DOCUMENT RESUME

ED 190 620 TM 800 414

AUTHOR	Caulley, Darrel N.: Smith, Nick L.
TITLE	Field Assessment Survey. Paper and Report Series No. 10.
INSTITUTION	Northwest Regional Educational Lab., Portland, Oreg.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Nov 78
NOTE	56p.
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Elementary Secondary Education: *Evaluation Methods: *Evaluation Needs: Evaluators: Program Descriptions: Program Evaluation: School Districts: *State Departments of Education: State Surveys
IDENTIFIERS	*Evaluation Problems

ABSTRACT

Heads of evaluation units in 21 state education agencies responded to a structured telephone interview about their unit's workload, evaluation methods, and associated problems and needs. Time committed to evaluation ranged from 10% to 100% and did not relate to staff size which ranged from 1-72. Eleven states contracted some of their evaluations. The dominant evaluation method (11 states) involved setting behavioral objectives and using tests and an experimental design to assess achievement. Three states used the discrepancy evaluation model: four states audited and accredited schools: the remaining states selected the method they felt was best suited to the program. Commonly cited problems were the lack of evaluators, time, funds, training, and particularly the lack of impact of evaluations. Difficulties in designing experiments, the need for school district data management systems, and evaluation training of school personnel were also cited. Because the problems and constraints of state evaluation units are so diverse, perhaps the strategy should be to develop innovative methods and then to determine what problems the methods solve. (The telephone survey questionnaire is appended). (CP)

paper and report series

Research on Evaluation Program

Northwest Regional Educational Laboratory
710 S.W. Second Avenue
Portland, Oregon 97204
Telephone: (503) 248-6800

No. 10 FIELD ASSESSMENT SURVEY

DARREL N. CAULLEY

NICK L. SMITH

November 1978

Nick L. Smith, Director
Research on Evaluation Program
Northwest Regional Educational Laboratory
710 S. W. Second Avenue, Portland, Oregon 97204

Published by the Northwest Regional Educational Laboratory, a private nonprofit corporation. The project presented or reported herein was performed pursuant to a grant from the National Institute of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

PREFACE

The Research on Evaluation Program is a Northwest Regional Educational Laboratory project of research, development, testing, and training designed to create new evaluation methodologies for use in education. This document is one of a series of papers and reports produced by program staff, visiting scholars, adjunct scholars, and project collaborators--all members of a cooperative network of colleagues working on the development of new methodologies.

What methodological problems do state and local evaluation units encounter? What types of new methods would be most useful to them? The purpose of the field assessment survey was to identify the characteristics of innovative methodological developments which would have maximum utility for improving the evaluations carried out by local and state evaluation units. The survey attempted to determine the evaluative activities of units as well as the problems, constraints, and conditions under which they operate.

ACKNOWLEDGMENT

According to the literature, telephone surveys experience refusal rates ranging from 4 to 20 percent. In the telephone survey reported here of state education agency evaluation units, all heads of units contacted agreed to be interviewed. In addition their cooperation was excellent, in that they took time and care to give comprehensive answers to the questions asked. We acknowledge with thanks their cooperation.

Grateful acknowledgment is given to those individuals who agreed to be interviewed with trial versions of the interview schedule and who also gave us useful comments for the improvement of the schedule. These individuals include Wayne Neuburger of the Beaverton School District, Kan Yagi of the Portland Public Schools, and Nelson Noggle and Barbara Williams of the Technical Assistance Center of the Northwest Regional Educational Laboratory.

We also wish to acknowledge the fine secretarial assistance of Judith Turnidge and Tammy Gann.

Although grateful acknowledgment is given these many individuals for their cooperation and support, any inadequacies remaining in the report are the responsibility of the authors.

D.N.C.
N.L.S.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIELD ASSESSMENT SURVEY

Purpose and Rationale

Most state education departments (SEAs) and large school districts (LEAs) have a unit which is wholly or partly devoted to carrying out evaluation. These units vary in the nature and scope of the evaluation work they carry out and the size of their staffs. As well as carrying out or supervising evaluations, staffs of these units are often also involved in activities which may be related or unrelated to evaluation.

The Research on Evaluation Program at The Northwest Regional Educational Laboratory has the aim of developing new evaluation methodologies for use in SEA and LEA evaluations. The purpose of this study was to ascertain the activities, evaluative methodologies, and problems and constraints of LEA/SEA evaluation units. The information obtained should be a useful guide in the development of new evaluation methodologies.

Our ultimate aim was to try to identify the characteristics of innovative methodological developments which would have maximum utility to improving LEA/SEA evaluations. To have utility, any new methodologies must be adaptable to the functions of evaluation activities carried out by LEA/SEA units. Therefore information was required on the functions of the various evaluation activities carried out by LEA/SEA evaluation units. To have utility, ideally new methodologies should also help to solve some of the problems that LEA/SEA evaluation units experience. New evaluation methodologies should be capable of utilization under the constraints imposed by the conditions and settings which units work under. Therefore, information was required on the problems and constraints that units experience.

Information on LEA Evaluation Units

Some information on LEA units was found in the literature.

Holley (1978) and Stephens and Barber (1978), all members of LEA evaluation units, write of the difficulty of trying to serve the various clients of an evaluation. Holley (1978, p. 10) states: "Since the evaluation unit really can't afford to lose touch with any of the potential clients, more time will be required to keep all bases of communication covered. In many cases, this dictates more staff. . ." Stephens and Barber (1978, pp. 1, 5, 6, 8) state:

> In most school districts teachers have specific things that they want from an RD&E unit, while superintendents, boards of directors, business offices, and personnel offices have other sets of expectations. Once an RD&E unit defines its clients, that definition limits the services available to other portions of the institution. By limiting the services through defining the major clients, cooperation that is necessary to carry out the basic charge may actually be inhibited. . . . At this point, we need to clarify that our primary service has been to provide information to the board and to the superintendent. . . . Now at the same time, the largest number of people in the district, i.e., classroom teachers, are not really being served as our client, and so they, through their trade union, have very little interest in seeing that the RD&E unit continues. . . . It is apparent that, regardless of the primary clients, other powerful groups feel neglected if their needs are not met. How to meet the variety of needs generated from within a school district and remain with the limited budget becomes a topic of major importance.

Holley and Lee (1977) describe a number of problems encountered by a LEA evaluation unit. "Finding someone who appropriately should make a decision is one of evaluation's more difficult tasks." Those who should take actions on the basis of an evaluation report will not. Another problem is that the treatment in a program changes over time and this invalidates any experimental design that has been set up.

One of the biggest changes that occurs in a program is changes in staff. The authors express their frustration with the lack of impact of evaluation on policy making. One of the problems is how to get decision makers to read evaluation reports.

Polemeni (undated), Director of the Office of Educational Evaluation, Board of Education of the City of New York, mentions a number of problems. These are well worth reading and are given in Appendix I.

To summarize, some of the problems of LEA evaluation units mentioned in the literature are as follows.

A. Carrying Out An Evaluation

(i) The disparity between the extent of evaluation expected and the amount of funding provided.

(ii) The difficulty of trying to serve a variety of clients at one time.

(iii) The lack of cooperation of school personnel in the collection of data.

(iv) Treatment in a program changes with time and this invalidates any design that has been set up.

B. Reporting And Impact

(i) The unwillingness to publicize evaluation findings which have politically undesirable implications.

(ii) Evaluation results are often not considered in making program or management decisions and evaluations lack impact on policy making.

(iii) The problem of communicating evaluation results to laymen.

(iv) The problem of getting decision makers to read evaluation reports.

A paper by Webster and Stufflebeam (1978) characterizes and assesses the different patterns of practice in educational evaluation that have emerged in large urban school districts during the past decade. Some of the findings of this study are:

1. The vast majority of evaluation units control the testing function.

2. Input evaluation is practically nonexistent in medium and small evaluation units.

3. Process evaluation is less emphasized as evaluation units become smaller.

4. Evaluation units, regardless of size, put most efforts into testing, product evaluations, and data processing.

5. The smaller the unit, the greater the amount of time that the unit head generally must spend selling evaluation activities to decision makers, including, in many instances, convincing them that they need information to make better decisions.

6. Small evaluation departments spend a comparatively small amount of resources on ad hoc information requirements.

7. Evaluation methodology and experimental design were consistently ranked among the most important competencies expected of evaluators in evaluation units.

8. Objectives-based evaluations were carried on by every unit in the sample, particularly in conjunction with process evaluation and utilizing some form of experimental design.

9. Most large urban districts perform an accountability function.

10. Policy studies are done at one time or another by all large urban evaluation departments.

11. Most district evaluation units concentrate on providing data for decision making.

Two studies concerning LEA evaluation units are at present in progress. It is hoped that these two studies will provide the Research on Evaluation Program with further information about LEA evaluation units.

Frank Chase of the Urban Education Studies in Dallas is conducting a national study of the conditions affecting utilization of knowledge from research and evaluation in urban school districts. The purposes of the proposed study include:

1. To develop an accurate picture of the amount and types of research and evaluation conducted in city school systems during the period 1973-1978.

2. To analyze the policies and processes governing the authorization of, and the quality controls for, research and evaluation projects.

3. To appraise the mechanics and processes for the interpretation, communication, and application of the knowledge gained to educational decisions and practices.

The most extensive study of LEA evaluation units is being carried out by Lyon (1978) at the Center for the Study of Evaluation at UCLA. An extensive questionnaire has been sent to LEA evaluation units throughout the nation. Questions being asked include: What are the activities of offices of evaluation? How are evaluation unit products being used? How are evaluation offices organized? What are the characteristics of evaluation personnel? How are offices of evaluation financed? What characteristics of school districts affect evaluation offices? What resource constraints and requests are reported by heads of evaluation organizations? This study should be a particularly valuable source of information on LEA evaluation units. This information should be a useful guide to the Research on Evaluation Program in the development of new evaluation methodologies suitable for LEA evaluation units.

Since extensive work is being done in studying LEA evaluation units, it was decided not to duplicate this work, but instead to concentrate on SEA evaluation units where nobody appeared to be doing any research and little information is available. Consequently this document reports on a study of SEA evaluation units throughout the nation.

Information on SEA Evaluation Units

As indicated, there is little information available on SEA evaluation units. Plog (1978, p. 2), an evaluator in the Illinois Office of Education (an SEA), mentions the constraints of time and money. "State evaluators are often more constrained by time than researchers from a university. Constraints of money are related to time constraints. A state bureaucracy is concerned about such things as travel costs, per diem, and salaries."

In order to obtain information on SEA evaluation units, twenty-five states were contacted by telephone and the heads of evaluation were interviewed. These interviews used a set of open-ended questions and lasted on the average for twenty minutes. The following pages describe the data collection, the results of the interviews and the implications of the data.

The Data Collection

Data were collected by means of a structured telephone interview schedule which is shown in Appendix II. Advice given by Dillman (1979) was used in construction of the schedule. The initial versions of the schedule were revised on the basis of trials. Initially this was done using staff members of the Title I Evaluation Technical Assistance Center of the Northwest Regional Educational Laboratory. These staff members provide technical assistance on evaluation to LEAs and SEAs and hence are familiar with their evaluation units. Staff members played the role of respondents, and deficiencies in the schedule were detected and corrected. Next, the schedule was tried with two nearby LEA evaluations unit heads and further revisions were made.

The interviewing of the administrative heads of SEA evaluation units was done in two stages. Initially the thirteen most western states of the U.S. were contacted. These were Alaska, Hawaii, Washington, Oregon, California, Idaho, Nevada, Utah, Arizona, Montana, Wyoming, Colorado and New Mexico. The reason for contacting units in these states was that these were the units that the Northwest Regional Educational Laboratory had most contact with. It was argued that given resource limitations, these states would be closest to the Laboratory in order to trial innovative evaluation methodologies. Later it was decided to obtain an overall picture of the units in the U.S. Thus, from the remaining 37 states, a random sample of 12 were selected. These states were Kansas, Minnesota, Arkansas, Illinois, Michigan, Alabama, Florida, Virginia, New York, New Jersey, Massachusetts, and Maine. The distribution of the total of 25 states that were chosen are shown on the map of the United States as Figure 1.

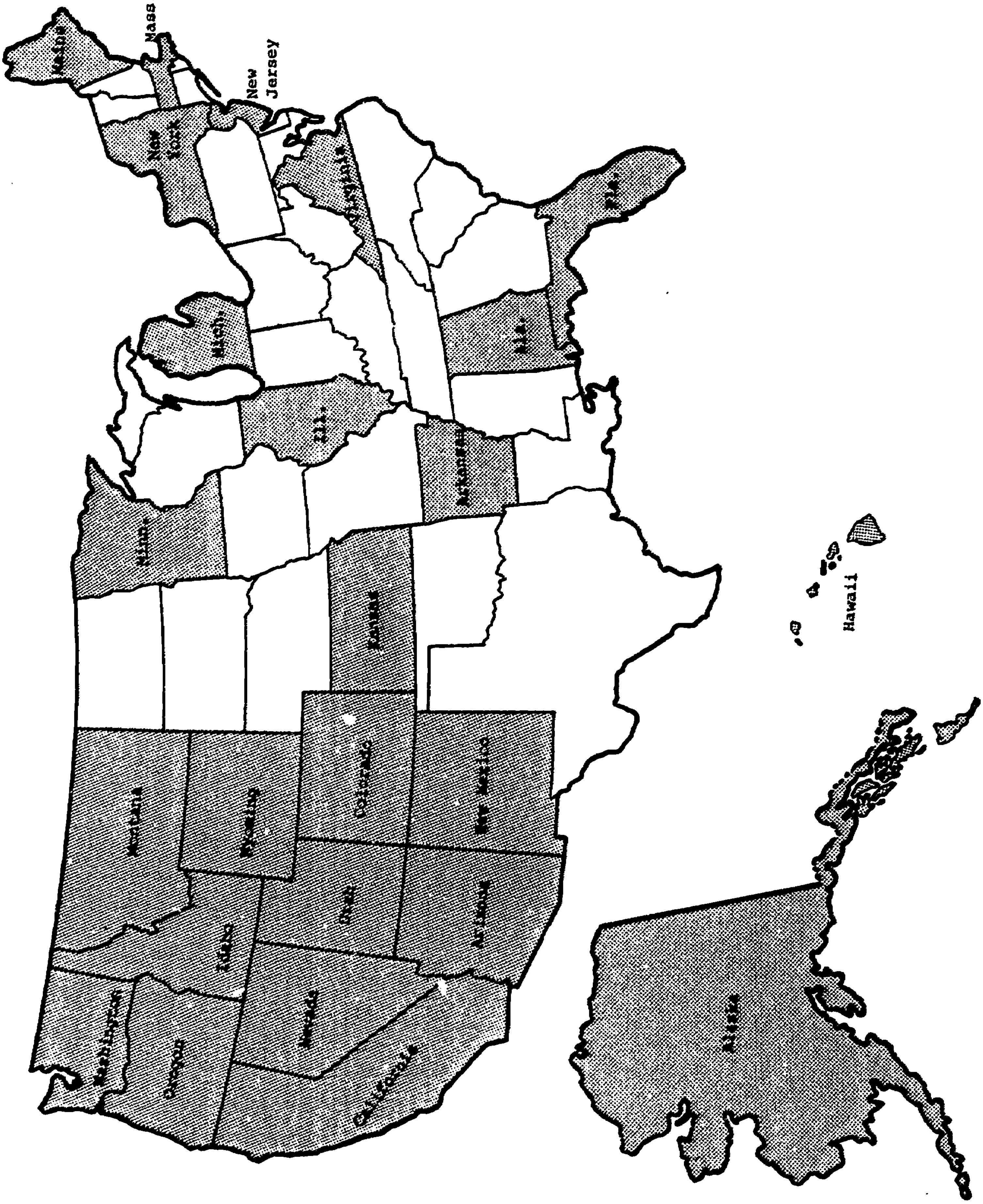
The names, titles, and telephone numbers of the administrators of SEA evaluation units were known for the western states. For the rest of the sample, the state education department telephone number was known only and the administrative head of the evaluation unit had to be located.

None of the administrative heads refused to be interviewed. Their cooperation and willingness to fully answer questions were quite outstanding. Interviews lasted from fifteen minutes to half an hour.

Results From the Survey of SEA Evaluation Units

For four of the twenty-five states the interview schedule was not appropriate. This was because these states did not have units that

Figure 1. Map of the United States Showing the States for Which the Evaluation Units Were Contacted in the Telephone Survey



carried out program evaluation. The unit head of one state indicated that his unit only contracted evaluation studies and did not carry out program evaluation studies. The unit head of a second state indicated that the unit simply supervised the statewide testing programs and carried out no program evaluation. The unit head of a third state said they carried out statewide assessment and were developing a basic skills improvement plan. He said that they did not visit schools as local autonomy was paramount. In the fourth state, staff members of the elementary and secondary divisions did visit schools and sit in on classrooms. These were in the nature of inspectorial visits. Reports were written but they could not write anything about teachers because the teachers' union was so strong.

In some states, all program evaluations were not necessarily carried out by the evaluation unit. At times, other units within a state department administration carried out evaluations, especially mandated evaluations for Federal programs.

The following sections describe the answers that were given to each of the questions of the interview schedule.

The nature of the work load and size of staffs of units. Answers to the questions related to the evaluative work load and the size of staff are shown in Table 1. It is clear from the Table that the percentage of time that the staffs of the units spend on program evaluation varies markedly from state to state, the range being from 10 to 100 percent. Unit heads often had difficulty in answering this question since it is not always easy to decide what to include in program evaluation. Examples of borderline activities are: arranging testing in the schools, evaluating facilities such as school buildings,

Table 1

The Evaluative Work Load and Size of Staff of the Evaluation Units in 25 State Departments of Education

State	Percent of Time Involved in Program Evaluation	Average No. of Evaluations Conducted Per Year	Average No. of Evaluations Contracted Per Year	The Number of Staff (Full-time Equivalents)
1	0	-	-	-
2	0	-	-	-
3	0	-	-	-
4	0	-	-	-
5	*	*	*	2 1/2
6	**	50	0	17
7	10	15-20	0	12
8	20	40	0	6
9	33.3	24	20-25	3
10	45	15	0	21
11	50	1	16	1
12	50	11	1	12
13	50	25	0	19
14	50	76	0	24
15	50-60	15	0	6
16	60	40	10-12	4
17	70	25	2	14
18	80	14	2	20
19	90	6	0	7
20	100	6	82	4
21	100	15	0	2
22	100	25	13-14	4
23	100	72	4	72
24	100	174+	0	1+
25	100	300++	Most	8

* The respondent found it impossible to answer. The respondent felt that they were not deeply into evaluation, most of their evaluations being contracted out.
**One person (full-time equivalent).
+ This person mainly serves to minitor the evaluations carried out by the local school districts and this explains the large number of evaluations.
++The evaluation unit in this state does not actually carry out evaluations but monitors and provides guidelines for others (especially local districts) to carry out evaluations.

accreditation of schools when this mainly involves filling out forms.

For those units not fully engaged in program evaluation activity, a major activity was planning. Another activity was program development. For some units, program evaluation, planning and development were seen to be interconnected activities. Overall, evaluation units were involved in a wide variety of activities including such things as test development, studies for departmental policies, accreditation, research into funding schemes, inservice workshops for school personnel and "any duties that the department requires."

Table 1 shows that the average number of evaluations conducted per year varies considerably from state to state and shows little relation to the size of the staff. There is a couple of reasons for this. One is that the size of a single evaluation is extremely variable. For example, a statewide testing program was counted as a single evaluation. A second reason is that there is a variation in the extent to which staff actually carry out evaluations themselves. For example, some units actually monitor or supervise evaluations carried out by others. As an example in Table 1, state number 16 conducted 174 evaluations which involved one staff member who monitored the evaluations carried out by local school districts. Thus, there is a great variation among states in the extent to which evaluation units are actually involved in conducting evaluation studies themselves.

Ten states did not contract out evaluations. The rest of the states varied greatly in the number of contracted evaluations.

The size of the staffs ranged from 1 to 72. The director with the 72 full-time staff indicated that his was the biggest state evaluation unit in the United States.

In summary, SEA evaluation units show a great diversity in the nature and size of their evaluative work load, and the size of their staff. Only six units engaged full-time in program evaluation. The other units engaged in a variety of activities besides program evaluation. The average number of evaluations conducted per year varies considerably from state to state and shows little relation to the size of the staff. Ten units did not contract out evaluations. The size of the staffs ranged from 1 to 72.

Strategies used in doing an evaluation. In answer to the question, "What strategies (or design) do you use in doing an evaluation?" some states indicated more than one strategy. Eleven states indicated that they used an approach which involved identifying objectives and determining the extent to which they have been achieved. Associated with this approach was the use of testing and the use of an experimental or quasi-experimental design such as involving pre- and post-testing and a control group. Stake (1976, pp. 21, 28) calls this approach student gain by testing. House (1978, p. 4) calls it the behavioral objectives approach. "The objectives of a program are spelled out in terms of specific student performances that can be reduced to specific student behaviors. These behaviors are measured by tests, . . ." Guba (1977) calls it the objectives approach or the Tylerian or neo-Tylerian approach, after its first proponent, Ralph Tyler. Three states indicated that their major strategy was the discrepancy evaluation model of Malcolm Provus (1973), which is a Neo-Tylerian approach. Some state units indicated that the use of the objectives approach affected a program so that it was better managed and it could be evaluated.

Four states are quite distinctive in that they use an auditing-accrediting approach to evaluation. While there is variation in the four states, all auditing-accrediting approaches involve a team of evaluators visiting the schools. One of the four approaches is heavily reliant on the self-evaluations done by schools. The approaches of two of the four states are described below.*

The criteria for the evaluations are minimum standards set by the state. These standards refer not just to curriculum but to a variety of aspects of the functioning of the schools. We use a team approach where all 21 evaluators go into a district. First we meet with the district administrators and explain the criteria and the evaluative process. Then we meet with teachers to do the same thing. We use a checklist which is an expanded version of the criteria. We spend one to three weeks in the schools. For schools that are selected in a district, we visit every classroom in these schools. After evaluation is complete we discuss it with teachers and administrators. Approximately 20 days after, we send back a written report and allow 30 days for any reply. If a district is not incompliance with the standards, they must develop a plan to comply.. This plan is given to the accrediting agency and is used by them when they go into the district the following year. We work on a six year cycle. There are 90 districts and we do 15 every year.

We have a requirement that all schools go through self evaluation - curriculum, buildings and grounds, administrative services etc. The evaluation unit prepares self-evaluation materials and visits the schools to explain them. A school has to prepare a five year plan and we check that they are following it. The plan goes to the accrediting committee. All elementary schools and non-accredited high schools have to go through the self-evaluation process. There are two problems with this. Firstly, it takes a lot of time for the local people to go through the self-evaluation process. They are given a year and they usually meet once a month. Secondly, community involvement is required and this is difficult to get. We have just gone through a 5 year cycle of self-evaluation. We pilot tested it for two years and got input from the local people. Now we are undergoing a revision of the process. Principals and superintendents say that the process is very useful. We are looking at ways of cutting down time for self-evaluation. School boards feel that it is very useful. They often serve on the self-study committee. We would like assistance with how better to aggregate data

*All quotes taken over the telephone are partly paraphrased.

from the reports in such a way that we can better see what the problems of schools are and the state can then respond to these problems. From several hundred reports it is difficult to discern what the problems are. Self-study leads to action on the part of those doing it.

Four states indicated that they used no particular strategy but employed the strategy that was most suited to the program being evaluated.

We use no particular strategy. We tailor the design to the program needs. We do not advocate one methodology over another. We are responsive to our target audiences.

Other evaluation strategies mentioned were as follows:

(2)* - Case studies.

(2) - Strategies oriented towards the needs of decision makers.

(2) - Monitoring programs.

Our evaluations basically involve monitoring programs to see if the programs are doing what they said they were going to do.

(2) - Monitoring evaluations.

LEAs prepare their evaluations for us. We give them guidelines which are based on RMC Model A of Title I. I also conduct workshops giving technical assistance for these evaluations.

(1) - Involve program personnel in all stages of an evaluation.

(1) - Adversary approach.

(1) - Encourage the use of Title I designs.

(1) - A three-member monitoring team.

A three-member monitoring team goes to a site visit for two days. The leader is a college professor who writes the report. The second member is an expert in the substantive area the project to be evaluated is concerned with. The third person of the team is a member of the advisory council. This team reports back to the advisory council which has oversight of all projects. The resultant evaluation report is valued by the council because it covers aspects of a project that cannot be

*Numbers in parentheses refer to the number of states involved.

measured. The report complements the evaluation report based on a quasi-experimental design.

In relation to the question, "Is there a difference in the methods used in the evaluations your unit does versus those studies you contract out?" ten states could not respond since they did not contract out evaluations. Three states said there was no difference. There were varied replies from other states, some of which are as follows:

Our evaluations are different from Title I evaluations (the contracted evaluations) which are just information collecting.

The contractor does not have the same concern about the program that we do--that it is going to work. That is why we engage in program management as well as program evaluation.

*Contract evaluations are summative with a pretest-posttest design and we attempt to backfill these evaluations with formative evaluation.**

The contracted evaluator is usually a substantive expert. Usually he does interviewing and observing and thus his evaluation is usually more informal and less structured than ours.

Yes, but it is really due to the nature of the contracted studies, not due to any difference in methodological philosophy. Contracted studies tend to be summative.

In summary, some states indicated using more than one strategy. By far the majority of states (eleven) used a behavioral objectives approach. Associated with this approach was the use of testing and the use of an experimental or quasi-experimental design. Four states used an auditing-accrediting approach to evaluation. Four states indicated that they used no particular strategy but employed the strategy that was most suited to the program being evaluated. In addition a variety of other strategies were mentioned. For those states that contracted out evaluations, most said there was some difference between contracted evaluations and evaluations which they did.

---

*By formative evaluation, the unit head appeared to mean process evaluation.

The problems and constraints of carrying out evaluations. In relation to the question, "What problems and constraints do you experience in carrying out evaluations?" most states mentioned more than one. There was a great variety of answers.

(7) - Shortage of evaluators and/or too many evaluations.

(6) - Costs mentioned were for personnel, travel, accommodation and food.

(5) - Lack of time.

(4) - Lack of training. Three states indicated that lack of knowledge of evaluation was a deficiency. Other areas of training mentioned were lack of knowledge of experimental design and statistical analysis, and lack of knowledge of the substantive area being evaluated.

(3) - The difficulty of setting up an experimental design.

The following is a list of other problems which were described. They have been divided into problems related to carrying out evaluation, problems of evaluative impact.

PROBLEMS RELATED TO CARRYING OUT AN EVALUATION

Lay persons assume that it is possible to answer questions in the social sciences like it is in the physical sciences. They assume, for example, that we can make causal connections where this is not possible. For example, a typical urban school will receive funds from multiple sources for multiple reasons. A child is receiving his education from multiple funding sources. It is difficult to partial out the effects of any one set of funds--to attribute any change to one set of funds.

Projects are planned independently of our involvement. We like to be involved in the planning of a project. We would like inservice training of project people on planning and evaluation.

The problem of defining the questions. After a set of questions have been developed, program persons will decide that they want other questions addressed. After an evaluation they will also decide that other questions should have been addressed. This effect is a result of the growth of their perceptiveness.

Schools obtain multiple requests for data--from federal and state agencies and from universities. It is difficult for a school to be responsive to all these requests. Thus at times it is difficult to get the data one requires from schools.

In the state-wide assessment survey we are seen as inundating districts with requests for information. Then we have to get tough though we prefer not to. The reason why districts are inundated is that in the districts close to the major urban area there are four institutions of teacher education. People want to collect information from the schools for Masters and Ph.D. research. In the small school districts, the administrative staff is small and is overloaded with work and so does not respond easily to information requests.

Evaluation is post hoc and it is difficult to establish a base line.

Sound sampling frames are not available.

Our problem is data collection. Data is only as good as the local schools send us. We need an improved data management system.

We have a small staff. There are 630 LEAs in our state. The LEAs do the actual evaluation. Personnel in the LEAs lack training in evaluation. For example, they think testing and evaluation are synonymous. They fear evaluation because they see it as accountability. They lack knowledge of testing and evaluation strategies.

PROBLEMS WITH REPORTING

There is a problem of producing a report that is readable by a variety of audiences. Different audiences have different needs.

We have trouble translating the evaluation results to program administrators so that action will be taken.

Problems with incorrect data in the Title I reports done by LEAs. They have not read the directions and there are typographical errors.

PROBLEMS WITH THE IMPACT OF EVALUATION

We would like to see the results of evaluations used more extensively. We would like to see change as a result of evaluation.

We don't see implementation or action taken on evaluation findings. We intend to do more followup to see if action has been taken. We often involve the program people in writing the recommendations and as a result action is more likely to be taken.

It is difficult to know what information to collect that will be useful to decision makers. It is also difficult to know when they will want the information to make decisions.

The problem of making school personnel realize the importance of evaluation and getting them involved in evaluation.

Credibility. There is a lack of trust of the department in conducting a survey and therefore we have to contract out.

Identifying evaluation studies which are critical as opposed to routine. A lot of evaluation studies are not really important. We don't get to do the evaluations we should be doing. This is because evaluations are threatening. People don't want the truth, positive or negative about a program.

It is difficult to get people to think rigorously about evaluation. They do not see it as essential but something that is a requirement that is laid on them by state and federal agencies. We would like them to get to see the utility of evaluation.

States which used an auditing-accrediting strategy experienced problems peculiar to the strategy. One state indicated that visiting schools took a lot of staff time and that it was physically and emotionally wearing. There was also a problem with the standards set by the states.

The minimum standards are general standards and are thus open to interpretation. There is a difference of interpretation within the evaluation staff and also a difference between the district personnel and the evaluators.

Another state indicated that evaluators had not previously visited the schools and were finding it difficult to adopt the auditing role. The state which emphasized self-evaluation in the auditing of schools, indicated that firstly the schools found it time consuming and secondly, community involvement, which is required, is often difficult to get. The state was also finding it difficult to aggregate data from across schools in order to see what common problems they experienced.

In summary, units gave a wide range of answers to the question of what problems and constraints they experience in carrying out evaluation. The commonest problems were shortage of evaluators and/or too many evaluations, costs and lack of time.

Strategies used in planning an evaluation. In relation to the question, "Are there any specific techniques or strategies that you use in planning an evaluation?" the following answers were obtained:

(9) - Attempt to be responsive to the audiences' needs.

We try to plan so that information will be collected that is useful to the program organizers, the superintendent and the funding source.

In planning we first of all generate questions to be answered--some of these questions are mandated and others we generate ourselves. There is a series of meetings with program managers and we determine what questions they have, especially formative ones.

First we review the written document that relates to the program and review the objectives. We then meet with the program manager and verify that these are the objectives. We then ask the program manager if he has additional questions that he would like addressed in the evaluation. We then ask the unit director, the assistant commissioner and the commissioner what questions they would like addressed. We get more data than is required by the funding source. We ask for evaluation questions from anybody in any way connected with the program.

We seek to involve the client in the evaluation such as the school district people. From department administrators we also obtain questions and issues. We also may add questions. We try to be responsive. This is important if people are going to use the data. We interact with the local district people to see how our evaluation can interface with their decision making needs.

The states indicated that their planning strategy was used in order to make the evaluation useful and to have impact.

It is important that program personnel obtain a sense of ownership of the evaluation and hence they are likely to take action on it.

(4) - Use the planning strategy associated with their auditing-accrediting approach.

(4) - Identify objectives and then design the evaluation. Two states indicated that helping program personnel to state their goals was useful to program personnel in planning their program and it also made the program more capable of being evaluated.

We attempt to get the evaluation planned at the same time that the program is planned. We try to find out what the intentions of the program planners are and help them to state their goals. This helps to make it possible to evaluate the program.

(3) - Use no particular strategy since it depends on the evaluation.

(1) - Use RMC Model A.

In summary, the commonest strategy used in planning an evaluation was one that was aimed at making the evaluation responsive to the audiences' needs. Such an evaluation was seen as more likely to be useful and to have impact.

Strategies in collecting data. In answer to the question, "Are there any specific techniques or strategies that you use in collecting data?" Many states mentioned more than one technique.

(12) - Testing.

(8) - Interviews.

(8) - Classroom observation.

(5) - Whatever data collecting strategy is appropriate to the particular evaluation.

(4) - Use some kind of document or record. For example one state uses the application form that schools fill out for what they are going to do for Title I as a source of data. Another state uses the individual education program for each child.

(3) - Questionnaires.

(1) - Uses a large, once-a-year collection of data from all schools. This has the effect of cutting down the amount of disturbance to each school.

In reply to the question, "What factors determine your choice of (data collecting) technique or strategy," most states indicated that it depended on the particular program being evaluated. Utility, credibility, and time were mentioned as factors. Other answers were:

We use the strategy which will give the hardest, most objective data, otherwise an evaluation degenerates down to the level of opinion.

Interviewing and classroom observation is better than hard statistical data.

Not wanting to disturb the school with our data collecting.

In summary, a wide range of data collecting techniques were used, the commonest being testing, interviews and classroom observation.

<u>Methods used in reporting an evaluation.</u> All states produced written reports of evaluations. Reasons given for using written reports were:

Written reports are required by clients.

Don't have a chance to talk to program personnel. We produce two written reports--one for the lay person, such as program personnel, and a technical report.

Oral reporting is limiting--it only becomes somebody's recollection. That is why we always use written reports.

All evaluations, even those that are federally funded have to be reported back to the legislators. Written reports are necessary. The state department has to forward the reports to the state board which forwards them to the legislators.

Written reports are required to meet OE requirements.

In addition to the use of formal written reports, five states also mentioned the use of written press releases. Because of public interest, one state said they had press conferences where there was a press release, an oral presentation and the use of pictorial material. The oral presentation was often done by the superintendent. At least eleven states mentioned giving some kind of oral report.

We have a reporting session with the people involved in the program. We discuss with them the questions of the evaluation, how to interpret the statistics and we assist them in making inferences. This procedure aids in the utilization of the evaluation results.

We give an informal five- to ten-minute oral presentation to the commissioner and assistant commissioners. We also give a more formal oral presentation to the state board of education. The aim of the oral presentation is to get commitment so action will be taken.

An oral report will often have more effect than a 200 page written report. The audience is more likely to take action.

Most states appeared to make an effort to report in a way that best suited their audiences in order that impact would be maximized. Thus, states often used more than one type of written report and used a combination of written and oral reports.

We try to identify audiences and target reports for the audiences. On state assessment for example, we use specific reports for specific audiences.

We prepare three reports - a one page summary, a limited page executive summary (three to five pages), and a total technical report. We always produce written reports. We also give informal and formal oral reports to the appropriate client.

First, we do an exit interview where the evaluation staff meet with the program administrators. Findings are present and interpreted. We link the result giving with the next evaluation cycle. We get questions and issues for the next cycle. Secondly, there is a written report which presents data and interpretations. Thirdly, there is an executive summary written in terms the layman can understand. This is given to the state board of education. We also at times will hold a press conference eg. for state assessment.

In summary, all states produced written reports of evaluations. In addition eleven states mentioned giving some kind of oral report.

<u>Are there evaluations that you have to do for which you have inadequate techniques or strategies?</u> Thirteen states answered in the negative. The positive replies were quite varied as the following illustrate.

At times we have trouble producing suitable tests.

We are unable to aggregate information so one can determine gains such as in Title I evaluations. Another problem is that if information from 300 schools (say) is aggregated, there is often no significant differences. However, the no significant differences ignores the fact that there are both good and bad programs. Consequently, we use case studies in order to document the programs. Another problem is that what is lacking is a set of criteria or categories for judging a program. When Consumer Reports investigates a product such as a new car, they have a checklist of criteria for judging the product.

Virtually all evaluations. In the beginning we did evaluations based on the social science research model which was a scientific model. There were enticing words like control group. The model attempted to get at causes. We pretended for a long time that this was the answer and lay persons came to expect this approach to evaluation. Now we are trying some process and case study evaluations using a modified form of Scriven's modus operandi approach.

We are not satisfied with Title IV designs.

Are there certain information needs of decision makers for which you feel that you have inadequate strategies or techniques?

(10) - Answered in the negative.

(4) - The difficulty of obtaining information, including a suitable management information system.

We do not have a comprehensive collection of information (a data base) on which to base planning.

We have massive data bases which are still hand tabulated. The legislature asks questions and as a result we take a long time to answer the questions. We need an improved data management system.

(2) - The problem of partialing out the effects of separate sources of funds when there are multiple sources of funds.

In showing program impact, for example in the evaluation of Title I by the models developed by OE and RMC. It is difficult to try to partial out the effects of Title I funds. For example we have students where the reading instruction is funded by multiple sources--local, state, migrant education, etc. It is difficult to partial the effects of these multiple sources of funds, yet policy makers want to know the impact of a single source of funds.

Other answers to the question were varied.

It is difficult to get a clear delineation of what decision makers want in the way of information. People cannot predict their needs for information well.

There is no formal way of approaching the legislature. The make decisions without information. They don't ask for evaluation.

Are there evaluation techniques or strategies that you are required to use that you think are unsuitable?

(14) - Answered in the negative.

(5) - The inadequacy of Title I evaluations.

Title I evaluations using NCEs (Normal Curve Equivalents). They are not sensitive enough for the measurement of change.

Concerned with Title I reporting model--there is over inflation of results. We see Model C as most technically sound. Model A is easier but it over inflates the results.

(1) - Have to use statistical significance but this is difficult to translate into educational significance.

(1) - Difficult to set up experimental and quasi-experimental designs, especially when random assignment is required

Are there any new techniques or strategies that you would like to try?

(7) - Answered in the negative.

(4) - Interested in trying new techniques but could not be specific.

(2) - Goal-free evaluation. One of these states said that it would be time-consuming but would produce higher quality data for decision making.

(2) - Self-evaluation techniques.

(1) - Techniques to assess affective outcomes.

(1) - "Big computer techniques" such as multivariate analysis.

Other answers to the questions were as follows:

Some of our evaluators are interested in the descriptive, anthropological approach and the use of case studies. We are interested in innovative measurement procedures especially for those things it is important to measure.

The use of unobtrusive measures to get at such things as school climate. The climate in a school is very important to the achievement of the student. It is very difficult and time consuming to get at.

We would like to try techniques and strategies that are more research oriented in the sense that we answer "why" questions about a program--going beyond a simple judgment of worth. For example, we are using the modus operandi approach.

What Kind of Assistance (If Any) With Techniques or Strategies Would You Like?

Twelve states indicated that they were not in need of any assistance. The following is a list of areas that other states would like assistance on. There was no area that was indicated in common by two or more states.

1. Assessing affective outcomes.
2. Checklists of criteria for judging the worth of a program.

3. How to report in a way that is brief but gets the important points across.

4. How to make site visits more productive.

5. Streamlined ways of getting information ready for punching and processing.

6. Assistance with goal-free evaluation.

7. We would like inservice training on new techniques.

8. Advice on the suitability and validity of test instruments.

9. In the area of sampling. How to sample student and public opinions.

10. Techniques for presenting data in an interesting way.

11. How to present data to lay persons.

12. Any kind of instruments that people have used for self study.

Summary of Results

Of the 25 state education departments contacted, four indicated that they did not have units that actually conducted program evaluation. Thus data was collected on only 21 state evaluation units.

It is clear from the data that the SEA evaluation units are extremely diverse both in their nature and the activities they carry out. They vary widely in the percentage of time they are involved in program evaluation, the average number of evaluations conducted and contracted per year, and the size of their staffs. Some of the units are involved in a multitude of activities besides program evaluation.

The most popular methodology (11 states) involved in identifying objectives and determining the extent to which they had been achieved. Associated with this approach was the use of testing. While they find it difficult, if at all possible, they attempt to set up an experimental or quasi-experimental design such as involving pre- and post-testing and a control group. Units help project managers and staff identify objectives and often assist in planning. By doing this, units aim to make projects more amenable to evaluation. Hence many units like to be involved in the planning stage of an evaluation. Four states used an auditing-accrediting approach to evaluation. Four states indicated that they used no particular strategy but employed the strategy that was most suited to the program being evaluated. Other than the commonalities mentioned above, there appears to be quite a diversity in the evaluative methodologies being used.

Units gave a wide range of answers to the question of what problems and constraints they experience in carrying out an evaluation. The commonest problems were shortage of evaluators and/or too many evaluations, costs and lack of time.

One theme emerging from the data is that relating to the impact of an evaluation. One of the problems mentioned by units was the lack of impact. One way units tried to increase impact was in the planning stage of an evaluation. They try to plan so that the information will be useful to their various clients. They seek to involve the clients in the evaluative process. Evaluation questions are solicited from likely audiences to the evaluation. Another way units seek to increase impact is in the reporting of evaluations. Most states appeared to make an effort to report in a way that best suited their audiences. Thus units often produced more than one type of written report, including a short report written in layman's terms. They also made use of oral reports to increase their impact.

A wide range of data collecting techniques was used, the commonest being testing, interviews and classroom observation.

Thirteen states answered negatively to the question of whether there are evaluations that they have to do for which they have inadequate techniques or strategies. The positive answers to this question were quite varied.

Ten states answered negatively to the question of whether there are certain information needs of decision makers for which they have inadequate strategies or techniques. Four states said that they had an inadequate information system for answering the questions of decision makers. Two states said that there is a problem of partialing out the effects of separate sources of funds when there are multiple sources of funds.

Fourteen states answered negatively to the question of whether there are evaluation techniques or strategies that they are required to use that they think are unsuitable. Five states mentioned the unsuitability of Title I designs.

In answer to the question of whether there are any new techniques or strategies that they would like to try, seven states answered negatively. Four states were interested in trying new techniques but were not specific. Two wanted to try goal-free evaluation and two wanted to try self-evaluation techniques.

Twelve states felt that they did not need assistance with techniques or strategies. Other states mentioned a wide variety of areas in which they would like assistance.

Implications of the Results

One of the major purposes of the survey was to try to infer what might be the nature of innovative methodologies that would have utility to SEA evaluation units. The following inferences are mainly based on the problems and constraints that SEA evaluation units experience in carrying out evaluations. The problems and constraints experienced by a unit will vary with the dominant methodology employed. For example states that employ an auditing-accrediting methodology will have problems that differ from states that mainly monitor evaluations. The following is a list of suggestions as to the nature of innovative methodologies that might have utility to SEA evaluation units.

1. Shortage of evaluators and/or too many evaluations, and time are interrelated problems and were mentioned by a number of states. The time it takes to do our evaluation is also related to costs. Thus a useful methodology would be one which enabled an evaluation to be carried out quickly and efficiently. Assuming time is money, this would reduce the cost of an evaluation and also enable more evaluations to be carried out. While we do not have a reference, we know that some British evaluators have been experimenting with ways of maximizing the amount of evaluative information that can be gleaned by a single day's visit to a school. Of course the speed with which an evaluation can be done is partly dependent on the attributes of the evaluator. For example an evaluator is likely to be more efficient if he is familiar with the "culture" of schools and is an expert in the substantive area being evaluated. Guba (1978b, p. 110) has reported on the use of the "fast study" in investigative journalism.

2. Methodologically, many states attempt to set up a Campbell and Stanley experimental design for an evaluation. There is a belief that this is the most rigorous methodology. However, they experience difficulties with setting up an experimental design. It is often difficult to set up control groups and to carry out random assignments.

The number of students involved may be too small to make statistical inference. It is impossible to set up an experimental design when evaluators are called in after a program has been operating for some time. Clearly what is wanted is a methodology which is equally rigorous as the experimental design but does not have its drawbacks. Guba (1978a, p. 25) argues for the rigor of naturalistic inquiry "To provide an alternative where it is impossible to meet the technical assumptions of the experimental approach in the real world." However his study of the rigor of naturalistic inquiry is still in progress.

3. Schools obtain multiple requests for data--from federal and state agencies and from universities. It is difficult for schools to be responsive to all these requests. It is clear that some states want a data management system. The nature of such a system would be that it avoids multiple requests for the same data, that it consumes as little time as possible of school personnel, that it produces as little disturbance as possible to school operations and that it be centrally located. Such a system would mean that basic data about pupils and schools would be readily available to evaluators. Evaluation studies that required sampling could be more easily carried out.

4. Lack of impact of evaluations is one of the problems that clearly troubles states. Any methodology that increased impact would certainly be welcomed. Besides their reporting strategy states attempt to increase impact through their planning strategy. This strategy is to obtain evaluative questions from a wide variety of audiences, particularly decision makers. The assumption behind this is that audiences will ask for information that will be useful to them and, if necessary, take action on it. However, as one unit head remarked:

"People cannot predict their needs for information well." Are there methodologies of planning and organizing an evaluation that would increase the impact of an evaluation? This is by no means a new question.

5. It is clear that states make strenuous efforts so that their evaluation reports will have impact that will lead to action. Ways of reporting data so that they will have most impact is clearly desired. Perhaps techniques from journalism, advertising or graphic design could be adopted for this purpose.

6. Four states use an auditing-accrediting approach to evaluating schools. Methods to streamline this approach would be appropriate. This approach requires a lot of staff, is time-consuming and is physically and emotionally wearing. One state, which used self-evaluation by the schools as well as site visits, experienced the problem that schools found self-evaluations very time consuming.

7. Some states do a considerable amount of monitoring of evaluation which is carried out by the local school districts. The problem experienced with this approach is that school district personnel lack training in evaluation and are unable to follow the guidelines provided by the state evaluation unit. Solutions to this problem could be one or more of the following.

(a) Train local district personnel in evaluation.

(b) Improve the guidelines provided.

(c) Produce simple models of evaluation that could be easily followed.

(d) Increase the support provided by the state unit.

8. Since the problems and constraints of the state evaluation units are so diverse, a certain problem often applying only to one

state, a strategy to take would be to develop innovative methodologies and then determine what problems they solve. This is an alternative to taking the problem and trying to develop a methodology to solve that problem.

In conclusion, one cannot say that the results of this survey of the methodology used by SEA evaluation units is at all surprising. One could probably do a survey of evaluators in general and come up with similar results. The behavioral objectives approach is the dominant methodology. Lack of time and tight budgets are typical enemies of the evaluator. The constant complaint of evaluators is about the lack of impact of their evaluations. But what special circumstances do SEA evaluators operate under? They are internal to the system they are evaluating, for unlike the external evaluator they do not negotiate a contract for each evaluation they do. They are captive evaluators in the sense that they do not often have a choice whether or not to do an evaluation. Within a state education department, evaluation is often closely tied to planning. This telephone survey is to some degree superficial. The next step would be to carry out intensive on-site visits to examine SEA evaluation units. Hopefully this would produce deeper insights into their operations and reveal the peculiar problems they experience.

REFERENCES

Dillman, D. A. _Mail and telephone surveys._ New York: Wiley & Sons, 1978.

Guba, E. G. Educational evaluation: The state of the art. Paper presented at the Annual Conference of the Evaluation Network, St. Louis, September 27, 1977.

Guba, E. G. Toward a methodology of naturalistic inquiry in educational evaluation. Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles, 1978a.

Guba, E. G. Metaphor adaptation report: Investigative journalism. Portland, Oregon: Research on Evaluation Program, Northwest Regional Educational Laboratory, 1978b.

Holley, F. M. Changing primary evaluation clients. Paper presented at the Annual Conference of the American Educational Research Association, Toronto, Canada, March 27, 1978.

Holley, F. M. & Lee, A. The real world of public school evaluation. Paper presented at the Annual Meeting of the American Educational Research Association, New York, April, 1977.

House, E. R. Assumptions underlying evaluation models. _Educational Researcher_, 1978, _7_, No. 3, 4-12.

Lyon, C. D. Evaluation and decision-making in school systems, _Division H Newsletter_, 1978, _4_, No. 1, 1-4.

Plog, M. The case for case studies. Paper presented at the Annual Conference of the American Educational Research Association, Toronto, Canada, March 27, 1978.

Polemeni, A. J. The politics of evaluation. New York: Office of Educational Evaluation, Board of Education of the City of New York, undated.

Provus, M. _Discrepancy evaluation._ Berkeley: McCutchan, 1973.

Stake, R. E. _Evaluating educational programmes._ Paris: OECD, 1976.

Stephens, C. E. & Barber, L. Serving some clients of research and evaluation inhibits serving others. Paper presented at the Annual Conference of the American Educational Research Association, Toronto, Canada, March 27, 1978.

Webster, W. J. & Stufflebeam, D. L. The state of theory and practice in educational evaluation in large urban school districts. Invited address at the Annual Meeting of the American Educational Research Association, Toronto, March, 1978.

APPENDICES

APPENDIX I

Some Problems of the Office of Educational Evaluation of the City of New York

Some of the problems mentioned by Polemeni (undated), Director of the Office of Educational Evaluation, Board of Education of the City of New York, are as follows:

> . . . the major problems today concern responses to the evaluation report by school superintendents, principals, teachers, unions and parent groups. The evaluation report has become politicized and has become the source of problems which must be faced by evaluators as they attempt to protect their findings from special interest groups. For example, there is often an unwillingness to publicize results which, though truly illustrative of the situation, might have negative political repercussions. It might be found, for instance, that ability grouping produces maximum growth in academic achievement. If, however, ability grouping would result in racially segregated classrooms, then such findings would, in many localities, be considered political anathema. . .
>
> For a variety of reasons, evaluation results are often not considered in making program or management decisions: The results may be considered politically inexpedient; the results are available too late for incorporation in the recycling design; the person in a key management position simply does not agree, philosophically, with the results. Whatever the cause, the net result is the same: The evaluation might just as well never have been conducted. This situation gives rise to a commitment problem in that the morale of evaluation workers suffers when they become aware that their work may be in vain. As in most such situations, reduced morale results in reduced quality of output. . . .
>
> Laymen--particularly on Boards of Education--demand gross over-simplification in reports of evaluation studies. . . . the evaluator trying to stress the limitations of his study, the lay user trying to make the findings universally applicable to satisfy his own purpose. . . .
>
> There exist no effective channels of communication between evaluators and field personnel and, as a consequence, research findings are seldom implemented. . . .

Frequently there is resistance to evaluation findings by supervisory personnel--including school principals--who "know in their bones" that the way they are doing it is the best way it can be done. This situation leads them to say such things as, "I know your findings prove my reading program is not working, but I feel the children are getting something out of it and I am going to stick with it." . . . .

Classroom teachers frequently oppose the collection of data because they are unable to see any profit to their own students. This is almost always the case with control groups from whom vast amounts of data must be taken without any program to compensate them for their time. In large degree their reluctance is well-founded since the period between data collection and report dissemination usually runs a year or more and the students who contribute the data are no longer with the teacher. . . .

Project managers are threatened by evaluation since, if the evaluation is negative, they might be out of a job. Such consideration causes all sorts of things to happen: data disappear, project personnel are unavailable for interview, students to be observed have suddenly gone on a class trip, the evaluator is incompetent, the evaluator is biased, and so forth, and so forth. While it is not absolutely impossible to evaluate a program without the project manager's approval, it is extremely difficult. . . .

School people are, increasingly, demanding that if evaluators are going to collect the data, they also provide a prescription for upgrading the program. . . .

Still another vice for the unwary is the one that says "Good news won't sell papers." This pithy aphorism probably accounts for the media's propensity to simplify or alter research and evaluation data. No matter how much the scores might have gone up within the city, there's always one school where they went down. Let's run that headline, Charlie. . . . What is needed is a balanced presentation; nothing is either all good or all bad. . . .

There is frequently a tremendous disparity between the amount of evaluation data required and the amount of funding provided for the conduct of the evaluation. This happens most often where a figure such as one-half of one percent is determined to be appropriate for the evaluation of each individual program. Where the program costs run to several hundred thousand dollars or more, this base rate is totally meaningful and applicable. Where the cost of a program runs to figures like thirty thousand dollars, the evaluation agency is being requested to perform its function for about one hundred and fifty dollars--which is absurd on the face of it.

APPENDIX II

Structured Interview Schedule for a Telephone Survey of SEA Administrators of Evaluation Units

Name ______________________________ Phone No. ______________

Name of State ______________________________

Presumably through secretary, identify and contact above person.

Say:*

HELLO. THIS IS DARREL CAULLEY OF THE NORTHWEST REGIONAL EDUCATIONAL LABORATORY IN PORTLAND, OREGON. *WE HAVE A FEDERAL GRANT TO DEVELOP NEW EVALUATION METHODS FOR LEA AND SEA EVALUATIONS. AS A PART OF THAT WORK WE ARE INTERVIEWING A FEW SENIOR ADMINISTRATORS IN LEAS AND SEAS TO BETTER UNDERSTAND THE PROBLEMS AND CONSTRAINTS THEIR STAFFS EXPERIENCE IN CARRYING OUT EVALUATIONS. I CAN ALSO TELL YOU A LITTLE ABOUT THE NEW KINDS OF METHODS WE ARE DEVELOPING.

SINCE WE ARE TRYING TO DEVELOP NEW METHODS THAT CAN BE OF USE IN OPERATIONS LIKE YOURS, I'D LIKE TO ASK YOU A FEW QUESTIONS ABOUT IT. THESE QUESTIONS SHOULD TAKE NO LONGER THAN 20 MINUTES. DO YOU HAVE THE TIME NOW OR SHOULD I CALL BACK LATER TODAY? (if late in the day, say EARLY TOMORROW MORNING)

CALL BACK: HI, THIS IS DARREL CAULLEY OF THE NORTHWEST REGIONAL EDUCATIONAL LAB. I TALKED WITH YOU ____________ AND YOU ASKED ME TO CALL BACK. AS I MENTIONED BEFORE, WE (to *). . . .

*Only the words in upper case are to be spoken over the telephone.

DATE	TIME	RESULT

Examples of Results:

1. Not in office. Will return at a certain time
2. Busy-Inconvenient. Phone back at specified time.
3. Refused.
4. Interview completed.

A.1 WHAT PERCENT OF THE TIME DOES YOUR STAFF SPEND IN DOING OR CONTRACTING PROGRAM EVALUATION STUDIES LIKE TITLE I EVALUATIONS OR EVALUATIONS OF, FOR EXAMPLE READING PROGRAMS? ______%

A.2 IN GENERAL, WHAT OTHER ACTIVITIES DOES YOUR UNIT ENGAGE IN?

______________________________ ______%

______________________________ ______%

______________________________ ______%

______________________________ ______%

A.3 (a) ON THE AVERAGE, APPROXIMATELY HOW MANY EVALUATIONS DO YOU CONDUCT IN A YEAR? ______

(b) HOW MANY EVALUATIONS DO YOU CONTRACT? ______

A.4 HOW LARGE IS YOUR STAFF? ______________Full-time equivalents

A.5 (i) WHAT STRATEGIES (OR DESIGN) DO YOU USE IN DOING AN EVALUATION?

Possible Answers:		Rank Order
(a) Identify objectives and determine the extent they have been achieved	____	____
(b) Do an experimental or quasi-experimental design, e.g., pretest-posttest, experimental-control group	____	____
(c) Testing students to see how they compare to norms	____	____
(d) Do a survey using a questionnaire	____	____
(e) Do interviewing and/or observations to find out how a program is working	____	____
(f) Other ______________________________		____
(g) Other ______________________________		____

(ii) COULD YOU PLEASE RANK THE STRATEGIES IN ORDER FROM T.E MOST USED TO LEAST USED.

On the above place the rank using 1 for the most used.

(iii) IS THERE A DIFFERENCE IN THE METHODS USED IN THE EVALUATIONS YOUR UNIT DOES VERSUS THOSE STUDIES YOU CONTRACT OUT?

B. WHAT PROBLEMS AND CONSTRAINTS DO YOU EXPERIENCE IN CARRYING OUT EVALUATIONS?

Possible Answers:

(i) Time

WHAT ASPECTS OF EVALUATIONS ARE TIME CONSUMING? ____

- (a) Planning the evaluation ____
- (b) Collecting data ____
- (c) Analyzing data ____
- (d) Writing reports ____
- (e) Travel to and from site ____
- (f) Other ______________________________

____

(ii) Cost

WHAT ARE YOUR MAJOR COSTS? ____

- (a) Personnel ____
- (b) Travel ____
- (c) Accommodation and food ____
- (d) Costs of analyzing data ____
- (e) Costs of printing reports ____
- (f) Acquisiton of tests ____
- (g) Other ______________________________

____

(iii) Shortage of evaluators and/or too many evaluations ____

PLEASE ELABORATE

(iv) Lack of training of personnel ___

WHAT AREAS OF TRAINING DO THEY LACK?

Instructional product development	___	Stochastic processes	___
Planning	___	Bayesian analysis	___
Management	___	Multivariate inferential stat	___
Communications	___	Univariate inferential stat	___
Operations research	___	Multivariate descriptive stat	___
Econometrics	___	Univariate descriptive stat	___
Path analysis	___	Politics of evaluation	___
Cost effectiveness	___	Objectives development	___
Case study	___	Evaluation theory	___
Historical research	___	Evaluation methodology	___
Experimental design	___	Instrument development	___
Library research	___	Scaling	___
Computer programming	___	Measurement theory	___
Canned program usage	___	Testing applications	___
Other ___		Other ___	

(v) Lack of cooperation ___

FROM WHOM?

School superintendents ___

Principals ___

Teachers ___

Students ___

Parents ___

Other ___

WHAT FORM DOES THE LACK OF COOPERATION TAKE AND WHAT ARE THE REASONS FOR IT?

(vi) Other ___

___

(vii) Other ___

___

C. 1. ARE THERE ANY SPECIFIC TECHNIQUES OR STRATEGIES THAT YOU USE IN PLANNING AN EVALUATION?

WHAT FACTORS DETERMINE YOUR CHOICE OF TECHNIQUE OR STRATEGY?

Possible Answers:

(a) Utility ____
(b) Credibility ____
(c) Time ____
(d) Costs ____
(e) Manpower availability ____
(f) Skills of the evaluator ____
(g) Ease of obtaining cooperation ____
(h) Other ____________________ ____

2. ARE THERE ANY SPECIFIC TECHNIQUES OR STRATEGIES THAT YOU USE IN COLLECTING DATA?

Possible Answers:

(a) Testing ____
(b) Classroom observation ____
(c) Questionnaires ____
(d) Interviews ____
(e) Other ____________________ ____

(f) Other ____________________ ____

WHAT FACTORS DETERMINE YOUR CHOICE OF TECHNIQUE OR STRATEGY?

Possible Answers:

(a) Utility ____
(b) Credibility ____
(c) Time ____
(d) Costs ____
(e) Manpower availability ____
(f) Skills of the evaluator ____
(g) Ease of obtaining cooperation ____
(h) Ease of analysis ____
(i) Other ____________________ ____

3. (i) WHAT METHODS DO YOU USE IN REPORTING AN EVALUATION?

Possible Answers:

(a) Written ____
(b) Oral ____
(c) Pictorial ____
(d) Other ______________________________
(e) Combination of the above ______________

(ii) WHICH TWO METHODS DO YOU MOST USE?

(iii) WHAT FACTORS DETERMINE YOUR CHOICE OF METHOD?

D. ARE THERE EVALUATIONS THAT YOU HAVE TO DO FOR WHICH YOU FEEL YOU HAVE INADEQUATE TECHNIQUES OR STRATEGIES? GIVE EXAMPLES.

E. ARE THERE CERTAIN INFORMATION NEEDS OF DECISION MAKERS FOR WHICH YOU FEEL THAT YOU HAVE INADEQUATE STRATEGIES OR TECHNIQUES? GIVE EXAMPLES.

F. ARE THERE EVALUATION TECHNIQUES OR STRATEGIES THAT YOU ARE REQUIRED TO USE THAT YOU THINK ARE UNSUITABLE? GIVE EXAMPLES.

G. ARE THERE ANY NEW TECHNIQUES OR STRATEGIES THAT YOU WOULD LIKE TO TRY?

H. WHAT KIND OF ASSISTANCE (IF ANY) WITH TECHNIQUES OR STRATEGIES WOULD YOU LIKE?

I. WHAT SHOULD WE KNOW ABOUT YOUR EVALUATION TECHNIQUES OR STRATEGIES THAT WE DID NOT ASK?

WE APPRECIATE YOUR COOPERATION IN ANSWERING THESE QUESTIONS. YOUR NAME OR UNIT WILL NOT BE IDENTIFIED IN ANY REPORT OF THE RESULTS, BUT I WOULD BE PLEASED TO SEND YOU A COPY OF THE RESULTS WHEN THE STUDY IS COMPLETED. ALSO, IF YOU WOULD LIKE, I COULD PLACE YOUR NAME ON OUR PERMANENT PROGRAM MAILING LIST TO RECEIVE OUR NEWSLETTER AND OTHER NOTICES OF NEW METHODS.

WOULD YOU LIKE A COPY OF THE SURVEY RESULTS?

____No ____Yes*

WOULD YOU LIKE TO BE ON OUR MAILING LIST?

____No ____Yes*

*Address ____

____

THANK YOU AGAIN. GOOD BYE.