DOCUMENT RESUME

ED 249 724 EC 170 826

TITLE	Program Evaluation for Effective Special Education. Research & Resources on Special Education: Issue IV.
INSTITUTION	merican Association of School Administrators, lington, Va.; Council of Chief State School Officers, Washington, D.C.; National Association of State Boards of Education, Alexandria, VA.; National Conference of State Legislatures, Washington, D.C.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC. Div. of Educational Services.
PUB DATE	Mar 84
NOTE	23p.
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Compliance (Legal); *Disabilities; Elementary Secondary Education; *Evaluation Methods; Program Administration; *Program Evaluation; *Research Methodology

ABSTRACT

The paper addresses issues in special education program evaluation. Introductory information examines the mandate for state monitoring under P.L. 94-142, The Education For All Handicapped Children Act, and reviews state monitoring approaches undertaken in Nebraska, Missouri, and North Carolina. A nine-step procedure designed to help policymakers become more effective consumers of evaluation reports is provided. Checkpoints for policymakers are listed under each of the following steps: reading the complete evaluation report, reviewing the evaluation model, reviewing the evaluation's goals and objectives, reviewing the evaluation criteria, reviewing the population and sample selection, reviewing the evaluation design, reviewing the data collection and analysis process, reviewing the overall evaluation report, and designing a program improvement management plan. Also included is a glossary of approximately 40 evaluation research terms. (CL)

RESEARCH & RESOURCES ON SPECIAL EDUCATION

ISSUE IV
MARCH 1984

PROGRAM EVALUATION FOR EFFECTIVE SPECIAL EDUCATION

Special education programs in the United States have expanded from state special schools for the deaf and blind founded in the late 1800s, to well over 15,000 state-authorized public school programs in 1983. These programs serve over three million pupils annually at a cost of millions of dollars of federal, state and local monies.

Over the last hundred years, evaluation of these programs has also expanded, from no evaluation of special education programs to intensive evaluations for both local and state use. Improvement of the quality of educational programming for students is the single most desirable outcome of evaluation. Program evaluation also helps assure accountability, and assists planning for budgetary, personnel, facility and materials needs. It provides local, state and federal decisionmakers with critical information; it fosters communication among diverse audiences; it clarifies program objectives and accomplishments.

Although there are many journal articles and monographs about evaluation theory and practice, there are few resources available to help policymakers be more thoughtful and effective "consumers" of the special education program evaluations they receive. This issue brief addresses this need through a focus on the following questions and their implications for state and local education policymakers:

- WHAT DOES SPECIAL EDUCATION PROGRAM EVALUATION MEAN?
- HOW ARE SOME STATES HELPING LOCAL DISTRICTS TO ASSESS AND IMPROVE THE QUALITY OF EDUCATIONAL PROGRAMMING FOR HANDICAPPED CHILDREN?
- WHAT ARE THE ESSENTIALS OF AN ADEQUATE EVALUATION?

A project for policymakers administered by the **National Association of State Boards of Education** *in cooperation with the Council of Chief State School Officers, National Conference of State Legislatures, and American Association of School Administrators.* **701 N. Fairfax St., Suite 340 Alexandria, VA 22314 (703)684-4000**

WHAT DOES SPECIAL EDUCATION PROGRAM EVALUATION MEAN?

The Mandate for State Monitoring Under P.L. 94-142

The intent of the monitoring provisions in P.L. 94-142 is to assure a single line of accountability for the education of all handicapped children within a state. The federal statute thus provides a broad mandate for state education agencies (SEAs), requiring that they:

> "...provide for procedures for evaluation at least annually of the effectiveness of programs in meeting the educational needs of handicapped children (including evaluation of individualized educational programs), in accordance with such criteria that the Commissioner shall prescribe...." (Section 613.(a)(11))

This mandate is made more specific in the federal regulations:

> "A general application must include assurances, satisfactory to the Commissioner...
>
> (3) That the state will adopt and use proper methods of administering each program, including:
>
> i) Monitoring of agencies, institutions, and organizations responsible for carrying out each program, and the enforcement of any obligations imposed on those agencies, institutions, and organizations under law;
>
> ii) Providing technical assistance, if necessary, to those agencies, institutions, and organizations;
>
> iii) Encouraging the adoption of promising or innovative educational techniques by those agencies, institutions, and organizations;
>
> iv) The dissemination throughout the state of information on program requirements and successful practices; and
>
> v) The correction of deficiencies in program operations that are identified through monitoring or evaluation." (Education Division General Administrative Regulations, Section 100b. 101 (3) (3))

Since P.L. 94-142's passage, SEAs have implemented monitoring systems designed to generate the information required by federal law and regulation. Underlying this approach is an assumption that compliance with P.L. 94-142's specific requirements will assure quality programming.

Assessing the Quality of Special Education Programs

Recently, state monitoring experiences have led many SEA officials to argue that "compliance monitoring efforts, at least as they have been conducted in the past, have missed some fundamental, qualitative aspects of educational programs for handicapped children" (Farrow, 1983,8). For example, it is relatively easy

(Continued on inside back cover)

to observe whether the necessary procedures are in place to develop an individualized education program (IEP) and make a placement decision. It is comparatively simple to note whether timelines are observed, whether parents are duly notified, and whether all written documents are in the student's file. However, a local education agency (LEA) can meet all these requirements and still be placing students in inappropriate settings or failing to develop the resources necessary to assure education in the least restrictive environment.

A growing number of state and local education agencies are ready to move beyond simply looking at procedural measures of program performance and are seeking to prove that their programs are not just "working"--but are effective and of high quality. This has led to a series of state and local efforts to develop evaluation procedures which measure the quality of special education programs.

Issues in Quality Evaluation

Although the issue of quality evaluation has become more important in response to both local program needs and national pressures on education, a number of factors make quality evaluation difficult (Farrow, 1983).

First, there is no consensus within either regular or special education about the meaning of education quality. Similarly, the approaches to evaluating quality are as varied as the goals by which different people attempt to define it. For example, should quality be judged solely in terms of program outcomes such as student achievement? Are cost factors relevant to assessing quality, or should evaluative judgments be made free of resource considerations?

Second, attempts to measure program quality are complicated by the need to define their relationship to compliance monitoring. Should these efforts be integrated into one system, or can they exist independently? Since resources are limited, would an emphasis on quality evaluation necessarily involve less intensive compliance monitoring?

Third, there is a paucity of materials designed to assess the quality of special education programs. Thus, the decision to measure program quality usually includes a commitment to develop appropriate evaluation materials as well.

Fourth, any attempt to assess program quality may encounter resistance from those whose work is being evaluated. While everyone endorses the concept of quality evaluation, the risk that results may be unfavorable rather than favorable can be intimidating.

In addition, research suggests that other factors such as the nature of SEA leadership, and the involvement of local "users" of the evaluation need to be considered and resolved to assure maximum success of any program quality assessment.

****

REFERENCE

Farrow, F. Effective state monitoring policies (Quality monitoring and monitoring of state operated programs). Washington, D.C.: The Center for the Study of Social Policy, 1983.

****

* *

ABOUT THE PROJECT

This material is made available through NASBE's Special Education Dissemination Project. Working in cooperation with the Council of Chief State School Officers, National Conference of State Legislatures, and American Association of School Administrators, NASBE has undertaken a variety of activities aimed at providing education policymakers with research and practice-based information on special education.

The project is funded by the Division of Educational Services, Special Education Programs, U.S. Department of Education. However, the views expressed herein do not necessarily reflect the position or policies of that Department. This material may be reproduced.

For more information about the project, contact Roberta Felker, Dinah Wiley or Cynthia Chambers at NASBE.

* *

QUALITY EVALUATION:
A CONSUMER'S GUIDE FOR POLICYMAKERS

Education policymakers' increasing concern with program quality is likely to lead to an increased demand for--and receipt of--systematic evaluation information. Many states have produced manuals detailing the process for obtaining this information, such as North Carolina's Special Education Program Quality Manual (North Carolina Department of Public Instruction, 1983); Maryland's Evaluation Resource Manual (Maryland State Department of Education, 1980); and Massachusetts' Special Education Program Manual: A Management Tool (Massachusetts Department of Education, 1981).

However, there are few resources available to assist policymakers in making sense of the often extensive reports which result from such processes. The following "checklist" is intended to help policymakers be more thoughtful and effective consumers of the evaluation reports they read by facilitating a clearer understanding of the components of an evaluation, and of the credibility of the reported results. The checklist outlines a series of steps which are arranged roughly in the order in which they appear in most evaluation reports. Each step includes representative activities, and provides "checkpoints"--questions to help policymakers assess each step's activities.

STEP 1: Read the complete evaluation report.

In order to assess the quality and programmatic implications of an evaluation report, it is necessary first to become familiar with the program itself, and the way in which it is evaluated. This is accomplished most readily through an overview reading of the report.

Checkpoints for Policymakers

- Is it clear exactly what program or program components the evaluation report is assessing?
- Can you describe the types of information contained in the report and in the appendices?
- Can you readily locate specific sections of the report?

STEP 2: Review the evaluation model.

Every evaluation has an implicit or explicit "model," an approach which defines the focus of the evaluation, including which questions are asked. Different models or approaches ask different questions, use data in different ways and require different levels of time, personnel, and resource commitment. Thus, the selection of an evaluation model defines the scope of the evaluation, frames the interpretation of the data, and helps everyone understand what types of information they can and cannot expect to receive.

A project for policymakers administered by the **National Association of State Boards of Education** *in cooperation with the Council of Chief State School Officers, National Conference of State Legislatures, and American Association of School Administrators* **701 N. Fairfax St., Suite 340 Alexandria, VA 22314 (703)684-4000**

Checkpoints for Policymakers

- Can you identify the particular evaluation model or approach which is used as the basis for the evaluation?
- Is a rationale included for its selection which clearly relates the assumptions, procedures and resource demands of the model to the assumptions, procedures and available resources of the special education program?
- Is it clear how the evaluation model guides the program evaluation procedures?

STEP 3: Review the evaluation's goals and objectives.

The purpose of the evaluation goals and objectives is to provide a framework for describing program results. Usually, the objective is to determine how well a program is succeeding in reaching its goals. Goal statements are generally broad in scope and reflect long-term aims of the program. The evaluator may narrow the goal statements to define one purpose or objective of the program, specify the objective in measurable terms, and define the timeframe within which the objective is to be accomplished.

The goals and objectives of the evaluation provide the focus for specific evaluation questions. For example, the CIPP Model (Stufflebeam et. al., 1971) delineates four broad goal areas from which evaluation questions are derived.

1) Context Evaluation

- What are the unmet needs of the special education student?
- What program obstacles or constraints exist that impede meeting these needs?
- What program goals or objectives are necessary to meet these needs?

2) Input Evaluation

- Does the special education program have the necessary prerequisite resources to deliver the needed services?
- What strategies will be employed to meet the program goals?
- How will these strategies be implemented?

3) Process Evaluation

- To what extent has the program implemented activities which were designed to meet program goals?
- What do existing program activities look like?

4) Product Evaluation

- To what extent have program goals been accomplished?
- How confident can we be that observed changes are a result of the special education program?

Checkpoints for Policymakers

- Do the evaluation goals, objectives and questions relate to the evaluation model selected?
- Are the goals and objectives of the evaluation clearly linked to the goals and objectives of the education program?
- Are the evaluation objectives stated in such a way that you can tell if they have been accomplished?
- Are the goals and objectives which you feel are important to investigate included? Do you have the data you want and need?

STEP 4: Review the evaluation criteria.

For a particular aspect of a program to be considered successful, it must meet certain expectations. These expectations are the evaluation criteria. Criteria or standards can be of two general types: 1) qualitative or descriptive standards, such as "commensurate with the student's abilities;" and 2) quantitative or numerical standards, such as "75 percent satisfactory." Standards can be established through the use of specialists, through a review of past performance to determine reasonable expectations, through reliance on measures of improvement, or through reliance on established norms and practices. Sometimes, the standards are stated in the policy under which the program was adopted.

Whatever standards are chosen, it is essential that they be agreed to by everyone involved in the evaluation, including the audience(s) for the evaluation report. If the standards for judging the program are not seen as credible and valued, the evaluation itself will not be taken seriously.

Checkpoints for Policymakers

- Is the type of standard to be used as the criterion for success clear for each objective?

- Is the way in which the standard was selected clear?

- Do you agree with the standard selected, i.e., do you feel the results as measured by this standard are believable and valued?

STEP 5: Review the population and the sample selection.

The population is the group about whom the evaluator is interested in gathering information and drawing conclusions. For example, a population might include "all students eligible for special education programs." When it is not feasible to include the total population in an evaluation, a sample must be drawn by specifying who will be included in or excluded from the population being studied. There are many commonly used sampling procedures (e.g., simple random sampling, stratified random sampling), all of which help to assure that the sample selected accurately represents the population being studied.

Frequently, including the entire population in an evaluation study involves a substantial commitment of time, personnel and money. The critical concern is to include enough individuals to produce a believable estimate of program effectiveness.

Checkpoints for Policymakers

- Is the population and/or sample for the evaluation clearly specified?

- If a sample is used, is the sampling procedure clearly delineated and a rationale provided?

- Are all groups which are affected by the program and which you believe are important included in the population and/or sample?

STEP 6: Review the evaluation design.

The purpose of evaluation is to provide a means of making valid decisions and judgments about a program's effect on the "real world." Viewed broadly, evaluation studies are either quantitative, based on the principles of experimental design (cf. Campbell and Stanley, 1963); or qualitative, drawn from social science field methods (cf. Patton, 1978).

Quantitative approaches assume the necessity, desirability and possibility of applying empirical standards to programs and problems involving human beings; they include experimental and quasi-experimental designs. Qualitative approaches, on the other hand, assume that human interactions are not all amenable to numerical and statistical reduction; they include case studies and other naturalistic methods which try to capture the meaning and successes of the program in descriptive terms. These two approaches to evaluation design are not discrete; it is not

necessary to choose between the two. Rather, both offer useful alternatives since different kinds of problems and questions require different approaches.

Ultimately, consideration must also be given to what the evaluation audiences believe constitutes valid and reliable findings or results. Measurement and design decisions are usually made within an explicitly political context; they are not simply a matter of expertly selecting the "best" techniques. As Patton observed, "Design and data collection decisions are a far cry from being neutral, objective, or rational; such decisions are political, subjective and satisficing" (Patton, 1978, 202).

Checkpoints for Policymakers

- Is the selection of an evaluation design clear and justified by the aims of the program, the types of information needed, and the timeframe and data collection burden involved?

- Do you believe that the types of information provided (e.g., statistical, descriptive) are valid and reliable data upon which you would be willing to take action?

STEP 7: Review the data collection and analyses processes.

It is possible to collect evaluation information in a variety of ways, such as questionnaires, observations, interviews, rating scales, document review, and performance tests. Care should be taken so that the method or methods chosen enhance the reliability and validity of both the data collected and the total evaluation design.

Once evaluation information has been collected, data analysis can be conducted. Most LEA and SEA procedures rely primarily on descriptive statistics such as frequencies, percentages and means. Quotations, illustrations and descriptions may be included to add to the readability and believability of the statistical data. One of the most important standards for data analysis is, "Is the evaluation information presented in the most clear and useful way possible?"

Checkpoints for Policymakers

- Does the evaluation address the extent to which the instruments measure what they are supposed to measure (i.e., are they valid)?

- Does the instrument design give you confidence that if the instruments were administered more than once to similar groups (or to the same group) they would yield consistent results (i.e., are they reliable)?

- Do the ways in which the data are analyzed and presented reflect the program objectives and make sense to you? Are the types of data provided (e.g., percent correct, percentile rank, grade equivalents) of use to you?

STEP 8: Review the overall evaluation report.

The evaluation report should be easy to understand, and should communicate what was done, how it was accomplished, and why, in a way which is credible to you as a primary evaluation audience. In addition, the evaluation report ought to be timely for your purposes and should focus at least in part, on your identified information needs. The question, "does evaluation make a difference?" is too often answered in the negative because of a failure to gear the evaluation report toward action by decisionmakers.

Checkpoints for Policymakers

Format Considerations:

- Is there an Executive Summary which provides a clear description of the evaluation procedures and results?
- Is the report arranged in such a way (e.g., with appropriate tables and section headings) that you can identify what information is of interest to you and find it easily?

Content Considerations:

- Does the report provide believable evidence that the positive results reported occurred as the result of the special education program or are other explanations, such as normal growth of the students, just as plausible? Are things better than what would be predicted without a special education program?
- Are alternative interpretations of the data presented, and the reasons for their rejection made clear?
- Are the reported results educationally significant, that is, are they "of nontrivial magnitude, in a content area generally accepted as important, which can be achieved at a reasonable cost" (Tallmadge, 1977, 34)?
- Are the data presented consistent, e.g., do the numbers in the text and tables agree? Are the inferences drawn from the data consistent with the evidence, e.g., does the evaluation claim educational significance only when the data support it? Are claims of causality substantiated by the evidence?
- Are there any "unanticipated consequences" of the program? Are they documented and explained in a way that makes their impact clear?

STEP 9: Design a program improvement management plan.

The value of an evaluation, especially one which addresses issues of program quality, rests in the usefulness of its recommendations. Recommendations cover two broad areas: 1) remediation of deficit areas; and 2) improvement of activities determined to be of primary importance. These require both an objective analysis of evaluation data and a subjective analysis of participant concerns. From such analyses, priorities for action can be selected, and a management plan for accomplishing these priorities can be generated.

Checkpoints for Policymakers

- Are the recommendations clearly based on the evaluation information presented? Do you believe these recommendations effectively address the appropriate program strengths and weaknesses? Is it clear who has responsibility for taking action on the recommendations?

- Do the policy implications and recommendations follow directly from the data? Does the report provide the rationale and information you need to act on these recommendations?

Conclusion

The preceding summary of program evaluation procedures illustrates representative components of a technically adequate evaluation. While these components can be easily delineated, in real life many factors limit their implementation. These include diversity in program objectives, heterogeneity of program staff, and swings in resource allocation as well as timelines, and politics.

Policymakers need to consider both the standards for a quality evaluation, and the constraints involved in operationalizing such standards. The key is to be involved and informed at all steps during the evaluation process!

****

REFERENCES

Alkin, M.C. Evaluation theory development. Evaluation Comment, 1969, 2, 2-7.

California State Department of Education. Guide to special education program evaluation. Sacramento: California Evaluation Improvement Program of the Office of Program Evaluation and Research, 1982.

Campbell, D.T. & Stanley, J.C. Experimental and quasi-experimental designs for research on teaching. In N.L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

Maryland State Department of Education. Program evaluation resource manual. Baltimore: Division of Special Education, 1980.

Massachusetts Department of Education. Special education program evaluation: A management tool. Quincy: Division of Special Education, 1981.

North Carolina Department of Public Instruction. Special education program quality evaluation. Raleigh: Division of Exceptional Children, 1983.

Patton, M.Q. Utilization-focused evaluation. Beverley Hills, Calif.: Sage, 1978.

Provus, M.N. Discrepancy evaluation. Berkeley: McCutcheon, 1971.

Stufflebeam, D.L., Foley, W.J., Gephart, W.J., Guba, E.G., Hammond, R.L., Merriman, H.O., & Provus, M. Educational evaluation and decision-making in education. Ithica, Ill.: Peacock, 1971.

Weiss, C.H. Evaluation research: Methods for assessing program effectiveness. Englewood Cliffs, N.J.: Prentice Hall, 1972.

SAMPLE STATE EVALUATION GUIDES

Massachusetts Department of Education. Special education program evaluation: A management tool. Quincy: Division of Special Education, 1981. Publication #12356. Limited copies available free of charge from Donna Toto, 1385 Hancock Street, Quincy, MA 02169; (617) 770-7490.

North Carolina Department of Public Instruction. Special education program quality evaluation manual. Raleigh: Division of Exceptional Children, 1983. Limited copies available free of charge from Valencia Woodward, Division of Exceptional Children, State Department of Public Instruction, Education Building, Raleigh, NC 27611; (919) 733-6081.

****

PROGRAM EVALUATION CHECKLIST

Step 1: Read the complete evaluation report.

- Is it clear exactly what program or program components the evaluation report is assessing?
- Can you describe the types of information contained in the report and in the appendices?
- Can you readily locate specific sections of the report?

Step 2: Review the evaluation model.

- Can you identify the particular evaluation model or approach which is used as the basis for the evaluation?
- Is a rationale included for its selection which clearly relates the assumptions and procedures of the model to the assumptions and procedures of the special education program?
- Is it clear how the evaluation model guides the program evaluation procedures?

Step 3: Review the evaluation's goals and objectives.

- Do the evaluation goals, objectives and questions relate to the evaluation model selected?
- Are the goals and objectives of the evaluation clearly linked to the goals and objectives of the special education program?
- Are the evaluation objectives stated in such a way that you can tell if they have been accomplished?
- Are the goals and objectives which you feel are important to investigate included? Do you have the data you want and need?

Step 4: Review the evaluation criteria.

- Is the type of standard to be used as the criterion for success clear for each objective?
- Is the way in which the standard was selected clear?
- Do you agree with the standard selected, i.e., do you feel the results as measured by this standard are believable and valued?

Step 5: Review the population and sample selection.

- Is the population and/or sample for the evaluation clearly specified?
- If a sample is used, is the sampling procedure clearly delineated and a rationale provided?
- Are all groups which are affected by the program and which you believe are important included in the population and/or sample?

Step 6: Review the evaluation design.

- Is the selection of an evaluation design clear and justified by the aims of the program, types of data needed and the timeframe and data collection burden involved?
- Do you believe that the types of information provided (e.g., statistical, descriptive) are valid and reliable data upon which you would be willing to take action?

Step 7: Review the data collection and analyses processes.

- Does the evaluation address the extent to which the instruments measure what they are supposed to measure (i.e., are they valid)?
- Does the instrument design give you confidence that if the instruments were administered more than once to similar groups (or to the same group) they would yield consistent results (i.e., are they reliable)?
- Do the ways in which the data are analyzed and presented reflect the program objectives and make sense to you? Are the types of data provided (e.g., percent correct, percentile rank, grade equivalents) of use to you?

Step 8: Review the overall evaluation report.

Format Considerations

- Is there an Executive Summary which provides a clear description of the evaluation procedures and results?
- Is the report arranged in such a way (e.g., with appropriate tables and section headings) that you can identify what information is of interest to you and find it easily?

Content Considerations

- Does the report provide believable evidence that the positive results reported occurred as the result of the special education program or are other explanations, such as normal growth of the students, just as plausible? Are things better than what would be predicted without a special education program?
- Are alternative interpretations of the data presented, and the reasons for their rejection made clear?
- Are the reported results educationally significant, i.e., are they "of nontrivial magnitude, in a content area generally accepted as important, which can be achieved at a reasonable cost"?
- Are the data presented consistent, e.g., do the numbers in the text and tables agree? Are the inferences drawn from the data consistent with the evidence, e.g., does the evaluation claim educational significance only when the data support it? Are claims of causality substantiated by the evidence?
- Are there any "unanticipated consequences" of the program? Are they documented and explained in a way that makes their impact clear?

Step 9: Design a program improvement management plan.

- Are the recommendations clearly based on the evaluation information presented? Do you believe these recommendations effectively address the appropriate program strengths and weakness? Is it clear who has responsibility for taking action on the recommendations?
- Do the policy implications and recommendations follow directly from the data? Does the report provide the rationale and information you need to act on these recommendations?

A POLICYMAKER'S EVALUATION GLOSSARY

Evaluation, like many other fields, has a language which describes its processes and products in a specialized manner. The following glossary is provided to clarify some of the basic evaluation research terminology. (Katzer et.al., 1982).

Action research: Research in which the decisionmaker using the results of the research takes part in the research.

Applied research: Research carried out to discover new knowledge that has immediate applicability.

Average: A measure of central tendency (the middle) of a set of scores. Usually, it is synonymous with the mean.

Case study: An analysis and explication of a single situation or single case (unit of analysis; subject), often using qualitative methods.

Before and after design: An experimental design in which subjects are measured on a given variable before and after the independent variable (or treatment) is administered.

Causality: The concept that some phenomena cause other phenomena.

Control group: In a study made up of several different treatment groups, the control group is the one against which the others are compared. Often, the control group does not receive any treatment. (Also known as the comparison group.)

Correlation: In informal usage, correlation is often used as a synonym for "association," although there is a technical distinction between these terms. Correlation is also used to stand for a particular measure of association, Pearson's r.

Demographic variable: An attribute variable of a person; usually a variable that classifies a person into social groupings. Common demographic variables are age, sex, social class, ethnicity and marital status.

Dependent variable: The variable that is thought to be affected by an independent variable (Also known as the criterion, outcome, predicted or response variable.)

Descriptive statistics: Statistics that summarize a set of data; inferences from the data to a larger population are not made.

Empirical: Subject to critical evaluation through observation or experiment rather than through speculation or theorizing.

Experimental group: A treatment group that receives a level of the independent variable (or treatment) which is of substantive interest, as contrasted with the control group.

External validity: The generality of research results.

Experimental research: Research in which the subjects have been subjected to a treatment and the purpose of the study is to assess the effect of that treatment or event.

Field experiment study: Non-experimental research conducted in a non-laboratory setting, usually the natural environment of the phenomenon under study.

Formative evaluation: Evaluation which is conducted during a program or process, usually for the purpose of pinpointing progress/problems and altering the process, if necessary, while it goes on.

Independence: No association or relationship.

Independent variable: The variable thought to produce a result on the dependent variable.

Internal validity: A consideration of whether the independent variable produced variation in the dependent variable. Thus, internal validity is a type of factual accuracy.

Longitudinal study: A study conducted over a period of time, although sometimes only at a few discreet points in time during that period.

Model: A simplified description of a phenomenon, often in words, pictures or symbols.

Naturalistic observation: A method of gathering data that involves making a detailed record of events as they occur in their natural setting while having as little effect as possible on those events.

Non-experimental design: A research design in which it is impossible to control all biases and protect internal validity.

Paradigm: A complex theoretical model that is used to explain phenomenon.

Policy research: Research conducted to help formulate or evaluate policy.

Population: All possible observations or units which could be used in a study.

Post-test: Measurement of experimental subjects on a given variable after a treatment is given.

Pre-test: Measurement of subjects on a given variable before the independent variable/treatment is administered.

Qualitative research: Research methods that attempt to describe and understand people or social entities from their own point of view.

Quantitative research: Research methods that attempt to categorize and summarize observations numerically.

Quasi-experimental research: An experiment in which some, but not all, sources of potential bias are under the control of the experimenter; at least random assignment of subjects is usually missing.

Reliability: Extent to which a measuring instrument would give the same value if used over and over providing the attribute measured did not change.

Research design: The strategy or plan of an experiment, often focusing on the control of possible biases and the scheduling of treatments and measurements.

Sample: The subset of the population actually used in the research.

Statistical significance: The outcome of significance tests in which the results are shown to have a low probability of being due to chance alone, thereby eliminating chance as a viable cause of the results.

Summative evaluation: Evaluation which is concerned with determining overall effectiveness of a completed program or process, usually for the purpose of deciding whether or not to repeat it, or what changes, if any, need to be made before repeating.

Validity: How well an instrument measures the phenomenon under investigation.

****

REFERENCE

Katzer, J., Cook, K., & Crouch, W. Evaluating information: A guide for users of social science research (2nd ed.). Massachusetts: Addison-Wesley, 1982.

****

SEA EFFORTS TO EVALUATE THE QUALITY OF SPECIAL EDUCATION PROGRAMS

Despite the formidable challenges involved in assessing program quality, a number of SEAs have committed themselves to "program quality evaluation," efforts which include a wide variety of methods, and which represent a multiplicity of views about the purpose of quality monitoring. Three such efforts are summarized in the following section. For a more detailed analysis of these and other state evaluation efforts, see Farrow (1983), and NASDE (1984).

Nebraska Department of Education's System of Program Effectiveness Evaluation

The Nebraska Department of Education (NDE), partly in response to state statute (Nebraska Statute No. 43650), is developing a system of program quality evaluation (PQE) which is designed to meet both the SEAs' and LEAs' needs for evaluative information. In 1981, a 20-member Program Effectiveness Development Committee developed a set of six standards or what they termed "service goals" based on their perceptions of the characteristics of an effective special education program. The six major topic areas are:

1. administration, relating to the internal operation of the district;
2. service delivery, relating to the identification of students and the provision of services;
3. curriculum/materials and facilities;
4. communication with consumers;
5. fiscal aspects; and
6. system aspects, relating to the relationship between the special education programs and other programs or agencies which may be affected by handicapped students.

The rogram Quality Evaluation model was field tested in 20 districts during the 1983-84 school year. Eventually, NDE plans to implement the system statewide.

For more information, contact: Mr. Don Anderson, Compliance Director, Special Education, Box 94987, 301 entennial Mall South, Lincoln, NE 68509-4987; (402) 471-2471.

* * * * * *

Missouri Department of Elementary and Secondary Education's System for Program Evaluation

The Missouri Department of Elementary and Secondary Education is developing a Special Education Evaluation (SEE) model to be used on a voluntary basis by LEAs as a way to document the effectiveness of their programs. The SEE model identifies six major components that address the quality of local special education programs:

1. identification of handicapped children;

2. evaluation/diagnosis or reevaluation;

3. development of individualized education programs (IEPs);

4. placement;

5. implementation of IEPs; and

6. annual program review.

For each component, the SEE model looks at resources, services, outcomes and feedback in order to view the program as a whole. The SEE was field tested during the 1983-84 academic year. The system will be implemented separately from the state's compliance monitoring system.

For more information, contact: Mr. Ted Nickell, State of Missouri, Department of Elementary and Secondary Education, P.O. Box 480, Jefferson City, MO 65102; (314) 751-2965.

* * * * * *

North Carolina Department of Public Instruction's Program Quality Evaluation

North Carolina's Program Quality Evaluation (PQE) system is designed to provide LEAs with a mechanism for determining program quality. PQE includes goals in three main areas:

1. determining learner gains/outcomes;

2. locating and evaluating learners; and

3. placing learners appropriately.

For each of these areas, PQE establishes program objectives and specific evaluation questions. LEAs can set the numerical standard for satisfactory performance at different levels, reflecting varying expectations for local programs.

The PQE was field tested in 16 LEAs during the 1983-84 academic year. LEAs were competitively selected through an incentive grant program which provided up to $10,000 to assist with implementation of the evaluation. Eventually, North Carolina plans to combine these evaluation activities with the on-going compliance monitoring procedures.

For more information contact: The Division for Exceptional Children, North Carolina Department of Public Instruction, Education Building, Raleigh, NC 27601; (919) 733-3921.

The chart on the following page highlights some major characteristics of the three state examples described in this section. Perhaps the most striking observation that can be made about these and other efforts to develop and implement program quality evaluation is that "the methods used by SEAs to measure quality are varied and represent divergent views about the goals of quality monitoring" (Farrow, 1983, 30). In special education, as in many other fields, quality means different things to different people.

****

REFERENCES

Farrow, F. Effective state monitoring policies (Quality monitoring and monitoring of state operated programs). Washington, D.C.: The Center for the Study of Social Policy, 1983.

National Association of State Directors of Special Education (NASDE). State responses to the "Annual Evaluation" requirements for state plans. Washington, D.C.: Author, 1984.

****

CHARACTERISTICS OF STATE PROGRAM QUALITY EVALUATION EFFORTS

	SEA ROLE	LEA ROLE	BASIS FOR EVALUATION	PRODUCTS	OTHER CHARACTERISTICS	CURRENT STATE OF DEVELOPMENT	RELATIONSHIP TO COMPLIANCE MONITORING
Missouri	Coordinated development of materials. Convened LEA representatives to review materials.	Jointly developed evaluation materials with SEA. Will self-administer evaluation methodology.	A systems analysis of educational programs, with questions tions that address program resources and policies.	Information for local school districts.	Discretion given to LEAs on which program areas to assess.	Field tested in 30 school districts in school year 1983-84.	None.
Nebraska	Coordinated development of materials and evaluation standards. Convened advisory committee. Will administer evaluation.	Reviewed and revised a draft of the proposed standards.	Standards and service goals which define the characteristics of a quality program.	Information for local school districts.	Use of evaluation is voluntary for LEAs at this point in time.	Field tested in four districts in school year 1983-84.	SEA wants to coordinate but not combine quality evaluation with compliance monitoring.
North Carolina	Developed materials. Will administer evaluation. Will aggregate data statewide.	Reviewed materials. Have option to compete for incentive grants.	Program goals, specific objectives, and evaluation questions which focus on program operations. Numerical standards for program accomplishments.	Information for LEAs. Information for SEA planning and budgeting system.	SEA will award incentive grants to participating LEAs.	Field tested in 1982-83. Implementation in 16 districts in 1983-84.	LEAs pilot-testing the system are exempt from compliance monitoring procedures. Eventual integration of quality evaluation and compliance monitoring.

Adapted from: Farrow, F. Effective state monitoring policies (Quality monitoring and monitoring of state operated programs). Washington, D.C.: The Center for the Study of Social Policy, 1983.