allocated given the scope of the evaluation question addressed--an important constraint on research strategies pursued by an evaluator. One area receiving little empirical investigation concerns the costs associated with program evaluation efforts. It seems that

the production costs of evaluations may increase if new data must be collected, current staff personnel cannot execute the evaluation effort, the data must be collected from a widely dispersed population, and the like. While there is some information about the absolute costs involved with an evaluation effort based on budget documents, little work has been done comparing the costs involved with different types of evaluations and, in turn, gauging the adequacy of funds allotted.

Time. Time is another resource allocated for an evaluation effort during the demand stages of the evaluation enterprise which serves as a constraint when implementing the evaluation effort. While time is often cited as a constraint on an evaluator's research efforts, few have examined or analyzed the variables which may affect the allocation of this resource variable. Time, like funds, is viewed here in terms relative to the task at hand rather than the absolute period allocated. A demand variable in the evaluation enterprise model which may be one important factor affecting the adequacy of time allotted is the form of the evaluation demand--a requirement or a request.

When an evaluation demand is in the form of a requirement, often the parameters concerning when the information is needed is relatively known. Required evaluations are often tied to budgetary cycles or legislative enactments with predictable target dates. With a required evaluation, the program manager can anticipate the evaluation effort and may be better able to devote time to aid evaluators.

On the other hand, a requested program evaluation may catch the program manager by surprise. In this case, he may not be able to devote time already budgeted for other activities to aid the evaluation effort. In addition, a requested evaluation often reflects a programmatic crisis or increasing doubts about a program tied to a pressing program decision at hand. If a rather pressing program decision is involved, then it seems more likely that the time period for an evaluation effort may be inadequate for the problem at hand. In order to generate information which can be injected into a governmental agent's decision calculus and bear on his program choice, fairly severe time constraints may be imposed. The following proposition suggests a relationship between one variable identified as a demand variable and time constraints imposed on an evaluation effort.

Proposition 6.4

When the demand for an evaluation is in the form of a request, the time allotted for the evaluation effort may be less adequate for the task than when the demand is in the form of a requirement.

This proposition, while admittedly speculative and partial, highlights the need for additional empirical work investigating the time needed for implementing various types of evaluation efforts: needs assessments, process monitoring studies, and impact studies. The variables affecting the allocation of resources as well as the dynamics of this process is an area that is relatively unexplored in the evaluation literature.

Analysis focused here on only two major demand variables, type of program decision pending and form of the demand, in an effort to initiate further speculation and research concerning the allocation of resources.

The Choice of An Evaluation Unit

Another aspect of demand central to the production of an evaluation product is the choice of an evaluation unit. This choice not only affects technical quality and useability characteristics, but is also related to two other product characteristics scrutinized by a governmental agent during his decision calculus--the producer and the nature of the findings. Yet, the evaluation literature has paid little attention to the variables or to the dynamics of this selection process. While a number of variables identified among the contextual variables may affect this choice, analysis is limited here primarily to one key demand variable in the evaluation enterprise model--the evaluation sponsor and his motives underlying the sponsorship of an evaluation. However, the type of program decision implied by an evaluation effort, another demand variable, will also be considered.

When examining the motives stimulating the decision to evaluate a public program from a rational choice perspective, two major purposes for demanding and sponsoring an evaluation effort were offered: Dissatisfaction and/or puzzlement about a program and generating evidence to support or oppose a program. ^{6/}

^{6/} See Chapter Three Three for a more complete discussion.

These motives may be linked to the amount of autonomy an evaluation sponsor may be willing to give an evaluator. In turn, this autonomy may affect the nature of the findings contained in the final program evaluation product.

When an evaluation sponsor is highly uncertain about the outcomes associated with various program strategies, he may be more willing to select an evaluator that is insulated from demands made by the program manager or sponsor himself concerning the nature of the program findings which should be generated. However, when an evaluation sponsor entertains few doubts about the program and sponsors an evaluation effort in order to generate evidence to alter the uncertainty calculations of other governmental agents, then it may be more likely that he will select an evaluator that is responsive to his demands. This may be particularly true with respect to demands made concerning the nature of the findings contained in the final evaluation product.

Proposition 6.5

The evaluation sponsor's motives for conducting an evaluation effort affects, in part, his choice of an evaluation unit.

- (a) An evaluation sponsor motivated primarily by dissatisfaction or puzzlement may be more willing to consider the range of institutional arrangement options and to be more willing to employ an evaluation unit which is autonomous from evaluation sponsor and program manager demands.
- (b) An evaluation sponsor motivated primarily by generating positive or negative program evidence is less likely to consider the range of institutional arrangement options and is more likely to select an evaluation unit who is responsive to his demands.

In terms of the four institutional arrangements discussed, there is some suggestion in the literature that internal and entrepreneurial evaluation units may be more responsive to evaluation sponsor demands placed on their efforts than the more autonomous governmental and academic evaluation units (eg., Coleman, 1972; Suchman, 1972; Weiss, 1972b; Pressman, 1975; Caro, 1971a; Bernstein & Freeman, 1975). An evaluation sponsor motivated primarily by dissatisfaction and/or puzzlement about a program may be more willing to consider all four types of institutional arrangements when deciding to employ an evaluator and more willing to employ an evaluation unit that remains relatively autonomous from evaluation sponsor and program manager demands.

Perhaps, however, as the program choice implied by the evaluation effort becomes more threatening, the evaluation sponsor may become more concerned with the issue of institutional arrangement autonomy. That is, when the type of program choice implied is rather threatening, as in the case of a pending decision to terminate possibly the current strategy, then the evaluation sponsor may be more concerned with insulating the evaluation effort from possible cooption by a program manager attempting to produce information that reflects his self-interest. In this case, then, the evaluation sponsor may select an institutional arrangement that is autonomous such as a governmental or an academic evaluation unit. The choice between these types of arrangements may be

a function of resource constraints: If a governmental evaluation unit has previously scheduled evaluations and the information is needed in a short period of time, the evaluation sponsor may opt to hire an academic evaluation unit; if additional funds are not allocated for the evaluation effort, then the evaluation sponsor may employ a governmental unit.

While the evaluation sponsor motivated by puzzlement and/or a high level of uncertainty may consider all the various options, the evaluation sponsor motivated by generating evidence may tend to limit his selection to institutional arrangements more susceptible to his demands--an internal or entrepreneurial evaluation unit. Here, the choice may not only be a function of time and resource constraints: If the evaluation sponsors desires to generate positive program evidence, then he may choose either an internal or entrepreneurial evaluation unit. Both, it seems, may be more susceptible to his demands in terms of the nature of the findings expected than a governmental or an academic evaluation unit. However, if the evaluation sponsor desires to generate negative program evidence, he may attempt to employ an entrepreneurial firm whose loyalty lies not in the program but with the evaluation sponsor.

While Proposition 6.5 focuses mainly on the level of uncertainty and motives stimulating the demand for an evaluation and is admittedly speculative, it directs research efforts to an area where there is little discussion or empirical work--the variables affecting and the dynamics involved with the choice of an

evaluation unit. While research examining the choice of the evaluation unit tends to be neglected by evaluation scholars, it seems that continued refinement and identification of additional important variables involved with this choice procedure would enhance ultimately the understanding of utilization. The type of evaluation unit selected is important in terms of the organizational setting, evaluator skills, and orientations which shape research strategies affecting directly the technical quality and useability potential of the final product. In addition, the choice of the unit in terms its reputation and autonomy from evaluation sponsor and program manager demands may directly affect two other characteristics scrutinized by a governmental agent when considering the information: The producer and the nature of the findings.

The Evaluation Environment

The final linkage between demand and supply aspect variables explored centers on the evaluation environment--the milieu between evaluators and program managers during a program evaluation effort. Unlike the allocation of resources and the choice of an evaluation unit, there is more guidance in the literature concerning possible linkages between contextual variables and the environment for conducting a program evaluation effort.

The program manager plays a pivotal role during an evaluation in terms of access to agency records and clients, guiding evaluators to feasible policy recommendations, and the like. It is posited

here that the degree of cooperation or tension present during the evaluation process is dependent on the amount of threat posed to the program manager by the evaluation effort. In turn, the evaluation environment serves as one constraint on the evaluation strategies that an evaluator pursues.

Assumption 6.3

The degree of cooperation or tension in the evaluation environment depends, in large part, on whether or not a program manager's self-interest is threatened.

If a program manager is threatened by an evaluation effort, (i.e., his security and the program's is threatened), then he may attempt to thwart an evaluator's efforts by employing strategies such as obstructing access to program data or by attempting to sidetrack the evaluator to tangential program issues. When a program manager's self-interest is less threatened by the evaluation effort, cooperation between the evaluator and program manager is more likely to result. When producing the information in a cooperative evaluation environment, the program manager tends to pursue strategies such as affording an evaluator access to clients and to data as well as sharing his program insights with the evaluator.

It is suggested here that a number of demand and organizational variables affect the evaluation environment. In addition, the degree of institutional arrangement autonomy, a production of information variable, may also contribute to the milieu for

conducting the program evaluation. Each of the variables listed in Proposition 6.6 is discussed briefly in terms of the potential tension generated between the evaluator and program manager during an evaluation effort.

Proposition 6.6

EVALUATION ENVIRONMENT = f [Evaluation Sponsor, Form of Demand, Type of Decision Implied, Program Funding Mixture, Type of Agency, Program Manager Characteristics, Institutional Arrangement Autonomy]

> TENSION = f [elected decision-maker, request, threatening decision, operating agency, trapped program manager, incongruent sponsor and funding, academic evaluation unit]

> COOPERATION = f [program manager sponsor, required, congruent funding source and sponsor, non-threatening decision, research agency, experimental program manager, internal evaluation unit]

The Evaluation Sponsor. The evaluation sponsor is one demand variable that may impact on the evaluation environment. If the self-interest of the evaluation sponsor and program manager conflict in terms of the proposed evaluation effort, tension may increase. For example, an elected decision-maker may sponsor an evaluation as a result of constituency pressures that express dissatisfaction with current program approaches. In order to produce information concerning program operations as well as to placate constituent

criticisms, the elected decision-maker sponsors a formal program evaluation. A program manager, interested in protecting his program's security and his own position, may be threatened by the prospects of this evaluation review. In this situation, then, a program manager may attempt to thwart an evaluator's research efforts, particularly if negative program findings are suspected. Thus, a tense evaluation environment may result.

However, if the self-interest of the evaluation sponsor and the program manager coincide, a more cooperative atmosphere for conducting the program evaluation may result. When the program manager is the evaluation sponsor, the evaluation environment also tends to be cooperative.

Proposition 6.6a

If an elected decision-maker sponsors an evaluation, a more tense evaluation environment results than when the program manager sponsors the program evaluation effort.

Form of Evaluation Demand. A required evaluation often poses relatively little threat to the security of the program. Enabling legislation or budgetary procedures may require periodic program evaluations. Since a program manager anticipates these required evaluations, program operations are not disrupted and cooperation may result. Part of this cooperative environment may be due in part to the program manager's structuring of the evaluation and timing program events in a manner that may shed favorable light on the program. Since a program manager can often anticipate required

evaluations, he is in a position to attempt manipulating evaluation parameters such as hiring evaluators who are sympathetic to program issues and practices; or, at a minimum, he may attempt to influence the choice of the evaluation unit. While a required evaluation gives a program manager prior notice and may lead to a cooperative evaluation environment, a requested evaluation may create more tension.

There are a number of potential sponsors that may request a formal program evaluation. Often, dissatisfaction with current program operations stimulate the demand for the evaluation: budgetary constraints may dictate reordering program priorities; dissatisfied clients may lobby elected decision-makers and policy administrators for change; a programmatic crisis may spur requests for a formal evaluation of the program. Unlike a required program evaluation effort, a requested one often catches the program manager and his staff by surprise. In turn, this unanticipated evaluation may disrupt program operations and may focus on program aspects that the program manager has neglected. A requested evaluation, then, can create friction and a tense environment for conducting the evaluation effort.

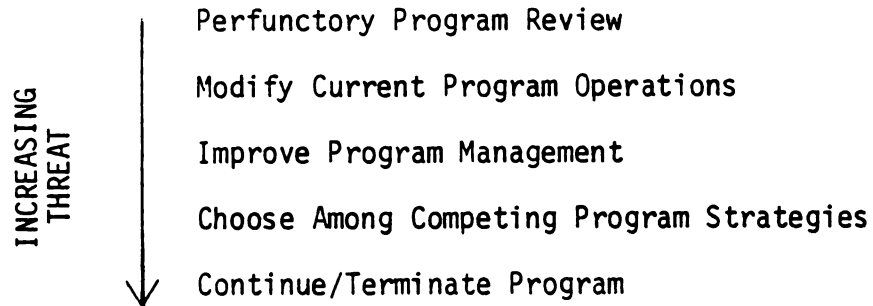
Proposition 6.6b

A required evaluation generally produces a more cooperative environment while a requested evaluation generally produces a more tense environment for conducting a program evaluation.

Type of Program Decision Implied. Another variable linked, in part, to the amount of tension permeating the evaluation environment is the type of program decision pending or implied by the evaluation effort (eg., Floden & Weiner, 1978). Given that program managers are concerned with maintaining their programs and their bureaucratic positions, the more threatening the decision is to a program manager, one would expect that greater tension is apt to result. A perfunctory review of program operations and impact may be least threatening to a program manager among possible reasons for evaluating a program. If an evaluation is sponsored in order to discover ways to modify current program strategies or to improve program management, one would still expect an evaluation context characterized by some cooperation between an evaluator and a program manager. However, evaluations that imply choice among competing program approaches or imply possible program termination create conditions for a tension-filled environment. Thus, a program manager, in an attempt to protect his self-interest, may withhold cooperation and attempt to sabotage evaluation research efforts so that the final product may be heavily discounted by other governmental agents.

Proposition 6.6c

As the decision implied by an evaluation effort becomes more threatening, tension in the evaluation environment increases.



Program Funding Mixtures. A program manager, in order to retain his position in the bureau and the program, must respond to evaluation demands placed upon him by the level of government that funds him. Even though a threatening decision may be implied by an evaluation effort, the program manager must cooperate to some degree with evaluators lest he loses his job. However, a program manager may view evaluation as a nuisance and be less cooperative if the demand for the evaluation arise from a source that does not directly fund program activities.

Proposition 6.6d

If the evaluation sponsor and the funding source is congruent, the evaluation tends to be more cooperative than when the sponsor and funding source is incongruent.

Type of Program Agency. Whether a program is embedded in an agency that is primarily an operating or a research agency may affect the program manager's predispositions concerning an evaluation effort. An operating agency is concerned mainly with delivering services to clients. Thus, any activity that disrupts program operations may be viewed as an annoyance and as dysfunctional to the program. Furthermore, personnel in an operating agency are likely to reflect norms and values of practitioners as opposed to those norms held by social scientists. Evaluators, particularly those with strong social science orientations, speak an unfamiliar language and often deal with issues of little concern to program personnel such as publishing research. This unfamiliarity may heighten suspicion and produce tension between an evaluator and a program manager (eg., Coleman, 1972; Pressman, 1975).

Individuals with social science orientations and familiarity with research activities often staff a research-oriented agency. Thus, a program manager operating within this organizational context is apt to be less threatened by the prospects of a formal evaluation. A program manager in a research agency is apt to speak social science language and be able to anticipate evaluation research strategies. Thus, he may be able to maintain better control over the evaluation endeavor than his operating agency counterpart.

Proposition 6.6e

A research agency tends to produce a cooperative evaluation environment while an operating agency tends to produce a tense evaluation environment.

Program Manager Characteristics. A trapped program manager is one whose political situation does not allow him to risk failure. Thus, a trapped program manager, whose security is threatened by an evaluation, is likely to obstruct evaluation attempts and tension may result between the program manager and the evaluator. The experimental program manager is the antithesis of the trapped manager (Campbell, 1965). Unlike the trapped manager, the experimental program manager is more willing to take risks. He is committed to the importance of solving the problem more than he is wed to current program approaches. The experimental manager sees his bureaucratic security resting on solving the particular problem the program addresses rather than by pursuing a current program strategy. Thus, the experimental program manager is more likely to be supportive of evaluation activities, especially if they will produce recommendations that result in improving a program's effectiveness in solving a problem.

Proposition 6.6f

An experimental program manager tends to be more cooperative during an evaluation effort than a trapped program manager.

Institutional Arrangement Autonomy. Those institutional arrangements that allow a program manager control over evaluation research strategies may be less threatening than institutional arrangements that provide evaluators autonomy from program manager demands. Institutional arrangements that are closely supervised by program managers may contribute to a cooperative environment while arrangements that afford evaluators autonomy from the program manager may contribute to a tense environment.

The four institutional arrangements may be ranked in terms of the amount of potential control a program manager may over research efforts. Other variables, particularly the sponsor of the evaluation, affect the actual amount of control that a program manager retains during the course of the program evaluation. However, if a program manager has input into structuring evaluation research strategies, certain institutional arrangements may be more conducive to program manager controls than others.

Evaluators who are attached to the program are likely to avoid sensitive areas lest they lose their jobs. In addition, a program manager and his staff may be least suspicious and threatened by evaluators housed within the program when compared to the other types of institutional arrangements. Thus, an evaluation conducted by an internal evaluation unit should contribute to a cooperative evaluation environment.

An entrepreneurial evaluation unit may also lead to a cooperative evaluation environment. Entrepreneurial evaluators, concerned with future contracts for organizational survival, may be susceptible

to demands and to controls placed on them by a program manager. Yet, since the entrepreneurial evaluation unit is located outside of the program, there may be slightly more tension created than found when an internal evaluation unit conducts the research.

The situation may be different for evaluators who are attached to a governmental evaluation unit or who are part of an academic evaluation unit. These institutional arrangements afford evaluators more autonomy from a program manager. Thus, more tension between the evaluator and program manager may result. A program manager has the least amount of potential control over academic evaluators since they are interested in pursuing research strategies that maintain their university positions as well as protect or enhance their professional reputations. Thus, academic evaluators are likely to bolt from a program manager's suggestions and directives if they conflict with their self-interest.

Proposition 6.6f

As the autonomy of the institutional arrangement increases, the tension in the evaluation environment increases as well.

Internal Evaluation Unit
 Entrepreneurial Evaluation Unit
 Governmental Evaluation Unit
 Academic Evaluation Unit



Summary. By identifying variables affecting the evaluation environment, Proposition 6.9 links a number of variables together that are often treated as separate issues in the program evaluation literature. While there have been discourses lamenting the fact that a program manager often disrupts evaluation activities, the model of the evaluation enterprise identifies systematically contextual and production of information variables that may contribute to tension or cooperation during the conduct of an evaluation. Given that the evaluation environment impairs or enhances research strategies affecting both technical quality and useability potential characteristics, systematic research into the weight of each variable in determining the degree of tension or cooperation would be instructive. Based on additional empirical work, evaluation sponsors could attempt to manipulate variables which could ease tension and enhance the technical quality and useability potential of the final program evaluation product.

CONCLUSION

The model of the evaluation enterprise proposed, based on variables cited in the literature, provides a framework for exploring possible linkages among the demand, supply, and consumption aspects. When exploring possible linkages between the supply and consumption aspects, analysis focussed on institutional arrangement reward structures, resource, and environmental constraints on an evaluator's strategies which may shape the technical quality and useability potential of the final product. Based on the assumptions posited and the analysis of each supply variable on these characteristics, a matrix of technical quality and useability potential characteristics was proposed.

When exploring possible linkages between demand and supply aspect variables, propositions concerning the relation between contextual variables and the choice of an evaluation unit, allocation of resources and the evaluation environment were offered. While this discussion was restricted in terms of the contextual variables examined and often speculative, the propositions generated provide initial direction for research in areas that remain relatively unexplored.

The proposed model, although continued refinement may occur, offers a number of advantages in terms of linking variables often treated in isolation by evaluation scholars. As an integrative device, the model serves to clarify possible linkages; as an

interpretive device, the model and logical underpinnings serve to describe and explain empirical findings; as a predictive device, the model generates a number of propositions for empirical testing. A conceptual and systematic approach to the evaluation enterprise has been taken in an effort to develop logical underpinnings, clarify linkages, and generate testable propositions. Models of the program cycle, the program decision calculus, the evaluation enterprise, and testable propositions associated with each aspect were formulated based on the rational choice perspective taken. While this approach provides insight into and organization of the program evaluation literature, the propositions generated must ultimately undergo empirical scrutiny. The concluding Chapter reports on a pilot study conducted to judge the feasibility of research, evaluate the operationalizations suggested, and provide a preliminary assessment of the propositions.

CHAPTER SEVEN

THE EVALUATION ENTERPRISE: A PILOT STUDY

As the field of program evaluation has evolved, so too has the literature discussing the various issues and topics involved in the evaluation enterprise. While this literature has given insight into various aspects of program evaluation, it tends to be fragmented and issue-oriented. In an effort to address this perceived shortcoming, this dissertation developed a systematic and conceptual approach to the evaluation enterprise. By making plausible assumptions, constructing simplified models, and developing some logical underpinnings, propositions for future research directions have been offered. While the primary emphasis of this dissertation is conceptual, the utility of the approach lies not only in terms of the plausibility of its assumptions and ability to generate propositions, but must also be judged in terms of its testability. Thus, a research strategy is proposed (Appendix A) and measures developed (Appendix B) in order to test the propositions. This Chapter reports the results of a pilot study conducted.

The pilot study focussed on only one program evaluation product. There are a number of benefits as well as limitations to the pilot study conducted. The benefits include gauging the practicality of the research strategy and examining the

measures developed before launching a wide-scale research effort. By implementing the research strategy, potential problems and obstacles can be identified and rectified before widening the scope of the research. The pilot study is also beneficial in checking the wording and the ordering of questions designed to measure various aspects of the evaluation enterprise. A preliminary assessment of biased questions, measures, and variability in responses can be made.

The major limitation of the pilot study lies in its inability to adequately test hypotheses. One case can hardly support or disconfirm a proposition. In order to test the hypotheses, a large number of cases (program evaluation products) need study. However, the information gathered may lend some initial insight into the credibility and feasibility of the rational choice perspective towards the evaluation enterprise.

THE PILOT STUDY

In general, the research strategy focuses on a program evaluation product as the unit of analysis and relies on interviews with a number of individuals involved with various aspects of the evaluation effort. In order to obtain information bearing on the demand aspect propositions and contextual variables, the evaluation sponsor must be interviewed. This information may also be obtained from the program manager. Interviews with the program manager and evaluator are essential in order to gather information bearing on supply aspect propositions and variables such as the evaluation environment. Interviews with potential users, identified by positional and reputational methods, yield information bearing on consumption aspect propositions. In order to rate characteristics of the final product which may affect use, potential users are asked a series of questions. The program evaluation product is also rated in terms of technical quality and useability potential by the researcher.

Obtaining a sampling frame of program evaluation products is the first step in implementing the research. An inventory of program evaluation products was compiled for one relatively narrow area of human services. Interviews with department personnel and departmental documents yielded the inventory of program evaluation products presented in Figure 7-1.

INSTITUTIONAL ARRANGEMENT	IMPACT STUDY Completed	IMPACT STUDY Planned	PROCESS STUDY Completed	PROCESS STUDY Planned	COMPREHENSIVE STUDY Completed	COMPREHENSIVE STUDY Planned
Governmental Unit	6	6	4	8	1	3
Joint Governmental	-	-	-	2	-	-
Internal Unit	1	-	-	-	-	-
Contract	<u>1</u>	<u>5</u>	<u>-</u>	<u>-</u>	<u>-</u>	<u>1</u>
TOTAL	8	11	4	10	1	3

Figure 7-1. An Inventory of Completed and Planned Program Evaluations in The Department

One program evaluation product was selected from this sampling frame for the pilot study. In order to pretest questionnaires, an impact study possessing high visibility and current interest among state officials was selected. A general description of the program, evaluation setting and political context follows. ^{1/}

The Program--Evaluation Setting and Political Context

The program selected was a "demonstration project", an experimental program for delivering social services where current types of strategies implemented seem to have about the same rate of client success and failure. The demonstration project approach is common in many of the human service policies developed by the federal government. Like many demonstration projects, the federal government funded 90 percent of the program while the state matched 10 percent of the program costs. After the first two years, the state's share of the costs, under conditions of the initial project grant increased to 50 percent. After five years, the decision must be made whether to continue the program and finance it entirely with state funds.

The case study was conducted as the state approached the decision to continue and possibly expand the demonstration project with state funds or to terminate the program strategy. The total

^{1/} The political sensitivity of this program and the easy identification of those interviewed in a case study prevent citing specific facts and program names in order to insure confidentiality of the respondents.

costs of the program were relatively small given the total expenditures of the state, the department and the expenditures for traditional program services. ^{2/} There were no plans to greatly expand the size of the program. However, the political climate and controversy surrounding the program and the program evaluation effort became more intense as the decision for state financing of the program approached.

Although the total amount of money involved did not seem to be the issue, there was increasing scrutiny of "seed projects" and their implications for the total state budget if demonstration projects were continued with state funds after federal monies were withdrawn. Since there were numerous demonstration projects operating that could eventually compete for state appropriations, the state budget office examined the impact and effectiveness of demonstration projects carefully. Thus, demands for solid evidence of program effectiveness as well as the cost-effectiveness of the program strategy were made.

Evaluation of program effectiveness was a condition of the demonstration project grant. A governmental evaluation unit, housed within the program's department, conducted the program evaluation. However, research activities did not commence until about six months after the program initially started. One of the

^{2/} The program costs comprised .005 percent of the total state expenditures, .02 percent of departmental expenditures, and 2.0 percent of expenditures made in the program area.

major constraints for the evaluators in terms of designing the research was the inability to randomly select program participants and to establish a control group which received no program treatments. Legal restrictions as well as ethical considerations prohibited randomization. Thus, a quasi-experimental design, attempting to match program participants with other program clients receiving standard program services, was implemented. Outcome measures used were based on other indicators widely use in this human services area. However, the small number of program participants led to a number of difficulties during the analysis phases of the program evaluation effort. Preliminary reports of the findings were released throughout the duration of the program.

The final program evaluation product released was fairly open to wide interpretation: The findings indicated some support for the program's effectiveness. It noted that the new program strategy, while not experienceing much greater success with clients than current program approaches, was at least as cost-effective. However, the small sample size, lack of randomization, and relatively short period of time to gauge client success rates provided room for a number of alternative explanations for the evaluation findings.

When the program evaluation product was released, it served as a reference point during the debates and interactions among state legislators, the governor, the program staff, budget officials, interest groups, and program personnel concerning the fate of the demonstration project. Upon receiving this report, interviews

were scheduled. Nine interviews were conducted across an array of different types of governmental agents.

The Demand, Supply, and Consumption Aspects of The Evaluation Enterprise

When examining each aspect of the evaluation enterprise from a rational choice perspective, propositions suggesting new research directions were offered. The evaluation enterprise begins when demands for a formal assessment are made and a governmental agent, upon weighing the potential costs and benefits, decides to sponsor the effort (Proposition 3.1). Programs that have fairly uncertain outcomes (programs which are new and controversial, require relatively large expenditures, experience a programmatic crisis, and/or are unsuccessfully defended by the program manager) tend to be the targets of formal evaluative scrutiny (Propositions 3.2 and 3.3). Program evaluation products, however, are not the only packages of information produced which may bear on program strategies and choices.

Individuals in the program cycle and action context supply information in an attempt to influence decisions when their self-interest is involved. When making the decision to produce information, the individual weighs the potential damage to his self-interest if a detrimental program choice is made and the costs incurred when producing the information (Proposition 4.1). The total supply of information with which a program evaluation product must compete depends, in part, on the number of individuals whose self-interest may be threatened. Situations and conditions which may

threaten a number of individuals' self-interest and, in turn, generate a large supply of information include programs with highly uncertain outcomes, a threatening type of program choice pending, and/or the sponsorship of a formal program evaluation effort (Proposition 4.2).

When turning to the consumption aspect of the evaluation enterprise, then, a governmental agent facing a program choice is supplied with competing packages of information (Proposition 5.1). The program decision calculus model and its discounting procedure provide a framework for analyzing the amount of weight a program evaluation product is given vis-a-vis other program information. Characteristics of the product as well as the self-interest of the governmental agent determine, in part, the amount of influence a program evaluation product makes, given competing information processed, during the decision calculus process. In addition, the availability of the information when it is needed (Proposition 5.2), reducing a governmental agent's search costs (Proposition 5.3) and features reducing processing costs (Proposition 5.4) also may affect the amount of influence a program evaluation product makes.

While the pilot study is useful in assessing measures, a comparison among a number of program evaluation products would be needed to test the demand and supply aspect propositions. The interviews with a number of users, however, lends more credibility to a preliminary examination of the consumption aspect propositions.

Demand Propositions. The evaluation effort was required by the federal government as a condition of the initial grant. This program reflected two characteristics of programs which tend to come under evaluative scrutiny: (1) The program strategy was new and a controversial approach to modify client's behavior, and (2) there was an indication that if the program continued as an alternative to present strategies, long-term budgetary expenditures could be incurred by the state (Propositions 3.2 and 3.3).

In terms of the decision to sponsor an evaluation effort, the sponsor weighed the potential damage to his self-interest and the potential benefits derived from the information (Proposition 3.1). The potential benefit gained by the evaluation was information, hopefully objective, assessing the impact of the program before it became permanently entrenched or its scope expanded. The production costs involved, deploying an evaluator from the agency evaluation unit, seemed worth the potential benefits derived. ^{3/}

Supply Propositions. The supply hypotheses presented in Chapter Four also were given some preliminary support by the case study. Individuals whose goals were affected by the program choice produced information which reflected their self-interest

^{3/} Information assessing the demand propositions was collected by interviewing the program manager and the evaluation sponsor (see Appendix B, Survey B-2).

(Proposition 4.1). In addition, the total supply of information seemed quite large since a number of individuals were threatened by the choice to continue or terminate the program (Proposition 4.2).

Program information was produced by individuals who felt the program may be terminated and who generated information which supported the program efforts. Here, there were a number of testimonials by program participants, program endorsements from interest groups, and subjective assessments by program staff supplied to governmental agents. Individuals opposed to the program tended to generate information which justified terminating the program. Here, budget officials supplied budget data and individuals testified the program strategy, a radical departure from traditional approaches, was ineffective. ^{4/}

Consumption Propositions. When confronting the choice of continuing the program, there were a number of packages of information supplied in an attempt to alter or to reinforce a governmental agent's uncertainty calculations (Proposition 5.1). The program evaluation product was just one source of information. Yet, the amount of competing information entering the decision calculus varied across users. Most users interviewed agreed that the timing of the program evaluation product coincided with the pending decisions and the findings considered when making program recommendations (Proposition 5.2). Search costs were

^{4/} Information assessing the supply propositions was collected by interviewing potential users (see Appendix B, Survey B-2).

generally reduced since the program evaluation was disseminated widely (Proposition 5.3). An attempt was made to reduce processing costs by including an executive summary in the final report. For individuals without an information processor, however, processing costs were higher (Proposition 5.4).

The model of the program decision calculus provides a framework for analyzing the amount of influence a program evaluation product makes on a governmental agent's recommendations and program choices. The weight given to the findings and characteristics of the information which affect discounting seemed to vary across respondents. While it is difficult to infer patterns to the weights given based on the small number of interviews, there does seem to be a preliminary indication that the level of uncertainty about program outcomes coupled with organizational incentives does account for some of the patterns noticed. The program evaluation product increased as well as reduced doubts about the program's effectiveness in terms of client success.

For governmental agents entertaining few doubts about the program's success, the evaluation findings weighed little in their final recommendation. Other program information, such as testimonials and the like, weighed more heavily. However, governmental agents experiencing dissatisfaction or puzzlement about the program (a high level of uncertainty), tended to weigh the information more heavily when making their program recommendations. Here, the technical quality of the information became the most important factor in terms of weighing the information.

Governmental agents with some degree of certainty about the program outcomes yet entertaining doubts as well tended to give the program evaluation moderate weight. Here, political considerations and other program information explicitly entered the decision calculus. These governmental agents, unlike ones expressing greater program uncertainty, cited the useability potential, nature of the findings, or the producer as the most important factor affecting in terms of weighing the information.

Given the different self-interests and organizational settings, governmental agents tend to focus on different aspects of the findings. Information processors tended to focus on the technical quality characteristics while individuals most proximate to the program decision tended to focus on the nature of the findings and the useability potential of the information. ^{5/}

The Model of The Evaluation Enterprise

During the pilot study, information to test measures developed for the model of the evaluation enterprise variables was also gathered. Limitations of a case study methods prevent tests of specific hypotheses offered in Chapter Six. However, some preliminary results are reported.

Linkages between Supply Variables and Product Characteristics.

The institutional arrangement employed to conduct the evaluation given resource and environmental constraints may affect the mixtures

^{5/} Information assessing the consumption aspect propositions was collected by interviewing potential users (see Appendix B, Survey B-2).

of technical quality and useability potential found in the final program evaluation product (Proposition 6.2). There is some preliminary support given by the case study. A governmental evaluation unit conducted the program evaluation. The time given, in terms of releasing the information when it was needed, was adequate. However, there were some funding constraints imposed given the scope of the evaluation problem. If additional funds were allocated, the evaluator would be able to expand the number of clients entering the program and develop more elaborate tracking mechanisms in order to gauge program effectiveness. The final product released can be characterized as possessing fairly high useability potential and average technical quality.

Technical quality and useability potential characteristics were rated by the researcher as well as the users. Based on researcher ratings, the program evaluation product exhibited relatively high useability potential (.71) when compared to technical quality (.43). The average ratings by users also reinforced this pattern. With both methods used, useability potential was rated approximately one-and-a-half times higher than technical quality. ^{6/}

There also seems to be support for the notion that both social science and practitioner norms are stressed in a governmental

^{6/} The researcher index scores are the percentage of the total number of items scored. The users, on the average, rated technical quality as 2.8 and useability potential as 4.0 on a five-point index (see Appendix A).

evaluation unit (Proposition 6.1). Based on interviews with the evaluation unit administrator and staff as well as departmental memos, the reward structure was sensitive to the credibility of the research effort (social science orientation) as well as cautious not to jeopardize relationships with the program staff (practitioner orientation).

Allocation of Resources. During the demand aspects of the evaluation enterprise, funds and time are allocated to conduct the research effort. As the demonstration project was started, there was little consideration given requests for additional funds; the evaluation effort was viewed as a ritual to continue funding during the first couple of years. As the decision concerning possible program termination approached, however, the evaluation sponsor and program manager became more receptive to requests for additional funds and personnel to conduct the research (Proposition 6.3). Since a formal evaluation was required as part of the grant, time parameters for releasing the information were predictable and fairly adequate given the evaluation problem (Proposition 6.4).

The Choice of An Evaluation Unit. One of the most important factors affecting evaluation products is the choice of the evaluation unit. This choice affects not only technical quality and useability potential, but also bears on the producer and nature of findings characteristics scrutinized by a governmental agent. Here, the evaluation sponsor, motivated by puzzlement about expected program outcomes, was not as concerned with the nature of the

findings produced as with the ability of the evaluation unit to monitor program developments, gauge the effectiveness of the program, and release information when it was needed for decision-making (Proposition 6.5). Due to funding constraints and considering the fairly reputable work produced, the evaluation sponsor selected the program agency's evaluation unit to conduct the evaluation research.

The Evaluation Environment. The evaluation environment, a constraint on evaluation strategies, is determined by contextual variables as well as the autonomy of the institutional arrangement (Proposition 6.6). The amount of threat the program manager perceives is a key feature of the evaluation milieu. While the evaluation environment was characterized as fairly cooperative, tensions did emerge as other governmental agents began to place demands on the research strategy in terms of its design.^{7/} Tension also increased slightly as the decision to continue or to terminate the program approached and more actors recommended termination of the program. Figure 7-2 summarizes the variables expected to contribute to the evaluation environment (Propositions 6.6a to 6.6f). The results of the case study are also noted in this figure.

Some of the conditions which were hypothesized to contribute to a cooperative evaluation environment were present. The program

^{7/} The program manager and evaluator were asked to characterize the evaluation environment (see Appendix A). They rated the environment as some cooperation to very cooperative.

VARIABLE	COOPERATION	TENSION	CASE STUDY
Evaluation Sponsor	Program Manager	Decision-Maker	State Liaison
Form of Demand	Required	Requested	Required
Type of Decision	Nonthreatening	Threatening	Shift from Nonthreatening to Threatening
Program Funding	Congruent	Noncongruent	Shift from noncongruent to congruent
Program Manager	Experimental	Trapped	Experimental
Autonomy	Internal	Academic	Governmental

Figure 7-2. The Evaluation Environment

manager was experimental and the form of the demand was a required evaluation. Other contextual variables tended to produce a cooperative environment as well: The evaluation sponsor was the liaison between the federal agency and the state program. In this case, the program manager did have some control over the evaluation contract which also lessened the potential threats to the program.

Two of the contextual variables shifted during the course of the evaluation research: The type of decision pending and program funding. As the decision implied by the evaluation shifted from a perfunctory review to possible program termination, tension increased. There was also some indication that as the congruency between the evaluation sponsor and program funding shifted to an incongruent mixture, tensions emerged as the program manager responded to state demands, not federal, concerning evaluation strategies.

Assessment of the Measures and Research Approach

The pilot study served to test measurement instruments as well as to judge the feasibility of the research approach.

The Measures and Survey Instruments. Most of the questions could be answered easily by the individuals interviewed. Perhaps the most serious problem emerged when users attempted to assign a specific weight to the amount of influence given the program evaluation product. Here, users were generally unable to give precise estimates such as 10 percent or 80 percent. However, users were able to answer the question when it was revised to ask

whether it contributed very little or a great deal when making a program recommendation or choice. Modification may also be necessary in a larger study with respect the rating technical quality and useability potential characteristics.

The technical quality index, a rating by the researcher, is a modification and extension of Bernstein and Freeman's quality index (1975). The items in the index were relatively easy to code from the evaluation product. However, an intermediate category for dichotmous indicators such as external validity and reliability could give more precision and allow for more variability among program evaluation products.

The useability potential index serves as a predictor of potential usefulness for a wide variety of governmental agents. The timing of the evaluation rindings is one of the more difficult items to score without the benefit of some interviews. The date of the report and knowledge of the budget cycle or legislative hearings is another way of coding the information..

For both researcher indices, the reliability among coders is an issue which must be addressed when implementing this research. Ideally, two or more individuals should rate an evaluation product. While the technical quality items are easier to specify, some of the useability potential items need clearer standards for categorizing products. Perhaps benchmark studies which serve as guides could help to ease some of these reliability problems.

The Research Strategy. The pilot study of the research design and operationalizations indicate that the propositions generated by the rational choice perspective are testable and the preliminary results are encouraging. No major difficulties were encountered in terms of identifying program evaluation products, identifying potential users, or securing cooperation for interviews. Respondents seemed genuinely interested in the research effort and willing to answer the questions. Identifying potential users and problems of recall may occur, however, with less visible type program evaluation efforts or if too much time has elapsed since the findings were released.

The major drawback of the proposed design is the amount of time needed to collect cases for a comparative analysis. Interviews with the program manager, evaluators and sponsor tend to take at least one hour. The average time to administer the user survey is approximately thirty minutes. The number of potential users will vary given the controversy of the program and the number of actors involved in a particular program decision. Thus, an extensive research effort, although needed to adequately test the propositions, is a time-consuming endeavor.

CONCLUSION

The practice and study of program evaluation activities have grown dramatically over the past decade. While the scholarly literature and empirical studies examining issues and topics is proliferating, it tends to be scattered, rather unsystematic, and often lacks a conceptual basis. In order for program evaluation to mature as a field of inquiry, it was argued, a more systematic and conceptual approach towards the evaluation enterprise needs to be taken. This dissertation is one effort in this direction. By making plausible assumptions based on a rational choice perspective, extending logical underpinnings, and constructing simplified models, it seems that some of the literature can be organized, past empirical findings can be integrated and explained, and new paths for empirical inquiry suggested.

Organizing The Literature. Defining program evaluation is essential in terms of conducting empirical studies as well as discussing various issues and topics involved with the evaluation enterprise. Yet, often the definitions offered seem incompatible or unclear. In an effort to analyze the various conceptualizations found, a typology of dimensions based on the methods used, focus on the program, and timing in the program cycle was constructed (Chapter Two). This typology provides some structure when turning to the evaluation literature and attempting to reconcile types of program evaluations often enumerated.

The typology of motives based on rational choice assumptions provides a framework for integrating some of the scattered and seemingly divergent motives cited for demanding and sponsoring a program evaluation (Chapter Three). While the literature often enumerates a number of possible motives, this typology groups motives into two broad categories: Evaluation efforts are motivated by dissatisfaction/puzzlement or in an attempt to generate program evidence. Besides developing typologies which may help to organize and clarify definitions of program evaluation and motives underlying program evaluation efforts, the dissertation also offers a framework for integrating and explaining past empirical findings concerning utilization.

Integrating and Explaining Past Empirical Findings. By extending rational choice assumptions to potential consumers of program evaluation products and constructing a model of a governmental agent's decision calculus, some of the divergent concepts of utilization may be integrated and past empirical findings explained (Chapter Five). This model and its discounting procedure accommodates both decisionistic and enlightenment perspectives of utilization. In addition, the program decision calculus model sheds some insight into empirical findings citing the lack of identifiable program changes made attributable to evaluation findings.

New Paths For Research. Perhaps the major contribution of this dissertation effort lies in directing research efforts to new areas. By proceeding systematically through the various aspects of the

evaluation enterprise, applying a rational choice perspective to each aspect, identifying important variables, and constructing a simplified model of the evaluation enterprise, a number of testable propositions were generated.

With respect to the demand aspect, an area often neglected in the literature, the propositions offered direct research efforts towards an analysis of program conditions leading to evaluative scrutiny and the costs and benefits involved in sponsoring an evaluation effort. Although the supply aspect of the evaluation enterprise has been an area of traditional concern for scholars, the propositions developed suggest research focussing on packages of information which may compete with a program evaluation product. Research in this area could enhance our understanding of utilization of program information.

Utilization of program evaluation products is an area which is receiving considerable empirical scrutiny. Yet, much of this research proceeds rather serendipitously, lacking a conceptual basis. The program decision calculus model and discounting procedure provide a framework for studying utilization and directs research towards examining the relationship between characteristics of a program evaluation product and the amount of influence the findings make in light of other information considered.

The model of the evaluation enterprise also presents a framework from which empirical studies can develop. While some of the propositions developed are admittedly speculative, they posit linkages among the various aspects of the evaluation

enterprise: The affect of institutional arrangements on technical quality and useability characteristics given environmental and resource constraints, the dynamics involved in the selection of an evaluation unit and the resource allocation process, and the relationship between contextual variables and the evaluation environment.

While the conceptual approach taken in the dissertation developed models and typologies as clarification devices and developed testable propositions, the utility of any conceptual approach lies in the ability to test the propositions derived. The pilot study confirms the testability of the propositions as well as provides some preliminary assessments.

APPENDICES

APPENDIX A

A PROPOSED RESEARCH DESIGN AND MEASURES

This appendix develops a research strategy for testing the models and propositions generated throughout the dissertation and suggests operational measures for the variables. Appendix B contains the survey instruments. A report of the pilot study can be found in Chapter Seven.

A RESEARCH STRATEGY

Preliminary Assumptions

Before proceeding to outline a research strategy, the assumptions made when designing this methodology for testing the models and propositions need clarification. The research strategy is designed for execution in a state government setting. However, with slight modification, the research design is equally applicable to federal evaluation efforts. Furthermore, evaluation research products constitute the primary interest for hypothesis testing. However, the general research strategy outlined in subsequent sections is equally applicable to program evaluation products produced by less formal methods.

Research Design

The following steps outline the design for implementing research:

(1) The research effort begins by compiling an inventory of program evaluation products and/or evaluation research products. The sample of cases drawn from this inventory for analysis depends, however, on research aims. For example, a researcher may be interested only in a particular policy area such as mental health. Thus, only evaluation research products which analyze mental health programs would be selected.

(2) The researcher must then secure copies of the evaluation research products selected for analysis. Each evaluation research

product must then be coded in terms of the producer, nature of the results as well as be rated in terms of technical quality and useability potential. After identifying the evaluator who conducted the evaluation research and the program analyzed, interviews with evaluators and program managers are scheduled.

(3) The next step involves interviewing evaluators. Evaluators provide data on evaluation process variables such as length of time devoted to the study, tensions arising during evaluation research, and autonomy from program managers. Evaluators are also a potential source of information for collecting demand variables such as evaluation research funding and sponsorship. In order to help identify potential consumers of evaluation research products, evaluators are also queried about the dissemination of the final study.

(4) Program managers provide data necessary to test all phases of the evaluation enterprise: evaluation research demand, production, and utilization. Program managers are one source of information concerning the politics surrounding the demand for evaluation as well as other demand variables such as evaluation sponsorship. It is also necessary to secure information about a program manager's attitudes towards evaluation research (production variables). A program manager is also a logical recipient of evaluation research information. Thus, a program manager must also be administered questions concerning the use of an evaluation research product.

(5) In addition to program managers, other potential consumers of evaluation research products must be identified. Evaluators and program managers may be helpful in identifying potential consumers of evaluation

research. However, biases may occur if strong advocates or strong opponents of evaluation research findings cite potential consumers reflecting their own biases. To mitigate against this bimodal bias, then, a supplementary strategy for identifying potential consumers is suggested. Governmental agents connected with a program and/or policy area should be interviewed and sampled. This list of potential consumers includes policy administrators connected with the program, legislators who serve on committees dealing program issues, staff of the budget office reviewing program activities, and the like. State directories and publications can be used to identify these individuals. Once potential consumers are identified, data concerning the utilization of evaluation research may be obtained through interviews.

In order to implement this research design, the variables of the evaluation enterprise models and propositions must be operationalized. The following section outlines a method for identifying evaluation research products and suggests measures for evaluation research demand, production, and utilization variables. The survey instruments and evaluation research product rating devices are based on these measures: Appendix A contains the survey instruments designed for evaluator, program manager, and potential consumer interviews.

Unit of Analysis

A program evaluation product/evaluation research product serves as the unit for analysis. The models and propositions generated in previous chapters encompass evaluation research demand, production, and utilization variables. These variables must be collected for each evaluation research project selected for analysis. Thus, providing an

operational definition for evaluation products is a central task.

Program evaluation, as defined in Chapter Two, is information assessing public programs produced for consumption in the action context. Thus, a program evaluation product refers to this package of information made available to individuals in the program planning-implementation-assessment cycle and other individuals in the action context. The term program evaluation product is a generic one encompassing process and impact studies; it also includes program information produced by informal methods as well as formal methods.

However, much of the concern in the evaluation literature as well as in this dissertation centers more specifically on evaluation research. Thus, a definition of evaluation research products must also be offered. Evaluation research products are a subset of evaluation products produced by scientific methods. These two general definitions allow comparing the weight of evaluation research products vis-a-vis other types of evaluation products in program decision-making.

Evaluation research products are defined by a reputational method. That is, information which governmental agents claim are evaluation research products identify the universe of cases for analysis. Identifying evaluation research products at the federal level of government using this reputational approach is relatively easy. The U.S. Comptroller General issues a directory entitled Federal Program Evaluations which lists federally funded program evaluations (U.S. Comptroller General, 1977). The current directory provides a listing of evaluations funded between 1975 and 1977. A subject and agency index allow restricting research to specific areas of interest.

While a sampling frame for federal program evaluation products is readily available, this is not generally the case for state evaluation research products. Although some of the evaluation research sponsored by the federal government may be conducted by state agencies and ultimately used by state government officials, relying on the federal directory alone would certainly omit a number of program evaluation products from the study. Thus, an inventory of program evaluation products produced for consumption within the state action context must be made.

An inventory of program evaluation products can be compiled by interviewing or sending a questionnaire to governmental agents likely to identify program evaluation products. These governmental agents include directors of departments, individuals housed in departmental research units, budget officials, the staff of legislative committees and legislative research units. The following question provides an operational definition for compiling an inventory of program evaluation products.

Operational Definition A.1 - Program Evaluation Product

Considering any information which assess program processes or impacts as program evaluations, what program evaluations has your department/unit sponsored? contracted? received?
--

In order to distinguish evaluation research products from other types of evaluation products, the following operationalization is suggested:

Operational Definition A.2 - Evaluation Research Product

An evaluation research product makes some claim or an attempt to gather program information according to scientific research principles.
--

Evaluation research products may be differentiated from other types of program evaluation products by two methods: A researcher may ask governmental agents to identify evaluation research products or they may be identified upon examining the program evaluation products cited in the inventory.

After completing an inventory of evaluation research products, a sample of cases for research may be drawn. By sampling cases for analysis, an attempt to minimize researcher bias in selection is made as well as generating research findings with some degree of generalizability. Yet, sampling does not completely eliminate bias from the analysis. There are some potential problems that should be noted with this reputational method. By relying on a reputational approach, the sampling frame itself may be biased. Two types of biases come to mind and need checking when executing research. It is possible that governmental agents may tend to overreport the "best" evaluation research efforts. Since this may be viewed as an academic research study, governmental agents may cite evaluation research studies which they feel best typify program evaluation ideals. If overreporting of sophisticated evaluation research occurs, then the variance of quality measures will be restricted and analysis will be impaired.

A second type of bias may also result if governmental agents tend to overreport evaluations which are extremely poor or extremely sophisticated. It is possible that if a governmental agent queried feels

hostile towards evaluation efforts in general, he may cite cases which exemplify poor quality or irrelevant information. On the other hand, a governmental agent may be a strong advocate of program evaluation efforts and may cite only the most elegant studies or generate controversial findings. If this situation occurs, reporting extreme examples of program evaluation products, then one would expect a bimodal distribution on technical quality and useability potential measures. Subsequent analyses would then be misleading.

Once a sample is drawn, data measuring evaluation research demand, production, and utilization variables must be gathered for each case. Thus, operationalizations for these variables must be made in order to test propositions and models suggested in earlier chapters. For purposes of clarity, the suggested measures are broken into three major groupings: demand variables, production variables, and utilization variables.

MEASURES DEVELOPED FOR TESTING DEMAND HYPOTHESES

A number of demand variables (DM) need measuring in order to test propositions hypothesizing conditions which lead to evaluation research (see Chapter Three).

7.1
(DM)

PROGRAM LONGEVITY

How long has the program been operating? _____yrs

0 = New program

1 = Established program

7.2
(DM)

COMPARATIVE PROGRAM EXPENDITURES

(a) What is the budget for this program? \$ _____

(b) What proportion of the total department's budget does this program comprise? \$ _____

(c) What proportion of the total state budget does this program's allocations make? \$ _____

7.3
(DM)

DEGREE OF PROGRAM CONTROVERSY

How would you characterize the debate surrounding the introduction of this program when it was considered by the legislature (Congress)?

1 = No controversy involved

2 = A little controversy involved

3 = Some controversy involved

4 = Some major controversy involved

5 = Extremely controversial

7.4
(DM)

SPECIFIC CONTROVERSY UNDERLYING REQUEST

(a) Was any particular event/controversy linked to the request for an evaluation? (0 = no; 1 = yes). Explain.

(b) Did any of the following stimulate the request?

No (0)

Yes (1)

Disatisfied program clients

Constituents demand
accountability

Programmatic crisis publicized

Other (Specify)

7.5
(DM)

BUDGETARY CONSTRAINTS

Was the request for evaluation research linked to any new budgetary crises? reorderings? (no = 0; 1 = yes). Explain.

7.6
(DM)

REQUIRED EVALUATION ENVIRONMENT

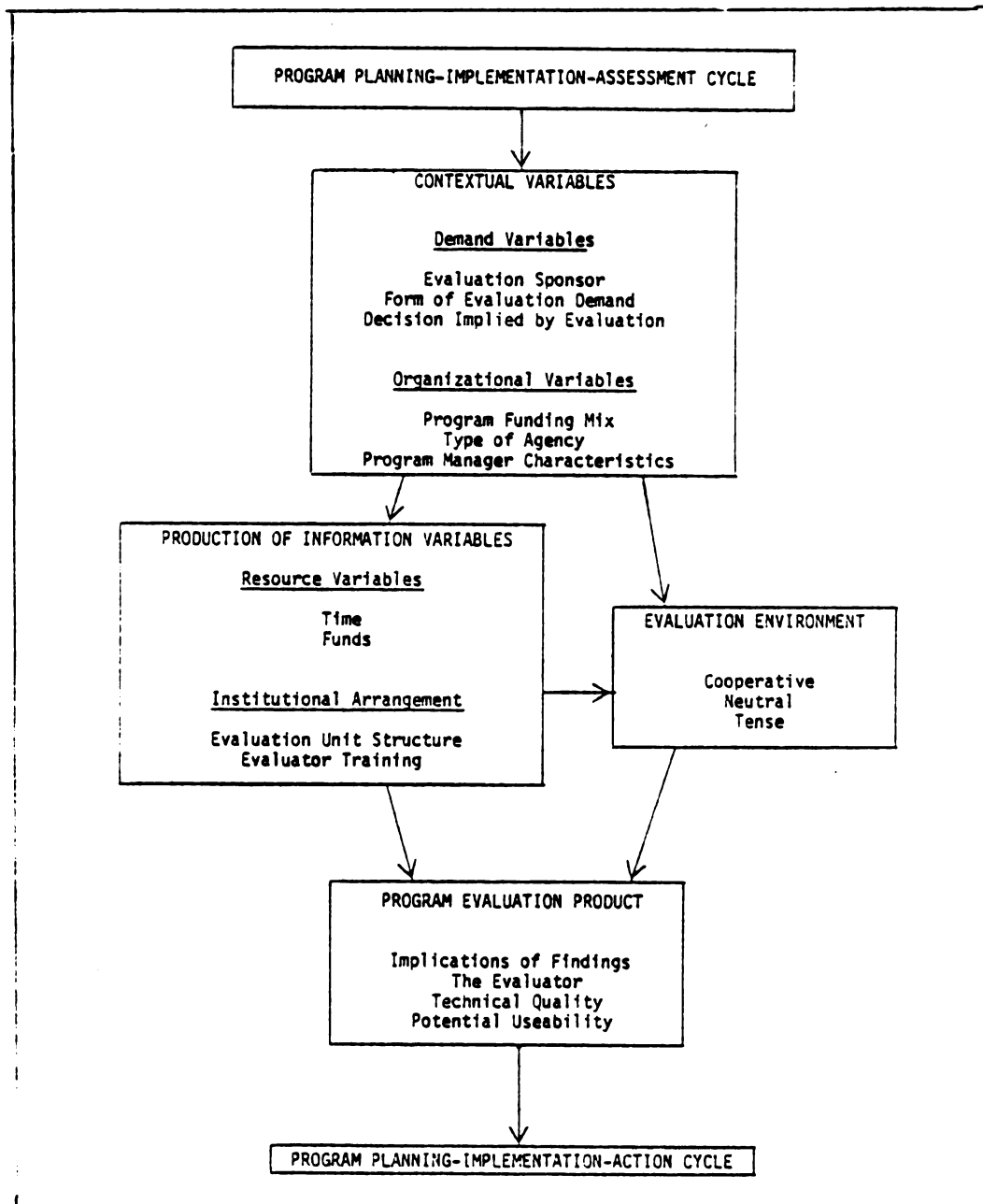
(a) When was the requirement for evaluation research written into legislation? _____yr

(b) Does this coincide with increased evaluation activities? (no = 0; yes = 1). List: Planning - programming - Budgeting; management by objectives, GAO Act.

MEASURES FOR TESTING THE EVALUATION ENTERPRISE HYPOTHESES

Chapter Six presented the proposed model of the evaluation enterprise and suggested a number of hypotheses. Measures for the variables identified in the model, then, are presented in this section (Model A.1).

Model A.1 - The Evaluation Enterprise



CONTEXTUAL

Demand Variables7.7
(PV)

DEMAND: SPONSOR OF EVALUATION RESEARCH

Who sponsored (required/requested) the study?

	State	Federal
Program Manager	1	4
Decision-maker	2	5
Policy Administrator	3	6
Combination	7	

7.8
(PV)

DEMAND: EVALUATION RESEARCH FUNDING

Who funded the evaluation research?

- 1 = State
- 2 = Federal
- 3 = State/Federal

7.9
(PV)

DEMAND: FORM OF THE DEMAND

Was the evaluation required or requested?

- 0 = Required
- 1 = Requested

7.10
(PV)

DEMAND: DECISION IMPLIED BY EVALUATION RESEARCH

What best describes the type of program decision implied by the requirement (request) for an evaluation?

- 1 = Perfunctory program review
- 2 = Modify current program operations
- 3 = Improve program management
- 4 = Choose among competing program strategies
- 5 = Decide whether to continue or to terminate program

CONTEXTUAL

Organizational Variables7.11 ORGANIZATIONAL: TYPE OF AGENCY HOUSING PROGRAM
(PV)

How would you characterize the agency which houses the program studied?

- 1 = Primarily a research agency
- 2 = Primarily an operating agency
- 3 = A mixture of both research and operating functions

7.12 ORGANIZATIONAL: PROGRAM FUNDING
(PV)

What proportion of program funds does the federal government provide? _____%

7.13 ORGANIZATIONAL: PROGRAM MANAGER CHARACTERISTICS
(PV)

(a) Background characteristics:

Level of Training 1 = B.A.
 2 = M.A.
 3 = Ph.D.

Field(s) _____

(b) What other experience have you had in this field? Explain.

(c) How many years have you been the manager of the program?

CONTEXTUAL

7.14 ORGANIZATIONAL: TRAPPED/EXPERIMENTAL PROGRAM MANAGER
(PV)

(a) What is the history of the program? What is the political situation for this program?

0 = Trapped (program future uncertain/
controversial)

1 = Experimental (program future secure/minor
adjustments)

(b) How receptive is the program manager to new program strategies?

1 = Very receptive

2 = Some interest

3 = Neutral

4 = Some resistance

5 = Openly hostile

(c) In general, how would you characterize your reaction when a new program strategy is suggested for your program?

1 I am always willing to experiment with new program strategies.

2 I am often interested in searching for new program strategies.

3 I am rather neutral about new program strategies suggested.

4 I am rather skeptical about new program strategies until there is sufficient information or experience with these approaches.

5 I feel that there are few new strategies which could improve the program.

CONTEXTUAL

7.15
(PV)

ORGANIZATIONAL: THREAT TO PROGRAM MANAGER

- (a) When the evaluation research was first started, what kind of findings did you expect? Explain.

0 = No challenge to the program
1 = Challenge to the program

- (b) Given these possible findings, how would you best characterize your feelings towards the evaluation research effort? (Use the following statements.)

1	.	.	.	5
Strongly				Strongly
Agree				Disagree

- The evaluation study would contribute little new information about the program.
- Some minor program changes may be indicated by the evaluation findings.
- The evaluation findings would directly challenge current program operations.
- The evaluation findings would be used as political ammunition for program supporters.
- The evaluation findings would be used as political ammunition for those who oppose the program.

PRODUCTION OF INFORMATION

Resource Variables7.16
(PV)

RESOURCE: AMOUNT OF FUNDS

How much did the evaluation research study cost?
\$ _____

7.17
(PV)

RESOURCE: TIME

- (a) How long did the study take to complete?
_____ months
- (b) How many manhours? _____

Institutional Arrangement Variables7.18
(PV)

INSTITUTIONAL ARRANGEMENT: EVALUATION UNIT STRUCTURE

- 1 = Internal, attached to program
2 = Internal, attached to program department
3 = External, attached to another government department
4 = External, academic
5 = External, entrepreneurial

7.19
(PV)

INSTITUTIONAL ARRANGEMENT: EVALUATOR CHARACTERISTICS

- (a) Regarding the study, how many persons are included on your professional/technical staff?

_____ full-time
_____ part-time

- (b) Background: Degree* Field Years Experience
- Project director
Principal investigator
People most essential to project

*1 = B.A.; 2 = M.A.; 3 = Ph.D.

PRODUCTION OF INFORMATION

7.20
(PV)

INSTITUTIONAL ARRANGEMENT: AUTONOMY

- (a) How would you best characterize the relationship between the agency funding the evaluation team? the program manager and the evaluation team?
- 1 = Confined to fiscal and related administrative decisions; conduct of research primarily determined by evaluation team
 - 2 = Formal reporting and review of major research decision with conduct determined by research team
 - 3 = Close supervision of research activities and major decisions in conduct by sponsor (program manager)
- (b) Under what circumstances can the findings be released?
- 1 = No restrictions
 - 2 = Submit copy before releasing
 - 3 = Sponsor/program manager must approve

Based on Bernstein & Freeman, 1975: 165

EVALUATION ENVIRONMENT

7.21
(PV)

ENVIRONMENT: TENSE/COOPERATIVE

- (a) How would you describe, in general, the working relationship between your evaluation team and the following individuals or groups?

Funding Agency	Departmental Director
Program Director	Governor's Office
Agency Director	Other Research Units

- 1 = Very cooperative
- 2 = Some cooperation
- 3 = Neutral
- 4 = Some tension
- 5 = Very tense

- (b) Did you encounter any particular political difficulties while conducting the research? Explain.

- (c) How would you describe, in general, the working relationship between your staff and the evaluation team?

- 1 = Very tense
- 2 = Some tension
- 3 = Neutral
- 4 = Some cooperation
- 5 = Very cooperative

- (d) How do you think the working relationship would have been with the following type of evaluation situations?

- 1 = Very tense
- 2 = Some tension
- 3 = Neutral
- 4 = Some cooperation
- 5 = Very cooperative

- An evaluation unit housed in the Department of Management and Budget.
- An evaluation unit housed within the Department.
- An evaluation team attached directly to the program.
- A private consulting firm.
- A university evaluation team.

EVALUATION PRODUCT

7.25
(PV)

PRODUCT: USEABILITY POTENTIAL INDEX

Recommendations

- 2 = Clearly stated
- 1 = Some implied
- 0 = No clear indication

Format

- 2 = Jargon free
- 1 = Some jargon
- 0 = Extensive jargon

Political Feasibility

- 2 = Studied manupuable variables
- 1 = Some manupuable variables studied
- 0 = Few manupuable variables studied

Findings

- 2 = Conclusive
- 1 = Some evidence
- 0 = Unable to draw conclusions

Budget Implications

- 2 = Less than current effort
- 1 = Same as current effort
- 0 = Greater than current effort

Timing

- 2 = Released in time for pending decision
- 1 = Late for decision
- 0 = Not concerned with releasing information

Based on discussion presented in Chapter Five

EVALUATION PRODUCT

7.26
(PV)

PRODUCT: TECHNICAL QUALITY

External Validity

- 1 = Applicable to other settings
- 0 = Restricted to setting of study

Research Design (Control of Internal Validity Threats)

- 3 = Experimental; Quasi-Experimental with both randomization and control
- 2 = Experimental/Quasi-experimental without randomization/control group
- 1 = Longitudinal/cross-sectional without control or comparison
- 0 = Descriptive/narrative, case

Sampling (Systematic Data Collection)

- 2 = Representative
- 1 = Possibly representative
- 0 = Haphazard

Measurement (Construct Validity)

- 1 = Adequate in face validity
- 0 = Inadequate in face validity

Theoretical Framework

- 1 = Reliance/reference to broader theoretical framework
- 0 = No reliance/reference to broader theoretical framework

Data Analysis

- 2 = Quantitative and qualitative
- 1 = Quantitative
- 0 = Qualitative

Methodological Appropriateness (Control of Statistical Conclusion Validity Threats)

- 2 = Few threats to statistical conclusion validity present
- 1 = Moderate threats to statistical conclusion validity present
- 0 = Severe threats to statistical conclusion validity present

Review of Findings

- 1 = Indication that findings preliminarily reviewed in wider context
- 0 = No indication that findings reviewed

Objectivity

- 2 = Stated assumptions and biases, examined alternative explanations/views
- 1 = Stated assumptions and biases clearly only
- 0 = No stated assumptions

Reliability

- 1 = Measures are reliable (stable)
- 0 = Measures are unreliable (unstable)

EVALUATION PRODUCT

7.22
(PV)

PRODUCT: TYPE OF PROGRAM EVALUATION

Type of evaluation product:

- 1 = Process study
- 2 = Impact study
- 3 = Combination/comprehensive

7.23
(PV)

PRODUCT: PRODUCER OF INFORMATION

Title of principal investigator:

- 0 = Evaluator
- 1 = Researcher

7.24
(PV)

PRODUCT: NATURE OF THE RESULTS INDEX

Support for current program approach:

- 0 = Overwhelmingly negative evidence
- 1 = Some negative evidence
- 2 = Unclear
- 3 = Some support for current approach
- 4 = Strongly support current approach

MEASURES FOR TESTING UTILIZATION HYPOTHESES

The final set of measures presented allow investigating the utilization of evaluation products. The program decision calculus model and its resultant propositions can be found in Chapter Five. The following set of questions are designed for identifying potential users of evaluation research products and for administering to potential users.

IDENTIFYING POTENTIAL USERS

7.27
(UV)

DISSEMINATION: IDENTIFYING CONSUMERS

- (a) Check the following that apply: Reports required
Major consumer
Maximum communication
- For the following:
- Federal Decision-Makers
 - State Decision-Makers
 - Staff-Federal Agency
 - Staff-State Agency
 - Program Manager/Staff
 - Other Researchers
 - University Affiliates
 - Community Groups
 - General Public
 - Mass Media
 - Other (Specify)
- (b) What other methods besides written reports have you employed?
- (c) Who do you think will utilize the results and how will they use them?
- (d) What other methods, besides written reports, have you employed for dissemination of the findings?

Based on Bernstein and Freeman, 1975: 166

The following utilization variables (UV) need measuring in order to test propositions concerning the use of evaluation research products by governmental agents. This section offers operationalizations for the variables related to the Program Decision Calculus model:

MODEL - PROGRAM DECISION CALCULUS

$p(\text{OUTCOME}) = f(\text{Information, Discount Factor})$

$\text{PROGRAM CHOICE} = f(p(\text{OUTCOME}), \text{self-interest})$

Where,

$\text{DISCOUNT FACTOR} = f(\text{Source of Information, Interpretive Bias, Reliability and Validity, Policy Relevance})$

Utilization Variables

7.27
(UV)

CONSUMPTION: BACKGROUND INFORMATION OF USER

- (a) Name:
- (b) Title:
- (c) Length of time in present position _____ yrs
- (d) Past positions
- (e) Academic field of training
- (f) Level of training:
 - 1 = B.A.
 - 2 = M.A.
 - 3 = Ph.D.

(a) Did you receive a Copy of _____?
When? How?

- (b) Did you read this report? read a summary of the report prepared by your staff? Explain.

- (c) Did the results of the study contribute directly to any program changes that you suggested or implemented? Explain.

0 = No
1 = Yes > Decisionistic

- (d) Did the results of the study affect your thinking (ideas or information) concerning the program?
How?

0 = No
1 = Yes > Enlightenment

- (e) Did you use the study (information) to affect the opinions of others concerning the program?

7.29
(UV)

(a) Did the study (information) help to reduce your uncertainty concerning the program?

- (b) Did the study (information) help to increase your uncertainty concerning the program?

7.30
(UV)

The _____ conducted this evaluation.

- (a) How would you characterize your past experience with this group?

- (b) How would you characterize their past work?

- (c) What is the professional reputation of these evaluators?

- (d) What is the professional reputation of their work?

Responses (a) to (d):

	1	2	3	4	5
Poor					
Excellent					

7.31
(UV)

CONSUMPTION: INTERPRETIVE BIAS

(a) To what extent do the findings agree with your sense of the situation?

1 . . . 5
| |
Strongly Strongly
Agree Disagree

(b) To what extent do you agree with the following statements?

1 . . . 5
| |
Strongly Strongly
Agree Disagree

- The findings are compatible with my ideas and values.
- The findings imply the need for major change in philosophy, organization, or services.
- The findings support a position I already held.
- The findings are consistent with a body of previous knowledge.
- The findings challenge existing assumptions and institutional arrangements.
- The findings raise new issues or offer new perspectives.
- The findings are unexpected or novel.

Based on C. Weiss and Bucuvalus, 1977: 231-232.

7.32
(UV)

CONSUMPTION: TECHNICAL QUALITY

(a) To what extent do you think the findings are valid and reliable?

1 . . . 5

Extremely	Extremely
Reliable & Valid	Unreliable & Invalid

(b) To what extent do you agree or disagree with the following questions?

1 . . . 5

Strongly	Strongly
Agree	Disagree

- The study adds to descriptive, causal, or theoretical knowledge.
- The study is generalizable to equivalent populations.
- The findings are internally consistent and unambiguous.
- The study is objective and unbiased.
- The research design appears sound.
- The measurement of the major variables is consistent with other measures in the field.
- The sampling techniques are systematic and unbiased.

Based on C. Weiss and Bucuvalus, 1977: 231-232

7.33
(UV)

CONSUMPTION: USEABILITY POTENTIAL

(a) To what extent is the general topic of the study relevant to issues your office deals with?

1 . . . 5

| |

Extremely Extremely
Relevant Irrelevant

(b) To what extent do you agree with the following statements?

1 . . . 5

| |

Strongly Strongly
Agree Disagree

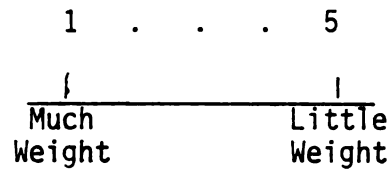
- The study deals with a high priority issue.
- The study adds to practical knowledge about the operations of policies and/or programs.
- The study adds to practical knowledge about the general operations of policies and/or programs.
- The study analyzes the effects of factors which decision-makers can do something about.
- The study has clear implications for a course of action.
- The study contains explicit recommendations.
- The implications of the findings are, in general, politically acceptable.
- The findings can be applied within existing agencies and programs.
- The study suggests strategies which are inexpensive to implement.
- The study is on time for a pending decision.
- The study was easy to read and to comprehend.

Based on C. Weiss and Bucuvalus, 1977: 231-232

7.34
(UV)

CONSUMPTION: WEIGHT IN DECISION CALCULUS

- (a) How much weight did you give the evaluation when considering the information?



- (b) What other information came to bear when using this study? Explain. What weight given?
- (c) Which factors were most important when considering this information and deciding whether or not to (1) suggest program changes or recommendation; (2) affect your ideas and thinking?

Rank order from most to least important factor.

- The source producing the information
- The nature of the findings
- The technical quality of the information
- The policy relevance of the study

Based on C. Weiss and Bucuvalus, 1977: 231-232

CONCLUSION

The research design and operationalizations developed in this Chapter facilitate empirically testing the models and the propositions formulated throughout the dissertation. In order to test the feasibility of this design and these measures, a pilot study was conducted. Chapter Seven, then, describes the pilot study, organizes the variables to test specific propositions, and analyzes the utility of the research strategy and measures. Survey instruments appear in Appendix B.

APPENDIX B

SURVEY INSTRUMENTS

B1. Evaluator Survey

EVALUATOR SURVEY

NAME: _____

DATE: _____

POSITION: _____

AGENCY: _____

* * * * *

BACKGROUND INFORMATION

1. Regarding the study, how many persons are included on your professional/technical staff? _____ (PV6.20a)

_____ Full-time

_____ Part-time

2. For each of the persons described below, what is their field of specialization, highest degree earned, and years experience? (PV6.20b)

a. Project Director--person responsible for the conduct of the evaluation research.

b. Principle Investigator--person who is responsible for the administration and organization of the evaluation research unit in which the study is housed. This person is typically responsible for fiscal matters, personnel practices, and other management functions. He is the person to whom the project director reports.

c. Essential Person--possibly a consultant who on a day to day basis participates most fully in the evaluation effort and whose work is most essential to the evaluation study.

a) Project
Director

b) Principle
Investigator

c) Essential
Person

____ SAME AS
(a)

____ SAME AS
(a)

____ SAME AS
(b)

FIELD OF SPECIALIZATION

DEGREE

YEARS OF EXPERIENCE

EVALUATION UNIT STRUCTURE

3. How is your evaluation unit organized? (PV-6.19)

____ attached to the program directly

____ attached to the program department

____ attached to another state government department

B1. Evaluator Survey, Continued

Page 2

- ☐ attached to academic organization
☐ attached to an entrepreneurial organization

RESOURCES & FUNDING

4. How much did the evaluation research study cost?(PV6.17) _____
5. How long did the study take to complete?(PV6.18)
- Beginning Date: _____ Termination Date: _____
- Approximate Manhours _____
- Any more reports planned? Explain.

DEMAND

6. Was the evaluation (a) required _____ or (b) requested _____? (DV-6.10)
- *IF (a) GO TO QUESTION 7. IF (b) GO TO QUESTION *.
7. When was the requirement for evaluation research written into legislation? (DV-6.7).
- *GO TO QUESTION 11*
8. Was any particular event/controversy linked to the request for an evaluation? Explain. (DV6.5)
9. Did any of the following stimulate the request? (DV-6.5b)
- ☐ Dissatisfied program clients
☐ Constituents demand accountability
☐ Programmatic crisis publicized
☐ Other (specify) _____
10. Was the request for evaluation research linked to any new budgetary crises or considerations or priorities? _____ Explain. (DV-6.7)

B1. Evaluator Survey, Continued

Page 3

11. What best describes the type of program decision implied by the requirement (request) for an evaluation? (DV-6.11)

☐ perfunctory program review
☐ improve program management
☐ choose among competing program strategies
☐ decide whether to continue or to terminate program
☐ modify current program operations
☐ other (specify)

SPONSOR & FUNDING

12. Who sponsored (required/requested) the study? (DV-6.8)

	State	Federal
Program Manager	<input type="checkbox"/>	<input type="checkbox"/>
Decision-Maker	<input type="checkbox"/>	<input type="checkbox"/>
Policy Administrator	<input type="checkbox"/>	<input type="checkbox"/>
Other:		

13. Who funded the evaluation research? (DV-6.9)

☐ State
☐ Federal
☐ Combination (specify %)

14. How would you best characterize the relationship between your evaluation team and: (a) the agency funding the evaluation;
(b) the program manager and program staff? (PV-621a)

☐ Confined to fiscal and related administrative decisions. Conduct of research primarily determined by the evaluation team.
☐ Formal reporting and review of major research decisions with the research conduct determined by the research team.
☐ Close supervision of research activities and major decisions in conducting research by sponsor (program manager).

B1. Evaluator Survey, Continued

Page 4

15. Under what circumstances can the findings be released? (PV-6.21b)

- ☐ no restrictions
- ☐ submit copy before releasing (specify)
- ☐ sponsor must approve
- ☐ program manager must approve
- ☐ other: specify

CONDUCTING EVALUATION RESEARCH

16. What is the history of this program? (PV-6.15a)

17. What was the political situation concerning this program when you began research? (PV-6.15b)

18. In your opinion, how receptive is the program manager (staff) to new program strategies/approaches? (PV-6.15c)

<input type="checkbox"/> very receptive	<input type="checkbox"/> some resistance
<input type="checkbox"/> some interest	<input type="checkbox"/> openly hostile
<input type="checkbox"/> neutral	

Evaluator Survey, Continued

Page 5

19. How would you describe, in general, the working relationship between your evaluation team and the following individuals or groups? (PV-6.22a)

1 = very cooperative	_____ Funding agency
2 = some cooperation	_____ Program Director
3 = neutral	_____ Agency Director
4 = some tension	_____ Departmental Director
5 = very tense	_____ Governor's Office
	_____ Other Research Units
	_____ Other (specify)

20. Did you encounter any particular political difficulties while conducting the research? Explain. (PV-6.22b)

21. Did you encounter any other difficulties when conducting the evaluation research?

Evaluator Survey, Continued

Page 6

DISSEMINATION

1. In your opinion, was the final report : ☐ released in time for a pending decision?
☐ late for a pending decision
☐ not concerned with a time for release

2. Check the following which apply:

Reports Required	Major Consumer	Maximum Communication	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Federal Decision-makers
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	State Decision-makers
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Staff--Federal Agency
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Staff--State Agency
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Program Manager/Staff
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Other Researchers
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	University Affiliates
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Community Groups
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	General Public
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Mass Media
			Other (specify)

3. What other methods besides written reports have you employed? (UV-6.28b)

4. Who do you think will utilize the results and how will they use them? (UV-6.28c)

B2. Program Manager Survey

PROGRAM MANAGER SURVEY

NAME: _____ DATE: _____

POSITION: _____

AGECNY: _____

* * * * *

BACKGROUND INFORMATION

1. What is your field of training? _____ Degree: _____ (DV6.14)
2. How many years did you manage the program? _____
3. What other experience have you had in this field? Explain. (DV-6.14)

PROGRAM CHARACTERISTICS

4. How long has the program been operating? (DV-6.3) _____ to _____
5. What is the budget of this program? (DV-6.4) \$ _____
6. What proportion of the total department's budget does this program comprise? (DV-6.4)
7. How would you characterize the debate surrounding the introduction of this program when it was considered by the legislature (Congress)? (DV-6.5)
 - ___ no controversy involved
 - ___ a little controversy involved
 - ___ some controversy involved
 - ___ some major controversies involved
 - ___ extremely controversial
8. How would you characterize the amount of controversy generated while implementing the program? (DV-6.5)
 - ___ no controversy involved
 - ___ a little controversy involved
 - ___ some controversy involved
 - ___ some major controversy involved
 - ___ extremely controversial
9. What proportion of the program funds does the federal government provide? (PV-6.13)
10. What best describes the type of program decision implied by the requirement (request) for the evaluation? (DV-6.11)
 - ___ perfunctory program review
 - ___ modify current program operations
 - ___ improve program management
 - ___ choose among competing program strategies
 - ___ decide whether to continue or to terminate the program

B2. Program Manager Survey, Continued

Page 2

CHARACTERISTICS

11. In general, how would you characterize your reaction when a new program strategy is suggested for your program?

☐ I am always willing to experiment with new program strategies.
☐ I am often interested in searching for new program strategies.
☐ I am rather neutral about new program strategies suggested.
☐ I am rather skeptical about new program strategies until there is sufficient information or experience with these approaches.
☐ I feel that there are few new strategies which could improve the program.

12. When the evaluation research was first started, what kind of findings did you expect? Explain. (DV-6.16)

13. Given these possible findings, how would you best characterize your feelings towards the evaluation research effort?

1 . . . 5
 Strongly Strongly
 Disagree Agree

☐ The evaluation study would contribute little new information about the program
☐ Some minor program changes may be indicated by the evaluation findings.
☐ The evaluation findings would directly challenge current program operations.
☐ The evaluation findings would be used as political ammunition for program supporters.
☐ The evaluation findings would be used as political ammunition for those who oppose the program.

14. How would you describe, in general, the working relationship between your staff and the evaluation team?

☐ very tense
☐ some tension
☐ neutral
☐ some cooperation
☐ very cooperative

B2. Program Manager Survey, Continued

Page 3

15. How do you think the working relationship would have been with the following type of evaluation situations?

very tense, some tension, neutral, some cooperation, very cooperative

- ☐ an evaluation unit housed in the Department of Management and Budget
- ☐ an evaluation unit housed within the Department
- ☐ an evaluation team attached directly to the program
- ☐ a private consulting firm
- ☐ an university evaluation team

Describe the program in general:

B3. User Survey

USER SURVEY

Page 1

NAME: _____

DATE: _____

POSITION: _____

AGENCY: _____

* * * * *

BACKGROUND INFORMATION

1. What is your field of training? _____ Degree: _____ (CV6.29)

2. How long have you served in your present position? _____ (UV-6.29)

3. What other experience have you had in the government? Explain.

TYPE OF USE

5. Did you receive a copy of _____? When? _____ How? _____

6. Did you read this report? _____

7. Did you hear about this study from someone else? a memo from your staff? Explain.

8. Did the results of the study contribute directly to any program changes that you suggested? _____ implemented? _____ Explain. (UV-6.30)

a. Did the study help to reduce your uncertainty concerning the program?

b. Did the study help to increase your uncertainty concerning the program?

B3. User Survey, Continued

Page 2

9. How much weight did you give the evaluation when considering this issue? 1 = little weight to 5 = much weight.
10. What other information came to bear when using this study? explain.
 - a. What weight was this information given? 1 = little weight to 5 = much weight
11. All in all, how much weight did the evaluation and other information receive? (example--50/50; 10/90, etc).
12. Which factors were most important when considering this information and deciding whether or not to suggest program changes?

1 = most important; 4 = least important

 - ___ the source which produced the information
 - ___ the nature of the findings (conclusions/inferences) included in the study
 - ___ the technical quality of the information
 - ___ the policy relevance of the study for the decision at hand.
13. Did the results of the study affect your thinking (ideas or information) concerning the program? ___ Explain. (UV-6.30)
 - a. Did you use the study to affect the opinions of others concerning the program? explain.
 - b. How much did this evaluation affect your ideas/thinking?

1 = very little to 5 = a great deal _____

Why this amount? explain. (previous knowledge, etc)
 - c. Did the study help to increase or to reduce your uncertainty concerning this program strategy? Explain.

B3. User Survey, Continued

Page 3

14. Which factors were most important in affecting your ideas or your thinking?

1 = most important; 4 = least important

- ☐ the source which produced the information
- ☐ the nature of the findings (conclusions/inferences) included in the study
- ☐ the technical quality of the information
- ☐ the policy relevance of the study for the issues and programs I deal with.

15. The _____ conducted this evaluation.

Responses: 1 = poor to 5 = outstanding

- ☐ How would you characterize your past experience with this group?
- ☐ How would you characterize their work based on your experience?
- ☐ How would you characterize the professional reputation of these evaluators in general?
- ☐ What, in general, is the professional reputation of their work?

16. To what extent do the findings agree with your sense of the situation?

Responses: 1 = strongly disagree to 5 = strongly agree

Explain. (GO TO A)

17. To what extent do you think that the findings are reliable and valid?

Responses: 1 = extremely unreliable and invalid to 5 = extremely reliable and valid

Explain. (GO TO B)

18. To what extent is the general topic of the study relevant to issues your office deals with?

Responses: 1 = extremely irrelevant to 5 = extremely relevant.

Explain. (GO TO C)

B3. User Survey, Continued

PART A.

To what extent do you agree with the following statements?

1 - - - 5
Strongly Strongly
Disagree Agree

- ☐ The findings are compatible with my ideas and values.
- ☐ The findings imply the need for major changes in the philosophy, organization, or services.
- ☐ The findings support a position I already hold.
- ☐ The findings are consistent with a body of previous knowledge.
- ☐ The findings challenge existing assumptions and institutional arrangements.
- ☐ The findings raise new issues or offer new perspectives.
- ☐ the findings are unexpected or novel.

Part B.

- ☐ The study adds to descriptive, causal, or theoretical knowledge.
- ☐ The study is generalizable to equivalent populations.
- ☐ The findings are internally consistent and not ambiguous.
- ☐ The study is objective and unbiased.
- ☐ The research design appears to be sound.
- ☐ The measurement of the major variables is consistent with other measures in the field.
- ☐ the sampling techniques are systematic and unbiased.

Part C.

- ☐ The study deals with a high priority issue.
- ☐ The study adds to practical knowledge about the operations or policies and/or programs.
- ☐ The study analyzes the effects of factors which decision-makers can do something about.
- ☐ The study has clear implications for a course of action.
- ☐ The study contains explicit recommendations.
- ☐ The implications of the findings are, in general, politically acceptable.
- ☐ The findings can be applied within existing agencies and programs.
- ☐ The study suggests strategies which are inexpensive to implement.
- ☐ The study is on time for a pending decision.
- ☐ The study was easy to read and to comprehend.

APPENDIX C

TECHNICAL QUALITY: FIVE CRITERIA

This Appendix discusses the five criteria of technical quality presented in Chapter Five in terms of the reliability and validity issues raised in a program evaluation product:

- (1) A Theoretical framework, (2) measurement reliability and validity, (3) research design and data collection techniques, (4) appropriate data analysis techniques and methods, and (5) objectivity and peer review.

Use of A Theoretical Framework

A theory is a set of laws, generalizations, and hypotheses which are related systematically and logically to one another. One of the major goals of science, it seems, is to build a body of knowledge which consists of theories that enable us to describe, explain, predict, and integrate information. Thus, theories serve three major functions: (1) They explain and predict phenomenon, (2) they raise new problems for inquiry, and (3) they help organize knowledge in a systematic and an efficient manner. ^{1/} As theoretical frameworks are central to basic scientific inquiry,

^{1/} These three functions are based on discussions found in McGaw & Watson (1976) and Legee and Francis (1974).

the use of a theoretical framework during evaluation research also helps to organize inquiry and to interpret findings. Many argue that public programs are based on theories either implicitly or explicitly. That is, a program strategy is implemented to set in motion a causal process which leads to the desired program effect (Suchman, 1972).

However, without basing an evaluation effort on a larger theoretical perspective, internal and external validity threats may arise. Threats to internal validity arise when factors other than the program may contribute to the measured changes in program clients or processes (Campbell & Stanley, 1963). However, if reference to a theory is made when designing evaluation research, the theory tends to specify other factors besides the program which may affect the program clients and outcomes. By identifying confounding factors, then, an evaluator can attempt to minimize their effects in the research design or at least attempt to measure their potential influence on program strategies and outcomes.

Threats to external validity arise when the evaluation findings concerning one program cannot be applied to other program settings, clients, or situations (Campbell & Stanley, 1963). Relying on a theoretical framework when designing evaluation research helps minimize some of these threats to external validity. That is, a larger theoretical framework may provide guidance concerning other settings and situations to which the evaluation findings may be applicable. Thus, reference to a theoretical

framework when designing a program evaluation effort and interpreting findings can potentially minimize some threats to external validity as well as to internal validity.

Measurement Validity and Reliability

The measurement process links abstract concepts to indicators that may be observed and assigned different values or properties (McGaw & Watson, 1976; Leege & Francis, 1974). Thus, during a program evaluation effort, the measuring process consists of assigning values to the observations made. However, the values assigned to an observation, test item, and the like reflect both the "true" score as well as measurement error (i.e., systematic and random error). Thus, two major issues emerge during this measurement process which effect subsequently the inferences which may be drawn from program evaluation efforts: construct validity and reliability.

Construct validity refers to accurately tapping the theoretical or abstract concept with the measure that one develops (McGaw & Watson, 1976; Leege & Francis, 1974). If a measuring instrument or device does not reflect the properties or dimensions accurately, systematic bias is introduced. Consequences of systematic error are severe during the analysis stage of research: Systematic error results in biased estimators and program evaluation findings, then, are often inaccurate and misleading.

Threats to construct validity are potentially most serious when measuring multifaceted, abstract concepts such as intelligence, attitudes, recidivism, and the like. When an evaluation effort

focuses on programs with more concrete concepts, such as the number of program recipients served, attaining construct validity may not pose as serious a threat. Devising and developing measures from a larger theoretical framework helps to minimize potential threats to construct validity during the measurement process.

Reliability is also an issue which emerges during the measurement stages of an valuation effort. Random measurement error results if the measures developed are not consistent from observation to observation. Unreliable measures also affect the analysis stages of program evaluation and the inferences which may be drawn from the findings. If the measuring process is unreliable, the explanatory power of the findings tends to be reduced. Thus, an evaluator may conclude erroneously that a program has little impact when in fact the low amount of program variation explained is merely a function of an unreliable measuring process.

Research Design and Data Collection

Research design is the plan that selects the people or processes for study, identifies the time frame for the research, and outlines the procedures for collecting the data. Research designs implemented in evaluation research vary in terms of their ability to control intrusions while measuring the program aspects of interest (internal validity issues) and in terms of their ability to allow application of the evaluation findings to other program settings and situations (external validity issues). There are four key features of a research design which attempt to control threats to internal and external validity: (1) Randomization--

a random sample of cases and the random assignment of cases to groups for study; (2) tests--pre-tests and post-tests; (3) comparison--control or comparison groups; and (4) Manipulation--the ability to manipulate the treatment (program) variables.

Randomization is a key aspect of design that helps to minimize measurement bias while random selection of cases for study facilitates accurate inferences to the population studies. Hence, the external validity (generalizability) of the evaluation findings is enhanced when randomization and random selection procedures are employed. Moreover, random assignment of participants to groups attempts to control internal validity threats such as selection bias.

Pre- and post-tests allow measuring changes which have occurred in program participants or changes in program services. While pre- and post-tests control some threats to internal validity, a pre-test often produces a sensitizing effect which may, in turn, confound the measurement process.

A control group to which cases are randomly assigned or a comparison group to which cases are not randomly assigned may control some threats to internal validity by gauging the amount of change which would occur without any program operating. In a comparison group, random assignment to the group does not occur--participants are either matched on characteristics of the group of interest or volunteer. A comparison group may reduce some confounding factors (internal validity threats). However, the generalizability of the evaluation findings may be somewhat restricted by the nature of the self-selection or the inability to match on relevant characteristics (external validity threats).

The ability to manipulate program treatment levels also helps to minimize some threats to internal validity. However, an experimental setting often limits the generalizability of the evaluation findings to a more natural setting. Hence, some threats to external validity may arise when an evaluator manipulates program variables.

Figure C-1 summarizes this brief discussion of research designs and their abilities to potentially control threats to internal and external validity threats. Figure C-1, displayed on the following page, groups a number of common evaluation designs in terms of experimental, quasi-experimental, and pre-experimental designs and identifies which of the four key features are generally present. The ability of the designs to control validity threats is also noted. ^{2/}

Selecting A Research Design. Cost, time, and feasibility constraints often determine the selection of a design for evaluation efforts. True experimental designs are often costly in terms of time and money (Rossi et al, 1979). Even if cost and time constraints are not present, feasibility considerations often plague evaluation efforts. For example, random assignment of participants to programs, a key feature of the experimental design, is generally prohibited either by law or by ethical considerations. Thus, quasi-experimental designs, which lack the random assignment component and manipulation of program variables, are often implemented

^{2/} See Campbell and Stanley (1963) and Cook and Campbell (1979) for a more thorough discussion of designs and validity threats.

KEY FEATURES.

RESEARCH DESIGN	EXTERNAL VALIDITY PROBLEMS	DECREASING CONTROL OF INTERNAL VALIDITY THREATS	(1) RANDOMIZATION		(2) TESTING		(3) COMPARISON		(4) CONTROL
			RANDOM SAMPLE	RANDOM ASSIGNMENT	PRE TEST	POST TEST	CONTROL GROUP	COMPARISON GROUP	MANIPULATE VARIABLE (X)
I. <u>TRUE EXPERIMENTAL</u> Soloman 4-Group Pre/Post Control Group Post-Test Only	*		yes yes yes	yes yes yes	yes yes no	yes yes yes	yes yes yes	- - -	yes yes yes
II. <u>QUASI-EXPERIMENTAL</u> Time Series Non-Equivalent Control	* *		{?} {?}	no no	2+ yes	2+ yes	no -	(?) yes	no no
III. <u>PRE-EXPERIMENTAL</u> Static Comparison Group One-Group Pre/Post Test One Shot Case Study	* * *		{?} {?} {?}	no no no	no yes no	yes yes no	- - -	yes no no	no no no

BASED ON CAMPBELL & STANTLEY (1973)

Figure C-1. Research Designs: Control of Threats to Internal and External Validity

instead. The pre-experimental designs, which in general do not adequately control threats to internal validity, are often implemented when time, money, or feasibility constraints are high.

Research design is a critical concern for the technical quality of a program evaluation product. It is, in effect, the blueprint for data collection and measurement. If threats to internal and external validity are not controlled adequately and the data not collected consistently (i.e. reliably), a governmental agent cannot make sound inferences and/or apply the evaluation findings to other situations and program settings.

Appropriate Data Analysis Techniques and Methods

The type of data analysis performed and the appropriateness of the methodology used are also criteria for judging the technical quality of a program evaluation product. After implementing a research design and measuring the phenomena of interest, the data must be analyzed and the program processes and effects analyzed.

Some studies tend to be strictly descriptive, relying on qualitative analyses. Other studies, however, may employ statistical analyses in order to make inferences and to test program hypotheses. There seems to be a growing concern that evaluation products include both qualitative and quantitative analyses of programs.

In terms of quantitative analyses, there is a wide array of techniques available. These statistical techniques vary in levels of sophistication--ranging from simple descriptive statistical

procedures, such as frequency distributions and percentages, to multivariate statistical procedures such as multiple regression analysis. While the more sophisticated multivariate techniques are often more robust and powerful tests, they may not be methodologically appropriate for the problem or the data at hand. Conversely, an evaluator may apply less powerful descriptive data analytic techniques to problems that require more sophisticated analyses.

The choice of the appropriate data analytical technique for the problem at hand is part of sound scientific procedures. The validity issue raised concerning data analysis and methodological appropriateness is statistical conclusion validity--threats during the data analysis stage which lead to false conclusions about covariation or statistically significant differences found. This type of threat to validity may be particularly important to program evaluation efforts since many final products fail to find statistically significant results. The reliability of the measuring procedure also affects statistical conclusion validity. If the measures are unreliable, the error terms become inflated and the chance of finding a relatively large correlation that is statistically significant tends to decrease (Cook & Campbell, 1979).

The data analysis stage, then, is also critical for technical quality considerations. If threats to statistical conclusion validity and reliability are not adequately controlled, accurate inferences may not be drawn about a program.

Objectivity and Peer Review

Bias during the evaluation process refers to a "pre-disposition or inclination on the part of the researcher to favor one definition, interpretation or inference over another" (Cherryholmes, 1978:2). This type of bias, or lack of objectivity, differs from the bias resulting from not controlling adequately the threats to internal and external validity. While it is never possible to eliminate completely the personal bias of the evaluator when making methodological choices concerning the measures used to tap concepts, design decisions, theoretical framework (paradigm) decisions, controlling the bias of an evaluator still remains a concern. Bias tends to increase when an evaluator responds unquestionably to demands by the program evaluation sponsors.

Ideally, the scientific process relies on peer review and intersubjective testability in order to minimize bias. Review of findings by the larger scientific community allows challenges of assumptions made as well as the falsification or modification of truth claims made. This open process, coupled with explicit statements concerning research assumptions, contributes to the relative "objectivity" of the information produced by scientific methods vis-a-vis less formal methods.

In terms of program evaluation, explicit statements of the value assumptions made, the desired goals one is measuring, and the like, lend an aura of objectivity to the evaluation effort as viewed by a governmental agent. Often, bias may be minimized

somewhat if the research is conducted in an autonomous fashion from the program manager or the evaluation sponsor control. Review and critique by the scientific or academic community also may act as a control on bias. Thus, the objectivity of the research, attempt to control bias, affects the technical quality of a program evaluation product and its subsequent weight given by a governmental agent.

All five criteria are aspects of technical quality and contribute to the reliability and validity of the information included in a program evaluation product. A governmental agent may discount the information when only one criterion of technical quality seems inadequate, such as the research design, or when a number of the criteria for sound and accurate inferences seem to be lacking.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Alkin, Marvin C. "Evaluation Theory Development." Evaluation
1969 Comment, Vol. 2, No. 1 (October), 2-7.
- Alkin, Marvin C., Richard Daillak, and Peter White. Using Evalua-
1979 tions: Does Evaluation Make A Difference? Beverly
Hills, Calif.: Sage Publications.
- Aronson, H. and C. Sherwood. "Research versus Practioner: Problems
1967 Social Action Research." Social Work, Vol. 12, No. 4,
89-96.
- Babbie, Earl R. Survey Research Methods. Belmont, Calif.: Wadsworth
1973 Publishing Company.
- Ball, Samuel and Scarvia Anderson. The Profession and Practice of
1978 Program Evaluation. San Francisco: Jossey-Bass, Inc.
- Banner, David D., I. Samuel Doctors, and Andrew C. Gordon. The
1975 Politics of Social Program Evaluation. Cambridge, Mass.:
Ballinger Publishing Company.
- Bartlett, Randall. Economic Foundations of Political Power. New York:
1973 Free Press.
- Bauer, Raymond A. Social Indicators. Cambridge, Mass.: The M.I.T
1966 Press.
- Bennett, Carl S. and Arthur A. Lumsdaine. Evaluation and Experiment.
1975a New York: Academic Press.
- Bennett, Carl A. and Arthur A. Lumsdaine. "Social Program Evaluation."
1975b in Bennett and Lumsdaine, (eds.). Evaluation and
Experiment. New York: Academic Press, pp. 1-38.
- Bernstein, Ilene N. Validity Issues in Evaluation Research. Beverly
1976 Hills, Calif.: Sage Publications.
- Bernstein, Ilene N. and Howard E. Freeman. Academic and Entrepre-
1975 neurial Research: The Consequences of Diversity in
Evaluation Studies. New York: Russell Sage Foundation.
- Bish, Robert L. The Public Economy of Metropolitan Areas. Chicago:
1971 Markham Publishing Company.

- Buchanan, James M. and Gordon Tullock. The Calculus of Consent: Logical Foundations of Constitutional Democracy. Ann Arbor: University of Michigan Press.
- 1962
- Brown, Robert D., Larry A. Braskamp, and Dianna L. Newman.
1978 "Evaluator Credibility as A Function of Report Style: Do Jargon and Data Make A Difference?" Evaluation Quarterly, Vol. 1, No 2 (May), 331-341.
- Caine, Glen G. and Robinson G. Hollister. "The Methodology of Evaluating Social Action Programs." in Peter H. Rossi and Walter Williams, (eds.). Evaluating Social Programs. New York: Seminar Press, pp. 107-137.
- 1972
- Campbell, Donald T. "Reforms and Experiments." American Psychologist, Vol. 24, No. 4 (April), 409-429.
- 1969
- Campbell, Donald T. and Julian C. Stanley. Experimental and Quasi-Experimental Design. Chicago: Rand McNally.
- 1966
- Caplan, Nathan, Andrea Morrison, and Russell J. Stambaugh. The Use of Social Science Knowledge in Policy Decisions at The National Level: A Report to Respondents. Ann Arbor, Mich.: Institute for Social Research, Center for Research on Utilization of Scientific Knowledge.
- 1975
- Caro, Francis G. "Evaluation Research: An Overview." in Francis Caro, (ed.). Readings in Evaluation Research. New York: Russell Sage Foundation, pp. 1-37.
- 1971a
- Caro, Francis G. (ed.). Readings in Evaluation Research. New York: Russell Sage Foundation, (First Edition).
- 1971b
- Caro, Francis G. (ed.). Readings in Evaluation Research. New York: Russell Sage Foundation, (Second Edition).
- 1977
- Carter, Reginald. "Clients Resistance to Negative Findings and Latent Conservative Functions of Evaluation Studies." The American Sociologist, Vol. 6 (May), 118-124.
- 1971
- Cherryholmes, Cleo. "Evaluation Research as A Linguistic Activity: An Exploration of Bias." (Mimeo). Paper Delivered at Political Science and The Study of Public Policy Conference, Michigan State University, (May 15-17).
- 1978
- Cohen, David H. "Politics and Research: Evaluation of Social Action Programs in Education." in Carol Weiss, (ed.). Evaluating Action Programs. Boston: Allyn and Bacon, Inc., pp. 137-166.
- 1972

- Cohen, David H. and Michael Garrett. "Reforming Educational Policy
1975 with Applied Social Research." Harvard Educational
Review, Vol. 45 (February), 17-44.
- Coleman, James S. Policy Research in The Social Sciences.
1972 Morristown, New Jersey: General Learning Press.
- Cook, Thomas D. (ed.). Evaluation Studies: Review Annual--Volume 3.
1978 Beverly Hills, Calif.: Sage Publications.
- Cook, Thomas D. and Donald T. Campbell. Quasi-Experimentation.
1979 Chicago: Rand McNally.
- Cook, Thomas D. and Charles L. Gruder. "Metaevaluation Research."
1978 Evaluation Quarterly, Vol. 2, No. 1 (February), 5-51.
- Cook, Thomas D. and Charles Reichardt. Qualitative and Quantitative
1979 Methods in Evaluation Research. Beverly Hills, Calif.:
Sage Publications.
- Davis, Howard R. and Susan E. Salasin. "The Utilization of Evalua-
1975 tion." in Elmer Struening and Marcia Guttentag. (eds.).
Handbook of Evaluation Research, Volume I. Beverly Hills,
Calif.: Sage Publications, pp. 621-666.
- deNeufville, Judith Innes. Social Indicators and Public Policy.
1975 New York: American Elsevier Publishing Company.
- deNeufville, Judith Innes. "Validating Policy Indicators." Policy
1979 Sciences, Vol. 10 (Winter), 171-187.
- Deming, W. Edwards. "The Logic of Evaluation." in Elmer Struening
1975 and Marcia Guttentag. (eds.). Handbook of Evaluation
Research, Volume I. Beverly Hills, Calif.: Sage
Publications, pp. 53-70.
- Downs, Anthony. "Some Thoughts on Giving People Economic Advice."
1965 American Behavioral Scientist, Vol. 9, No. 1
(September), 30-32.
- Fairweather, George W. and Louis Tornatzky. Experimental Methods for
1977 Social Policy Research. New York: Pergamon Press.
- Fiorina, Morris P. "Formal Models in Political Science." American
1975 Journal of Political Science, Vol. 19 (February), 133-159.
- Floden, Robert E. and Stephen S. Weiner. "Rationality to Ritual: The
1978 Multiple Roles of Evaluation in Governmental Processes."
Policy Studies, Vol. 9, 9-18.

- Franklin, Jack L. and Jean H. Thrasher. An Introduction to Program Evaluation. New York: John Wiley and Sons.
1976
- Frankenlious, Jeanne Pierre. "How to Get A Good Consultant." Harvard Business Review, Vol. 55, No. 6 (November/December), 133-139.
1977
- Freeman, Howard E. "Evaluation Research and Public Policies." in Gene M. Lyons. (ed.). Social Research and Public Policies. Hanover, New Hampshire: The Public Affairs Center, Dartmouth College, pp. 141-174.
1975
- Gabriel, Richard A. Program Evaluation: A Social Science Approach. New York: MSS Information Corporation.
1975
- Glass, Gene V. (ed.). Evaluation Studies: Review Annual--Volume 1. Beverly Hills, Calif.: Sage Publications.
1976
- Guba, Egon G. "The Failure of Educational Evaluation." in Carol H. Weiss. (ed.). Evaluating Action Programs. Boston: Allyn and Bacon, pp. 250-266.
1972
- Guttentag, Marcia. "Subjectivity and Its Use in Evaluation Research." Evaluation, Vol. 1, No. 2, 60-65.
1973
- Guttentag, Marcia. (ed.). Evaluation Studies: Review Annual--Volume 2. Beverly Hills, Calif.: Sage Publications.
1977
- Guttentag, Marcia and Elmer Struening. (eds). Handbook of Evaluation Research: Volume II. Beverly Hills, Calif.: Sage Publications.
1975
- Hatry, Harry P., R.E. Winnie, and Donald M. Fisk. Practical Program Evaluation for State and Local Officials. Washington, D.C.: The Urban Institute.
1973
- Hatry, Harry, Louis Blair, Donald Fisk, and Wayne Kimmel. Program Analysis for State and Local Governments. Washington, D.C.: The Urban Institute.
1976
- House, Ernest R. "The Politics of Evaluation in Higher Education." Journal of Higher Education, Vol. 45, No. 8 (November), 618-624.
1974
- Larsen, Judith K. and Daryl G. Nichols. "If Nobody Knows You've Done It, Have You . . .?" Evaluation, Vol. 1, No. 1 (Fall), 39-45.
1972
- Leege, David C. and Wayne L. Francis. Political Research: Design, Measurement, Analysis. New York: Basic Books.
1974

- Longood, Robert and Arnold Simmel. "Organizational Resistance to
1972 Innovation Suggested by Research." in Carol Weiss.
(ed.). Evaluating Action Programs. Boston: Allyn and
Bacon, pp. 311-317.
- Lyons, Gene M. (ed.). Social Research and Public Policies. Hanover,
1975 New Hampshire: The Public Affairs Center, Dartmouth
College.
- Mann, Floyd and Rensis Likert. "The Need for Research on the
1952 Communication of Research Results." Human Organization,
Vol. 11, No. 4 (Winter), 15-19.
- Mann, John. "Evaluating Educational Programs." Urban Review, Vol. 3,
1969 No. 4, 12-13.
- Mann, John. "The Outcome of Evaluation Research." Changing Human
1965 Behavior. New York: Charles Scribner's Sons, pp. 191-214.
- Mann, John. "Technical and Social Difficulties in The Conduct of
1965 Evaluative Research." Changing Human Behavior. New York:
Charles Scribner's Sons, pp. 177-189.
- McGraw, Dickenson and George Watson. Political and Social Inquiry.
1976 New York: John Wiley and Sons, Inc.
- Meltsner, Arnold J. Policy Analysts in The Bureaucracy. Berkely,
1976 Calif.: University of California Press.
- Merton, Robert. Qualitative and Quantitative Methods. New York:
1979 Free Press.
- Minnesota Systems Research, Inc. Methodological Adequacy of Federal
1975 R & D Projects. Washington, D.C.: National Technical
Information Service, U.S. Department of Commerce.
- Morris, Lynn and Carol Fitz-Gibbon. How to Deal with Goals and
1978 Objectives. Beverly Hills, Calif.: Sage Publications.
- Mushkin, Selma. "Evaluations: Use with Caution." Evaluation, Vol. 1,
1973 No. 2, 31-35.
- Nachmias, David. Public Policy Evaluation. New York: St. Martins
1979 Press.
- National Research Council. Study Project on Social Research and
1978 Development. Knowledge and Policy: The Uncertain
Connection. Laurence E. Lynn. (ed.). Washington, D.C.:
National Academy of Sciences.

- National Research Council. Study Project on Social Research and
1978 Development. The Federal Investment in Knowledge of
Social Programs. Washington, C.D.: National Academy of
Sciences.
- Office of Management and Budget. "Resources for Program Evaluation."
1977 in U.S. Senate. Committee on Human Resources. Costs,
Management, and Utilization of Human Resources Program
Evaluation, 1977. Washington, D.C. Government Printing
Office, pp. 12-20.
- Orlans, Harold. Contracting for Knowledge. San Francisco: Jossey-
1973 Bass.
- Patton, Michael Quinn. Qualitative Evaluation Methods. Beverly Hills,
1980 Calif.: Sage Publications.
- Patton, Michael Quinn. Utilization-Focused Evaluation. Beverly Hills,
1978 Calif.: Sage Publications.
- Pressman, Jeffrey L. "Decisionmakers and Evaluators: Some Differences
1975 in Perspective and Possible Directions for The Future."
in Gene M. Lyons. (ed.). Social Research and Public
Policies. Hanover, New Hampshire: The Public Affairs
Center, Dartmouth College, pp. 198-205.
- Rein, Martin and Sheldon H. White. "Policy Research: Belief and
1977 Doubt." Policy Analysis, Vol. 3, No. 2 (Spring),
239-271.
- Riecken, Henry W. "Memorandum on Program Evaluation." in Carol H.
1972 Weiss. (ed.). Evaluating Action Programs. Boston:
Allyn and Bacon, pp. 88-104.
- Riecken, Henry W. and Robert F. Boruch. Social Experimentation.
1974 New York: Academic Press.
- Riker, William and Peter C. Ordeshook. An Introduction to Positive
1973 Theory. Englewood Cliffs, New Jersey: Prentice-Hall.
- Rivlin, Alice. Systematic Thinking for Social Action. Washington,
1971 D.C.: The Brookings Institution.
- Rivlin, Alice and P. Michael Timpane. Ethical and Legal Issues of
1975 Social Experimentation. Washington, D.C.: The
Brookings Institution.
- Rodman, Hyman and Ralph Kolodny. "Organizational Strains in the
1964 Researcher-Practitioner Relationship." Human Organization,
Vol. 23, 171-182.

- Rossi, Peter H. "Boobytraps and Pitfalls in the Evaluation of Social
1972 Programs." in Carol H. Weiss. (ed.). Evaluating
Action Programs. Boston: Allyn and Bacon, pp. 224-235.
- Rossi, Peter H. "Observations on the Organization of Social Research."
1972 in Peter H. Rossi and Walter Williams. (eds.).
Evaluating Social Programs. New York: Seminar Press.
- Rossi, Peter H. "Evaluating Educational Programs." The Urban Review,
1969 Vol. 3, No. 4 (February), 17-18.
- Rossi, Peter H., Howard Freeman, and Sonia Wright. Evaluation:
1979 A Systematic Approach. Beverly Hills, Calif.: Sage
Publications.
- Rossi, Peter H. and Walter Williams. (eds.). Evaluating Social
1972 Programs. New York: Seminar Press.
- Rutman, Leonard. Evaluation Research Methods. Beverly Hills,
1977 Calif.: Sage Publications.
- Sechrest, Lee. (ed.). Evaluation Studies: Review Annual--
1979 Volume 4. Beverly Hills, Calif.: Sage Publications.
- Scioli, Frank P. and Thomas J. Cook. (eds.) Methodologies for
1975 Analyzing Public Policies. Lexington, Mass.:
Lexington Books.
- Scriven, Michael. "The Methodology of Evaluation." in Carol H.
1972 Weiss. (ed.). Evaluating Action Programs. Boston:
Allyn and Bacon., pp. 123-136.
- Scriven, Michael. "Evaluating Educational Programs." The Urban
1969 Review, Vol. 3, No. 4 (February), 20-22.
- Stake, Robert E. "The Countenance of Educational Evaluation."
1967 Teachers College Record, Vol 68. No. 7 (April),
523-540.
- Stephan, A. Stephen. "Prospects and Possibilities: The New Deal
1935 and The New Social Research." Social Forces, Vol. 13,
515-521.
- Struening, Elmer L. and Marcia Guttentag. (eds.). Handbook of
1975 Evaluation Research: Volume I. Beverly Hills, Calif.:
Sage Publications.
- Suchman, Edward A. Evaluation Research: Principles and Practice in
1967 Public Service and Social Action Programs. New York:
Russell Sage Foundation.

- Suchman, Edward A. "Action for What? A Critique of Evaluation Research." in Carol H. Weiss. (ed.). Evaluating Action Programs. Boston: Allyn and Bacon, pp. 52-84.
1972
- Tullock, Gordon. The Politics of Bureaucracy. Washington, D.C.: Public Affairs Press.
1965
- U.S. Comptroller General. Federal Program Evaluations: July 1975 through June 1977. Washington, D.C.: Government Printing Office.
1977
- U.S. General Accounting Office. Status and Issues: Federal Program Evaluation. Washington, D.C.: Government Printing Office, (October).
1978
- U.S. Senate. Committee on Human Resources. Costs, Management, and Utilization of Human Resources Program Evaluation, 1977. Washington, D.C.: Government Printing Office.
1977
- Ward, David A. and Gene G. Kassebaum. "On Biting The Hand That Feeds: Some Implications of Sociological Evaluations of Correctional Effectiveness." in Carol H. Weiss. (ed.). Evaluating Action Programs. Boston: Allyn and Bacon, pp. 300-310.
1972
- Weiss, Carol H. (ed.). Using Social Research in Public Policy-Making. Lexington, Mass.: Lexington Books.
1978
- Weiss, Carol H. "Research for Policy's Sake." Policy Analysis, Vol. 3, No. 4 (Fall), 531-545.
1977
- Weiss, Carol H. "Evaluation Research in The Political Context." in Elmer Struening and Marcia Guttentag. (eds.) Handbook of Evaluation Research: Volume I. Beverly Hills, Calif.: Sage Publications.
1975
- Weiss, Carol H. "Between the Cup and The Lip." Evaluation, Vol. 1, No. 2, 49-55.
1973
- Weiss, Carol H. (ed.). Evaluating Action Programs: Readings in Social Action and Education. Boston: Allyn and Bacon.
1972a
- Weiss, Carol H. Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs, New Jersey: Prentice-Hall.
1972b
- Weiss, Carol H. "Utilization of Evaluation: Toward Comparative Study." in Carol H. Weiss. (ed.). Evaluating Action Programs. Boston: Allyn and Bacon, pp. 318-326.
1972c

- Weiss, Carol H. "The Politicization of Evaluation Research."
1970 Journal of Social Issues, Vol. 26, No. 4, 57-68.
- Weiss, Carol H. and Michael J. Bucuvalas. "The Challenge of Social
1978 Research to Decision Making." in Carol H. Weiss. (ed.).
Using Social Research in Public Policy-Making. Lexington,
Mass.: Lexington Books, pp. 213-233.
- Weiss, Janet A. "Using Social Science for Social Policy." Policy
1976 Studies, Vol. 4, No. 3 (Spring), 234-238.
- Weiss, Robert S. and Martin Rein. "The Evaluation of Broad-Aimed
1972 Programs: Difficulties in Experimental Design and An
Alternative." in Carol H. Weiss. (ed.). Evaluating
Action Programs. Boston: Allyn and Bacon, pp. 236-249.
- Wholey, Joseph F., H.G. Duffy, J.S. Fukumoto, J.M. Scanlon, M.A.
1972 Berlin, W.C. Copeland, and J.G. Zelinsky. "Proper
Organizational Relationships." in Carol H. Weiss. (ed.).
Evaluating Action Programs. Boston: Allyn and Bacon,
pp. 118-122.
- Wholey, Joseph F., John W. Scanlon, Hugh G. Duffy, James Fukumoto,
1970 and Leona M. Vogt. Federal Evaluation Policy.
Washington, D.C.: The Urban Institute.
- Williams, Walter. "The Capacity of Social Science Organizations to
1972 Perform Large-Scale Evaluative Research." in Peter Rossi
and Walter Williams. (ed.). Evaluating Social Programs.
New York: Seminar Press, pp. 287-313.

MICHIGAN STATE UNIV. LIBRARIES



31293101654121