DOCUMENT RESUME

ED 275 747 TM 860 664

AUTHOR St. John, Mark
TITLE Evaluation Design: Selecting Methods. Guide Number 14. Evaluation Guides Series.
INSTITUTION Northwest Regional Educational Lab., Portland, OR. Research on Evaluation Program.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE [85]
CONTRACT 400-80-0105
NOTE 17p.; For other guides in this series, see ED 253 952-954, ED 256 637-638, ED 257 629, and TM 860 663-666. Printed on colored paper.
PUB TYPE Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Data Analysis; Data Collection; Evaluation Criteria; *Evaluation Methods; Evaluation Needs; *Evaluation Problems; Evaluators; *Program Design; *Program Evaluation; Program Guides; *Research Tools

ABSTRACT

The choice of methods is part of the overall evaluation design process. The process consists of the following steps: (1) analyzing the problem context; (2) asking a few general questions; (3) selecting the methods (strategies) to use; and (4) selecting the specific techniques (tactics) to use. To operate successfully the evaluator needs to know many different kinds of methods, understand their purposes, appreciate their limitations, and use courage and imagination. While method reflects the strategies of the evaluator, techniques reflect the tactics. Techniques are specific and can be viewed as tools. They are used for gathering and analyzing data. The following conclusions are made about the planning and implementation of an evaluation: (1) the evaluation problem is undefined (there is no one unique solution); (2) successful evaluation design cannot be prescribed (at best, a few heuristics may prove useful); (3) the power and flexibility of the evaluator is increased as his or her repertoire of methods and skills increases. Various selecting methods, methods to help in the evaluation design, tools for gathering and analyzing information, and tools for communicating findings are presented in tables following the article. (JAZ)

THE NEED FOR A REPERTOIRE OF METHODS

When three blindmen in India set out to discover the true nature of the elephant, they all found it to be quite different--one said it was like a wall; another found it like a snake; a third was sure that elephants were like trees. By looking in one way, in one place, each thought he had understood the whole.

Evaluators have the same tendency. By identifying themselves with an approach or a method, they limit their senses and restrict their view. And since what they see depends on how they look, they are able (however accurately) to learn about only one aspect of the program. Thus, when evaluators adopt a one-dimensional approach, they end up in a position where they are able to claim "the program is like this. . ." If, on the other hand, evaluators have at their command a wide repertoire of approaches and methods, and if they know how to combine these methods appropriately and flexibly, then, at the very least, they will be able to say, "The program is like this, and this, and this . . ."

But you may argue, "Of course it is better to have more information, and therefore more views of a program, but resources are limited. We have to choose; we need to decide which methods really work, and which are really feasible and useful in the situations we face."

This is true, but it is not as absolutely true as we usually believe. That is, we do not always need to choose a single method for all situations. To understand this, consider, as an analogy, the way that a carpenter works.

A carpenter does not choose whether a hammer or saw is more useful--he knows they are both tools of his trade, and they are equally and exclusively useful, depending upon the situation. Carpenters do not find a need to debate the relative merits of the hammer and saw. Nor do carpenters have a preference for one tool over the other; they don't try to use a hammer when a saw is required.

Different evaluation methods (experiments, case studies, surveys . . .) can be seen to be like the carpenter's hammer and saw--they are the evaluator's tools. Evaluators would do well to think of themselves as artists and craftsmen, and take pride in learning to use a wide range of tools skillfully. As carpenters

do, evaluators can learn to eye a situation and know instinctively which tools will work. As their skill increases, evaluators can learn how to extend the range of use of their tools, and even how to combine their uses in innovative ways to accomplish more difficult tasks.

Carpenters' tools have evolved over centuries of continuous use; consequently, modern carpentry tools are simple, elegant, and ideally suited for their uses. By comparison, evaluators have had a short history of tool development.

Fortunately, other disciplines have developed and refined methods for evaluation in their own fields. For example, wine tasters, film critics, accreditation agencies, investigative journalists, test drivers, and senate committees have all developed approaches that are potentially useful for the educational evaluator. For several years, the Research on Evaluation Program has been collecting and adapting these kinds of methods for use in the evaluation of educational, social, or health programs. A sizeable literature describing the evaluation methods and techniques of a wide range of disciplines now exists. For the evaluator who is willing to experiment and learn new skills, these methods can become very useful tools.

GUIDELINES FOR SELECTING METHODS

The development of many new and varied methods is perhaps a mixed blessing for the evaluator. For, while it empowers the evaluator, it also makes the task of deciding what to do more difficult. When should the evaluator experiment with a new method? and which method? What is the basis on which choices are to be made?

To answer these questions, it is useful to remember that the choice of methods is part of the overall evaluation design process. As outlined in Guide Number 12, the design process consists of the following steps:

1. analyzing the problem context;
2. asking a few general questions;
3. selecting the methods (strategies) to use;
4. selecting the specific techniques (tactics) to use.

Analyzing the Context

Like any other design process, an evaluation begins with an analysis of the problem context--that is, with an investigation of the program as it is, and with an attempt to learn what is important in the setting. In a typical case, an evaluator faced with complexity may become lost in all the factors of problem context, as shown in Figure 1.

As the evaluator investigates each of these factors and begins to "get the lay of the land," a working plan (a design) begins to form in his or her mind. Moving from the general to the specific, the evaluator begins to form some general and guiding questions in his or her mind. These questions provide a focus and purpose to the evaluation. They prioritize what is important to study. Examples of such general questions include:

1. How can we best understand what is happening in this program?

2. How could this program be made to work better?

3. What are the outcomes of this program?

4. What important variations are there in the program's activities or effects?

5. How worthwhile is the program?

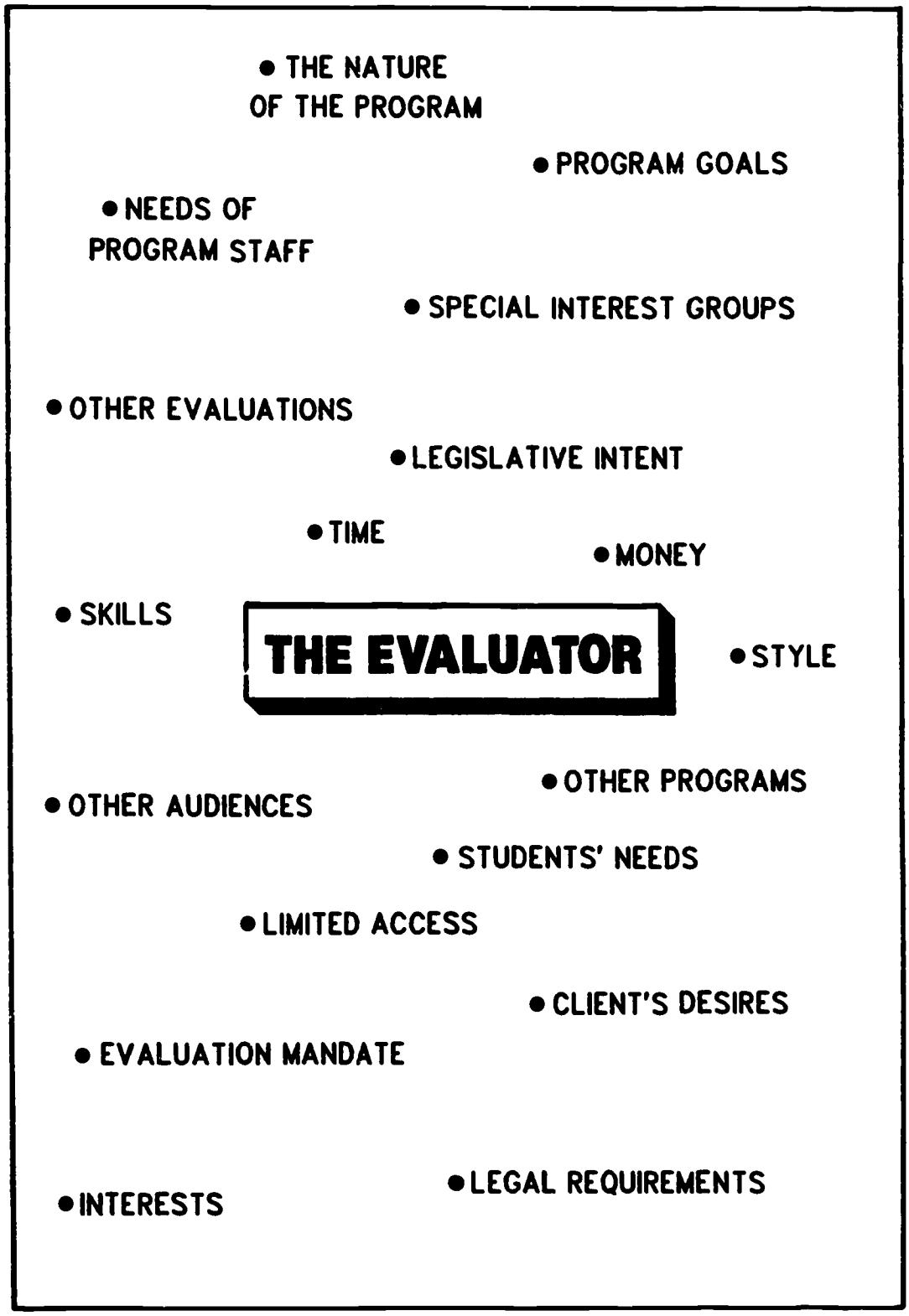
Example:

An evaluator in a state evaluation agency is given the task of evaluating a fourth grade reading program. The predominant factors in the problem context are the requirements set by the legislature for evaluation and the political pressure of parental groups for basic skills improvement. Thus, the evaluator in this case has little leeway--the general question that must be addressed is something like this: how much have the fourth grade students improved their reading skills over the year?

Example:

A museum has been given funding to train middle school science teachers. The evaluator has been called in early in the program to assist in any way possible. Here the evaluation problem context is more loosely defined, and the predominant general questions that emerge are: (1) What is happening in the program? (2) Who are these teachers? and What are their actual needs? and (3) How can the practices of the program be improved?

Figure 1


• THE NATURE OF THE PROGRAM
• PROGRAM GOALS
• NEEDS OF PROGRAM STAFF
• SPECIAL INTEREST GROUPS
• OTHER EVALUATIONS
• LEGISLATIVE INTENT
• TIME
• MONEY
• SKILLS
THE EVALUATOR
• STYLE
• OTHER PROGRAMS
• OTHER AUDIENCES
• STUDENTS' NEEDS
• LIMITED ACCESS
• CLIENT'S DESIRES
• EVALUATION MANDATE
• LEGAL REQUIREMENTS
• INTERESTS

Selecting Methods

When the evaluator begins to pose general questions, he or she is implicitly beginning to select responsive and appropriate methods. That is, certain methods are better than others for answering the relevant questions of the evaluation. In the previous example of the evaluator assessing the improvement of the reading skills of the fourth graders, experimental methods and achievement tests are obviously called for. The evaluator working with the museum would more likely use case study or journalistic methods.

The choice of methods also suggests a stance or role that the evaluator may adopt. Usually, almost unconsciously, evaluators adopt the role of the experimenter and tester. In using other methods, evaluators may bring a different approach with different perspectives and values to the evaluation. The evaluator can adopt not only the techniques, but also the perspective of the discipline from which the method is derived. Thus, in using journalistic methods, the evaluator begins to think and act like a journalist; in using methods of operations research, the evaluator takes on the role of the efficiency expert.

Techniques

If methods reflect the strategies of evaluators, techniques reflect the tactics. Methods come from the parent discipline, and they are complete with rationale, perspective, and techniques. The techniques are more specific and can be viewed as tools--the evaluator's tools. Techniques are adaptive and flexible in their use. They are used for gathering and analyzing data, for organizing findings, and for presenting results. Interviews, t-tests, surveys, research briefs, and thematic analysis are all examples of techniques used in evaluation.

Putting it all Together

As the evaluator gains an understanding of the evaluation context, general questions begin to emerge. These questions in turn suggest methods and approaches with which to structure the evaluation. Within the context of the methods, specific techniques are used by the evaluator. The chart below shows a few examples of how these elements go together to form a rough working plan.

General Question(s)	Method(s)	Technique(s)	Evaluator's Stance
What are the main issues for the program participants?	Case Study	Interviews	Anthropologist
How could the program be run more efficiently?	Modeling	Queueing theory	Operations Researcher
Is this method more effective?	Pre-post Control	Analysis of variance	Experimental Designer
What is it like to be in this program?	Story-telling	Stream of consciousness	Storyteller

This movement from general questions to methods and techniques is illustrated more fully in Table 1. Part A of Table 1 lists five general questions and associated methods that cover a wide spectrum of evaluation purposes. Part B of Table 1 suggests questions and methods that can help the evaluator in the design process itself.

Table 2 lists more specific evaluation techniques (tools), describing the discipline from which they arise and the purpose for which they are most suited.

SUMMARY

Because the evaluation process can be viewed as a problem in design, the following conclusions can be made about the planning and implementation of an evaluation:

- The evaluation problem is undefined; there is no one unique solution.
- Successful evaluation design cannot be prescribed; at best, a few heuristics may prove useful.
- The power and flexibility of the evaluator is increased as his or her repertoire of methods and skills increases.

Thus, the evaluator's search for a best method or in a best way to choose among methods is futile. Rather, the ator is left in a less well defined but freer world. To ope successfully in this world, what the evaluator needs is a k dge of many different kinds of methods, an understanding of ir different purposes, and an appreciation of their lim ions--and then, in addition, a little courage and imagination . . .

Table 1

A. Selecting Methods

I: How Can We Best Understand What is Happening in the program?
What is the nature or character of the program?
What are the conditions and activities like?
What are the central issues, themes, conflicts, trade-offs?
What seems important?

Useful Methods	Relevant Tools	General Purpose	References*
Investigative Journalism	Interviewing Tracking Shuffling, circling Files and profiles	To discover that which is important, but not immediately apparent	8, 15
Case Study	Interviewing Field observations Qualitative analysis Aggregation techniques	To gain insight into a program by understanding the many facets of the actual (not ideal) functioning of the program	
Storytelling	Oral histories Narrative techniques Stream of consciousness	To convey humanness and complexity of a program; to create images and establish mental connections	22, 26
Criticism	Thematic matrix analysis Connoisseurship Composing techniques	To illuminate forms; to demystify; to enhance sensibilities To provide an artful representation of reality	5, 6
Photography	Photo interviewing Sampling techniques	To capture, and portray images of reality; to illustrate themes or issues; to deepen insights with visual images	25
Exploratory Data Analysis	Stem and leaf displays Box plots Functional transformation of data	To discover relationships not immediately apparent in accumulated data; to select appropriate analytical methods	2

II: How Could This Program Be Made to Work Better?
Are resources being used optimally?
Where is there a critical lack of feedback?
What are the barriers to improvement?
What are the critical weaknesses?

Useful Methods	Relevant Tools	General Purpose	References*
Operations Research	Assignment model Transportation model Queueing theory	To maximize the use of existing resources	3, 16, 17
Service Delivery Assessment	Qualitative methods Debriefings Briefings	To provide policymakers near-immediate feedback about the conditions and activities of the program	9, 10
Hearings	Committee hearings Panel reviews	To gather and share testimony from involved and affected parties about the program	24

(Cont.)

Table 1 continued

III: What are the Outcomes of the Program?
What objectives are/are not met?
What "side effects" does the program appear to have?

Useful Methods	Relevant Tools	General Purpose	References*
Achievement Testing	Experimental designs Quasi-experimental designs	To assess if program has statistically significant effect on participants' skills or knowledge	
Survey Research	Questionnaires Surveys Interviews	To assess the perceptions and feelings of individuals and groups about the program	18
Product Evaluation	Critical competitor Systematic check for side effects	To make comparative judgments about product's overall quality	19
Exploratory Data Analysis	Stem and leaf displays Functional transformations Box plots	To make a search for unanticipated results and patterns in accumulated data	2

IV: What Important Variations Are There in the Program's Activities or Effects?
To what extent are different groups affected in different ways?
In what ways has the program varied over time?
How do the program's resources, services, or outcomes vary geographically?

Useful Methods	Relevant Tools	General Purpose	References*
Hearings	Committee hearings Jury methods	To gather testimony from advocates of different points of view	24
Document Analysis	Tracking Legislative history	To trace over time the changes in the operations or character of a program	2, 4, 23
Geographic Methods	Geocode analysis Trend surface analysis Social area analysis	To assess the distribution of program parameters over regional areas	14, 21

VII: How Worthwhile is the Program?
Overall, how good is the program?
Is the program cost-effective?

Useful Methods	Relevant Tools	General Purpose	References*
Product Evaluation	Needs assessment Cost analysis Synthesis procedures	To come to an overall judgment about a program's quality; to aid in decision making	19
Cost Analysis	Feasibility studies Cost-effectiveness analysis Cost-benefit analysis	To render questions of cost into useful forms; to generate information to aid decision making	12, 13, 20

(Cont.)

Table 1 continued

B. Methods to Help in the Design of the Evaluation

I: What Should the Evaluation Focus On?
What are the critical or pay-off issues?
What dimensions are important to include in the study?

Useful Methods	Relevant Tools	General Purpose	References*
Product Evaluation	Checklist	To aid the evaluator in making a comprehensive assessment of a program	19
Investigative Journalism	Quick study	To get the "lay of the land" and review relevant background	8, 15
Case Study	Observation and interview techniques	To gain insight into the whole by studying a single part	
Document Review	Legislative history Content analysis	To learn the historical or legislated intent of a program; to discover program themes or characteristics	2, 4

II: How Can We Move from the General to the Specific?
Do we agree on the meaning of key terms?
How specific do we wish to be?

Useful Methods	Relevant Tools	General Purpose	References*
Philosophical Analysis	Concept analysis	To clarify thinking about general and abstract questions; to see how concepts function in language and thought	1, 7, 11

*See reference list at end of this guide.

Table 2
Tools for the Evaluator

A. Tools for Gathering Information

Tools	Methods	Purpose
Investigative	Investigative Journalism	To confirm hunches; discover new leads
In-depth interview	Case Study	To probe: to gain insight
Testimony	Committee Hearings Panel Reviews	To gather evidence and viewpoints of different interests
Observation	Case Study Phenomenonology SDA	To obtain "snapshops" of reality; to discover patterns
Document review and tracking	Investigative Journalism Legislative History	To substantiate inferences; to learn history of issue or program
Achievement tests	Experimental Design	To determine if groups are statistically different
Operational tests	Product Evaluation	To measure the qualities of performance
Surveys and questionnaires	Market Research	To discover the distribution of opinion
Photographs	Photography	To capture images of reality

(Cont.)

Table 2 continued

B. Tools for Analyzing Information

Tools	Methods	Purpose
Factor analysis t-test	Statistical Analysis	To determine if observed differences are statistically significant
Cost analysis: feasibility effectiveness utility benefit	Cost Analysis	To determine if programs are feasible, or to measure costs against results
Operational analysis: assignment transportation queueing	Operations Research	To determine maximum use of resources; to minimize costs
Box plot Function transformation Stem and leaf display	Exploratory Data Analysis	To discover relationships, and patterns hidden in accumulated data
Geocode, Trend surface, and Social area analyses	Geographic Methods	To portray the spatial distribution of program variables
Thematic matrix analysis	Literary Criticism	To identify predominant themes
Concept analysis	Philosophy	To clarify thinking, language and ideas
Content analysis Tracking	Document Analysis	To substantiate themes; to substantiate a hypothesis
Debriefing	Service Delivery Assessment	To arrive at consensus of perceptions
Connoisseurship	Criticism	To offer personal, expert analysis and opinions
Hearings	Government Committee Hearings	To synthesize evidence in an open public format
Juries	Legal Proceedings	To judge evidence in the form of adversary testimony

(Cont.)

Table 2 continued

C. Tools for Communicating the Findings

Tools	Methods	Purpose
Narrative Prose	Storytelling	To convey the reality, humanness of program
Briefs	Journalism	To convey highlights in headline form
Graphics	Art/Design	To translate information into clear, insightful, graphic form
Maps	Geography	To illustrate relationships using mapping formats
Pictures	Photography	To use pictures to heighten sense of program reality
Oral Briefings	Service Delivery Assessment	To give oral presentation of findings
Hearings	Committee Hearings	To present all testimony and evidence publicly
Vignettes	Case Study	To present in writing typical illustrative scenarios

REFERENCES

1. Caulley, D. N. (1981). Concept analysis in evaluation (ROEP Paper and Report Series No. 61). Portland, OR: Northwest Regional Educational Laboratory.

2. Caulley, D. N. (1982). Legislative history and evaluation, Evaluation and Program Planning, 5, 45-52.

3. Caulley, D. N. (1982). The use of assignment and transportation models in evaluation (ROEP Paper and Report Series No. 68). Portland, OR: Northwest Regional Educational Laboratory.

4. Caulley, D. N. (1983). Document analysis in program evaluation, Evaluation and Program Planning, 6, 19-29.

5. Della-Piana, G. M. (1981). Literary and film criticism. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: Sage Publications.

6. Della-Piana, G. M. (1982). Film criticism and microcomputer courseware evaluation. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

7. Gowin, D. B. (1981). Philosophy. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: Sage Publications.

8. Guba, E. G. (1981). Investigative reporting. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: Sage Publications.

9. Hendricks, M. (1981). Service delivery assessment: Qualitative evaluations at the cabinet level. In N. L. Smith, Federal efforts to develop new evaluation methods (New Directions for Program Evaluation Series). San Francisco: Jossey-Bass.

10. Hendricks, M. (1982). Oral policy briefings. In N. L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: Sage Publications.

11. Lane, C. A. (1982). Using the tools of philosophy: Metaphor in action. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

12. Levin. H. M. (1981). Cost analysis. In N. L. Smith (Ed.), New techniques for evaluation. Beverly Hills, CA: Sage Publications.

13. Levin, H. M. (1983). Cost-effectiveness: A primer. Beverly Hills, CA: Sage Publications.

14. Monk, J. J. & Hastings, J. T. (1981). Geography. In N. L. Smith (Ed.), Metaphors for evaluation: Sources of new methods. Beverly Hills, CA: Sage Publications.

15. Nelson, D. E. (1982). Investigative journalism methods in educational evaluation. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

16. Page, E. B. (1979). Educational evaluation through operations research (ROEP Paper and Report Series No. 30). Portland, OR: Northwest Regional Educational Laboratory.

17. Page, E. B. (1979). Operations research as a metaphor for evaluation (ROEP Paper and Report Series No. 15). Portland, OR: Northwest Regional Educational Laboratory. (ERIC Document Reproduction Service No. ED 206 681)

18. Rasp, A. Jr. (1982). Interviewing to augment large scale survey data: The Washington high school and beyond story (ROEP Paper and Report Series No. 71). Portland, OR: Northwest Regional Educational Laboratory.

19. Scriven, M. (1981). Product evaluation. In N. L. Smith (Ed.), New techniques for evaluation. Beverly Hills, CA: Sage Publications.

20. Smith, J. K. (1983). Case reports of Northwest Regional Educational Laboratory cost studies (ROEP Paper and Report Series No. 82). Portland, OR: Northwest Regional Educational Laboratory.

21. Smith, N. L. (1979). Techniques for the analysis of geographic data in evaluation. Evaluation and Program Planning, 2, 119-126.

22. Smith, N. L. (1980). Bibliography of evaluation utilization (ROEP Paper and Report Series No. 39). Portland, OR: Northwest Regional Educational Laboratory.

23. Smith, N. L. (1982). Investigative tracking in library evaluation. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

24. Stenzel, N. (1979). Committee hearings as an evaluation format. In N. L. Smith (Ed.), Field assessments of innovative evaluation methods (New Directions in Program Evaluation Series). San Francisco: Jossey-Bass.

25. Templin, P. A. (1982). Still photography in evaluation. In N. L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: Sage Publications.

26. Wachtman, E. L. (1982). Storytelling: The narrative structure of educational evaluation. In N. L. Smith (Ed.), Communication strategies in evaluation. Beverly Hills, CA: Sage Publications.

RECENT GUIDEBOOKS IN THIS SERIES

No.	Title
1	Microcomputers and Evaluation
2	Cost-Outcome Analysis: Measuring Costs
3	Microcomputers: Word Processing
4	Cost-Outcome Analysis: Measuring Outcomes
5	Microcomputers: Statistical Analysis Software
6	Investigative Journalism Techniques
7	Microcomputers: Data Base Management Software
8	Committee Hearings: Their Use in Evaluation
9	Microcomputers: Spreadsheet Software
10	Methods of Product Evaluation
11	Microcomputers: Instrument Generation Software
12	Evaluation Design: The Evaluator and the Architect
13	Microcomputers: Communication Software
14	Evaluation Design: Selecting Methods
15	Service Delivery Assessment Techniques

Printed by the Northwest Regional Educational Laboratory, a private nonprofit corporation. The work upon which this publication is based was performed pursuant to Contract No. 400-80-0105 of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.