DOCUMENT RESUME

ED 131 861 95 IR 004 245

AUTHOR Webster, William J.
TITLE The Evaluation of Instructional Materials.
INSTITUTION Stanford Univ., Calif. ERIC Clearinghouse on Information Resources.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Oct 76
CONTRACT NIE-C-74-0027
NOTE 30p.; Figures 2 and 3 may be marginally legible due to print quality of original

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Annotated Bibliographies; Conceptual Schemes; *Evaluation Methods; Formative Evaluation; *Instructional Materials; Instructional Media; Material Development; *Models; *State of the Art Reviews; Summative Evaluation

ABSTRACT

A review of four models of instructional materials evaluation is presented to synthesize the state of the art of such evaluation. A model for evaluation is then developed which includes the following elements: (1) context evaluation and needs assessment, (2) input evaluation, (3) determination of resources, (4) program planning, (5) program implementation, and (6) applied research. The use of the model to evaluate the Targeted Achievement in Reading Program (TARP) in the Dallas Independent School District is discussed to show its application in a large urban school district. An extensive annotated bibliography on evaluation of instructional materials is included. (EMH)

The Evaluation of Instructional Materials

William J. Webster

October 1976

Published by the ERIC Clearinghouse on Information Resources,
Stanford Center for Research and Development in Teaching,
School of Education, Stanford University, Stanford, California 94305

INTRODUCTORY COMMENTS

In this paper, William J. Webster provides educators with a review of current theory on instructional materials evaluation, along with a brief description of how a practical system of materials evaluation has been applied in a large urban school district. As a result, local educational decision-makers who are interested in improving their knowledge of the still-infant discipline of instructional materials evaluation are given an opportunity to acquaint themselves with what amounts to a theoretical/practical description of the advanced state of the art of materials evaluation.

The designation "advanced" is certainly appropriate when one compares the approaches described in this paper with what passes for materials evaluation in most school systems. Indeed, if as many as five percent of this country's 15,000-plus school systems are using anything approaching the systematic, empirically-based discrepancy model described by Dr. Webster, this writer will happily eat every available evaluation report, staples and all. In fact, one of the large discrepancies in American education today is between the enormous effect instructional materials have on classroom instruction and the little time, effort, and money expended by school districts in answering these questions: Which materials shall we use? How shall teachers use them? For which students are specific materials best suited?

There is ample evidence of this discrepancy. A recent nationwide study by EPIE, the Educational Products Information Exchange (National Survey and Assessment of Instructional Materials, 1974-75), summarizes reports from some 13,000 teachers within 9,000 school buildings indicating that 90-95 percent of their classroom time involves the use of instructional materials of one sort or another. Yet most of these teachers spend little or no time (one to three hours per year) selecting the materials they are using. At the school system level, there seldom is any systematic information being gathered that could be used for the purpose of materials evaluation if, indeed, there were the commitment to engage in data-based empirical evaluations.

What usually passes for evaluation in schools today is some form of a "materials selection checklist" or a set of "materials selection guidelines"--whose validity and reliability are invariably questionable. The result is that most school systems are basing what amounts to major instructional decisions on simple descriptive information that has been more or less idiosyncratically organized from school system to school system. Even in states that attempt to evaluate materials there is little recourse to empirical evaluation either before or following statewide adoption of materials. Here too the reliance on the ubiquitous checklist is heavy. Dr. Webster's paper clearly points to the shortcomings of this reliance.

The question implicitly raised by this paper is: Is it realistic to assume, or even to hope, that most school districts will adopt in the near future a materials evaluation system of the sort described by Dr. Webster? It would seem that before this can happen there will need to be massive consciousness-raising about the role and the importance of instructional

materials on the part of school consumers. The objective of this consciousness-raising would not be to improve materials evaluation, per se, but to improve the all important three-way "fit" across materials, learners, and teachers.

Dr. Webster's report in his practical example that "it was discovered that certain types of students do better in some programs when taught by specific types of teachers" is tantalizing. After a decade of research on the "fit" question in instructional materials evaluation, we at EPIE know that we have only begun to scratch the surface. Currently, EPIE's research on materials-to-learner-to-teacher fit is being expanded to include data-gathering on instructional materials use and performance across a national network of school systems. At present 22 systems are involved, ranging in size from 2,900 to 250,000 students, with the broadest sort of socioeconomic and ethnic spread. This growing network is a permanent part of EPIE's longitudinal National Survey and Assessment of Instructional Materials which completed its first biennial cycle in June 1976. Using a model similar to that described by Dr. Webster, EPIE hopes that this network of cooperating school districts will contribute to the much needed consciousness-raising on this matter that must occur if instruction is to be consistently and rationally improved in American schools.

P. Kenneth Komoski
Educational Products Information
Exchange Institute
475 Riverside Drive
New York, New York 10027

TABLE OF CONTENTS

PURPOSE OF THIS PAPER

The evaluation of instructional materials is an extremely important yet often overlooked component of the total instructional process. Often the term evaluation is operationally defined as "checklist" by many curriculum specialists who perform a weak form of input evaluation on instructional materials, using some variation of a survey form or checklist. In presenting alternatives to inadequate evaluation, this paper attempts to accomplish three objectives:

1. To synthesize, through a selective review of the literature, a brief description of the state of the art of evaluation;

2. To present a working model demonstrating the functions of various forms of evaluation in assessing the relative merits of instructional materials;

3. To provide an annotated bibliography of sources for readers seeking further information on evaluation.

A BRIEF STATE OF THE ART

Four Evaluation Frameworks

The CIPP Model. Probably the most comprehensive of existing evaluation models is the CIPP Model developed by Stufflebeam, et al. (1971). Evaluation is defined as the process of delineating, obtaining, and providing useful information for judging decision alternatives. The model identifies four major types of evaluation: Context evaluation to feed planning decisions, input evaluation to feed programming decisions, process evaluation to feed implementing decisions, and product evaluation to feed recycling decisions.

Briefly, context evaluation provides a rationale for determining educational objectives by defining relevant environment, describing desired and actual conditions of the environment, identifying unmet needs, and diagnosing problems that prevent needs from being met. Input evaluation assesses relevant capabilities of responsible agencies, identifies strategies for achieving the objectives determined through context evaluation, and suggests designs for implementing selected strategies. Once a strategy has been selected, process evaluation provides periodic feedback to help predict or detect faults in procedural design or implementation so that interim adjustments may be made. Finally, product evaluation provides interim and final assessment of the effects of educational programs. That is, product evaluation assesses the effects of the strategy selected through input evaluation to meet the need identified by context evaluation. Such assessment is completed in light of process evaluation data.

Scriven. Scriven (1967) has conceptualized an extremely straightforward and widely accepted evaluation framework. Not nearly as comprehensive as the CIPP Model, it is largely concerned with the process-product portion of Stufflebeam's structure. According to Scriven, the major goal of evaluation is to credibly judge the merits of educational programs. To accomplish the goal, he introduces the concepts of formative and summative evaluation.

The focus of formative evaluation is upon program improvement. Thus, formative evaluation attempts to provide feedback to program personnel in order to upgrade or improve an educational program while it is in the developmental stage. In the CIPP vernacular, interim product and process data provide formative evaluation information to program personnel.

The focus of summative evaluation is upon the determination of the ultimate worth of a program or project. This type of evaluation should be implemented when a program has reached some stability. Summative data feed recycling decisions; as a result of summative evaluation information, a program may be terminated, restructured, continued, or expanded. In the CIPP vernacular, final product evaluation information, interpreted in consideration of context, input, and process data, is used to draw summative conclusions about the merits of an educational program and to feed recycling decisions.

Stake. Stake (1967) suggests that evaluation ought to be concerned with three classes of conditions: antecedents, transactions, and outcomes. Antecedents are those conditions that exist prior to program implementation, i.e., the educational context. Transactions are interactions between students, teachers, and materials. Outcomes are the intended products of transactions.

Three classes of activities are suggested by Stake. The first provides assistance to program staff by generating a clear statement of the program or project rationale. The second activity generates descriptive data, including statements about intended and actual antecedents, transactions, and outcomes. Thus, congruence between planned and observed antecedents, transactions, and outcomes can be checked.

The third class of activities generates judgments about the worth of educational programs. Stake suggests that such judgments be made by a variety of individuals on the basis of both absolute and relative criteria. In other words, programs should be assessed in terms of: 1) The degree to which they attain absolute and sometimes arbitrary goals, and 2) The degree to which they attain those goals relative to other programs with similar goals or objectives.

Provus. Provus (1971) suggests that all projects move through design, installation, process, and product stages. During each stage the evaluator must delineate, in conjunction with project staff, a set of standards that can be used as a basis for comparison with program performance. It is the evaluator's function to make comparisons with program performance, to identify discrepancies at each stage, and to report those discrepancies to project management who have the option of terminating the program, proceeding to the next stage, or modifying the program. The product of the design stage is a set of standards used to judge the effects of program efforts in each of the three succeeding stages. At every stage the object of the evaluation is to provide useful data for decisions about program improvement or recycling.

General Discussion

These four generic evaluation frameworks are supplemented by many articles dealing with methodology, purposes, instrumentation, strategies, and variables to be considered in evaluation. Among these are the Differential Evaluation

Model (Tripodi, Fellin, and Epstein, 1971), the Decision Oriented Classification Model (Alkin and Wooley, 1969), the Apex Evaluation Model (Morgan, 1970), the Synergistic Evaluation Model (Hunter and Schooley, 1973), the Ontological Evaluation Model (Peper, 1973), and the Ott Model (Ott, Fletcher, and Turner, 1966). In addition, the IPI Formative Evaluation Model (Lindvall and Cox, 1970) and the New Start Evaluation System (Lamrock, Smith, and Warren, 1971) were developed specifically for the evaluation of individualized curriculum packages while the Trade-off and Comparative Cost Approach (Glass, 1972) and the Weighted Criteria Approach (Crane and Abt, 1969) were designed to aid in the evaluation of educational materials.

The research literature boasts literally thousands of studies dealing with research on most areas of the cognitive, affective, and psychomotor domains. Of particular relevance to the evaluation of instructional materials are the summaries by Bracht (1969) and Cronbach and Snow (1969) and the book edited by Wittrock and Wiley (1970).

The point that all the aforementioned literature attempts to make is that mere checklists provide insufficient information about instructional materials. The evaluation of instructional materials requires a system that takes the user condition and needs into consideration, examines process and product, and renders a decision about the effectiveness of a given piece or set of materials for a specified target audience under a specified set of conditions. The work of Stufflebeam, Scriven, Bracht, Cronbach and Snow, Solomon (1972) and Webster and Mendro (1974, 1975), is used in this paper to develop a straightforward operational model for the evaluation of instructional materials.

A MODEL FOR INSTRUCTIONAL MATERIALS EVALUATION

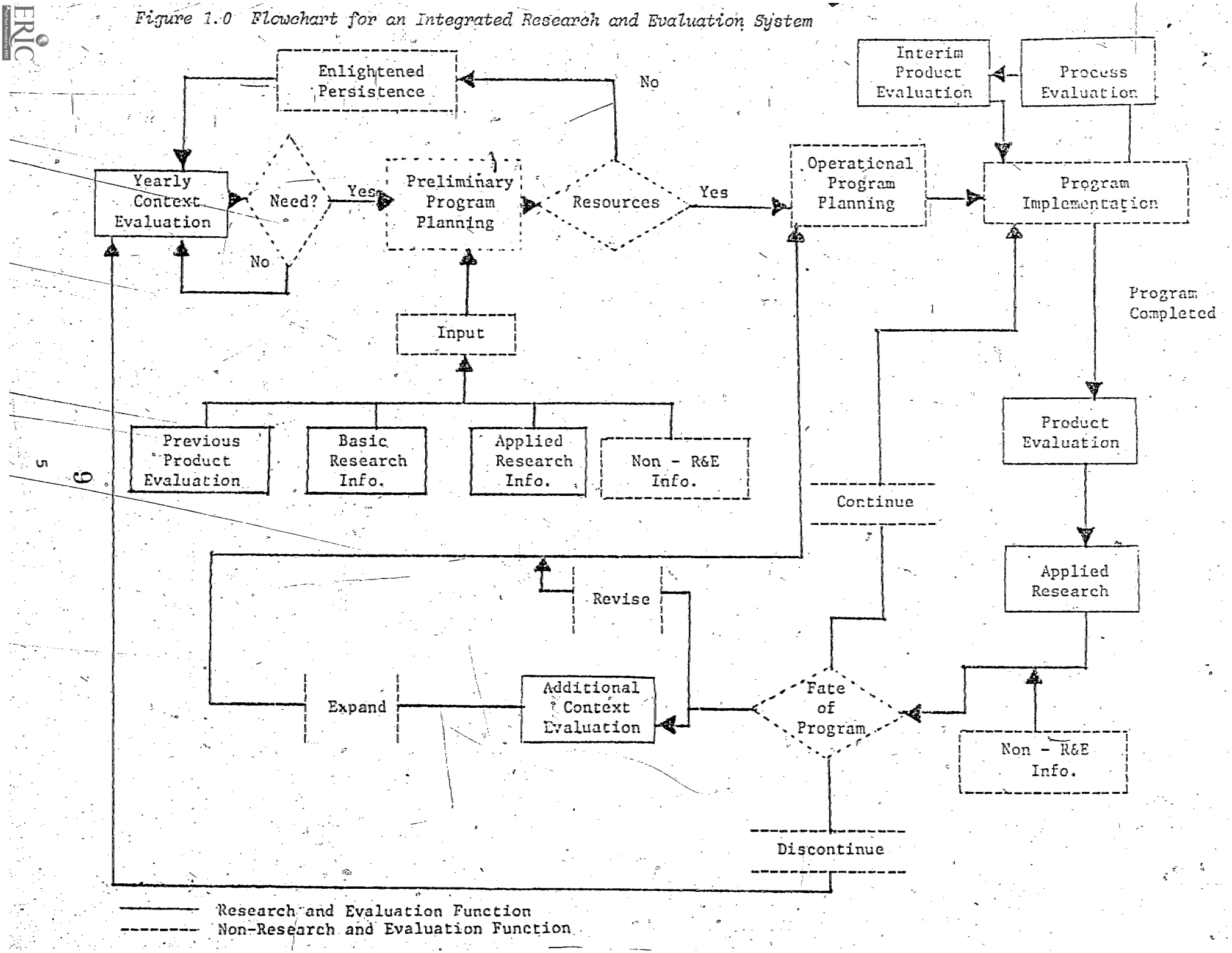
Figure 1.0 presents a flowchart for an integrated model for evaluating instructional materials. The events depicted in Figure 1.0 may take place within the period of one day, ten years, or a lifetime, depending on the scope of the materials to be evaluated.

Figure 1.0

Context Evaluation and Needs Assessment

A prerequisite to improvement must be a knowledge of existing performance levels. Thus, before beginning the search for instructional materials, you must establish a need for such materials. Stufflebeam et al. (1971) define context evaluation as the provision of baseline information that delineates the environment of interest, describes desired and actual conditions pertaining to the environment, identifies unmet needs and unused opportunities, and diagnoses the problems that prevent needs from being met and opportunities from being used. Thus an adequate context evaluation would point to problems that need resolution, such as low reading or math achievement, high student dropout, high teacher turnover, low student or staff morale, and so on. An example of an operating context evaluation system was presented by Webster and Schuhmacher (1973).

Figure 1.0 Flowchart for an Integrated Research and Evaluation System



Once the context evaluation system has identified needs, those needs must be prioritized, and management must focus upon reducing the discrepancy between desired and existing conditions by establishing goals for highest priority needs. It should be obvious that the felt need to select new materials properly begins with dissatisfaction with those in use. In order for you to be legitimately dissatisfied, however, you must at least suspect that current materials are not meeting needs. The fact that there is a discrepancy between desired outcomes and needs suggests the existence of objectives or goals. Thus a loop is formed--context evaluation, needs assessment, goal setting. Or is it the other way around? Denton (1976) presents a concise discussion of the complex interrelatedness of these three mutually dependent concepts. Some methodology for needs assessment has been discussed by the Center for the Study of Evaluation (1973), Feldmesser (1973), the Florida Educational Research and Development Council (1968), and Kaufman (1972). Bolin (1973), Cook and Walbesser (1973), Uhl (1971), and Weatherman and Swenson (1974) have presented models for goal setting.

Input Evaluation

Input evaluation helps determine the best resources to accomplish program goals. It logically follows the context evaluation, needs assessment, goal setting stage. The program planner can use four major sources of information in selecting instructional materials to meet specific needs: (1) Previous summative product evaluation information; (2) basic research information; (3) applied research information; and (4) non-empirical information.

(1) Summative product evaluation information concerns the extent to which specific project or program goals are achieved. When product evaluation information is available on a given program with goals similar to those identified in response to context evaluation information, that information helps decision makers determine the probability that the program would reduce the identified discrepancy between desired and existing conditions.

(2) Basic research information pertains to fundamental relationships that affect student learning. Before making a decision to implement a given program, decision makers should know if that program is or is not consistent with the principles established by basic research in learning and development.

(3) Applied research information concerns the interaction between student characteristics, teacher characteristics, and instructional systems. Applied research differs from basic research in that the information is more closely related to specific decisions in an applied educational setting. Decision makers need information about the types of students (e.g., high anxiety versus low anxiety) who function best in given instructional systems as implemented by teachers with different types of characteristics or traits.

(4) Nonempirical information also enters into any materials selection decision. The influence of materials distributors, costs, political feasibility of material adoption in given communities, capabilities and attitudes of staff members, and existing facilities are among the many nonempirical considerations that influence such decisions.

Many checklists or evaluation forms have been developed for aiding educators in the materials selection process. These forms are useful for organizing information about instructional materials, but in themselves do not provide for sufficient materials evaluation. In many cases such forms are used to summarize information of the type previously referenced. A particularly organized approach was presented by Armstrong (1973); other instruments were developed by Carpenter and Froke (1968), Caldwell (1968), Kovac (1976), Grobe (1976), Miller (1969), and the National Institute of Education (Product Rating Form, 1974). In addition, the Educational Products Information Exchange has published a number of guidelines for product selection (Komoski, 1967, 1975; Educational Products Information Exchange, 1971).

Determination of Resources

After the collection of relevant input information feeding the preliminary program planning stage, it must be determined whether sufficient resources are available to make the desired instructional changes. Quite often, adequate resources are not available and some compromise is necessary. In many cases, the lack of resources is not limited to the realm of cost and political feasibility but rather stems from an insufficient base of research information. Thus, educators are often in the position of having sufficient material resources but insufficient information resources.

If sufficient material resources are not available, the system may have to exist for some time in a state of enlightened persistence. Periodic context evaluation will continue to highlight the extent of discrepancy between that which is desired and that which exists.

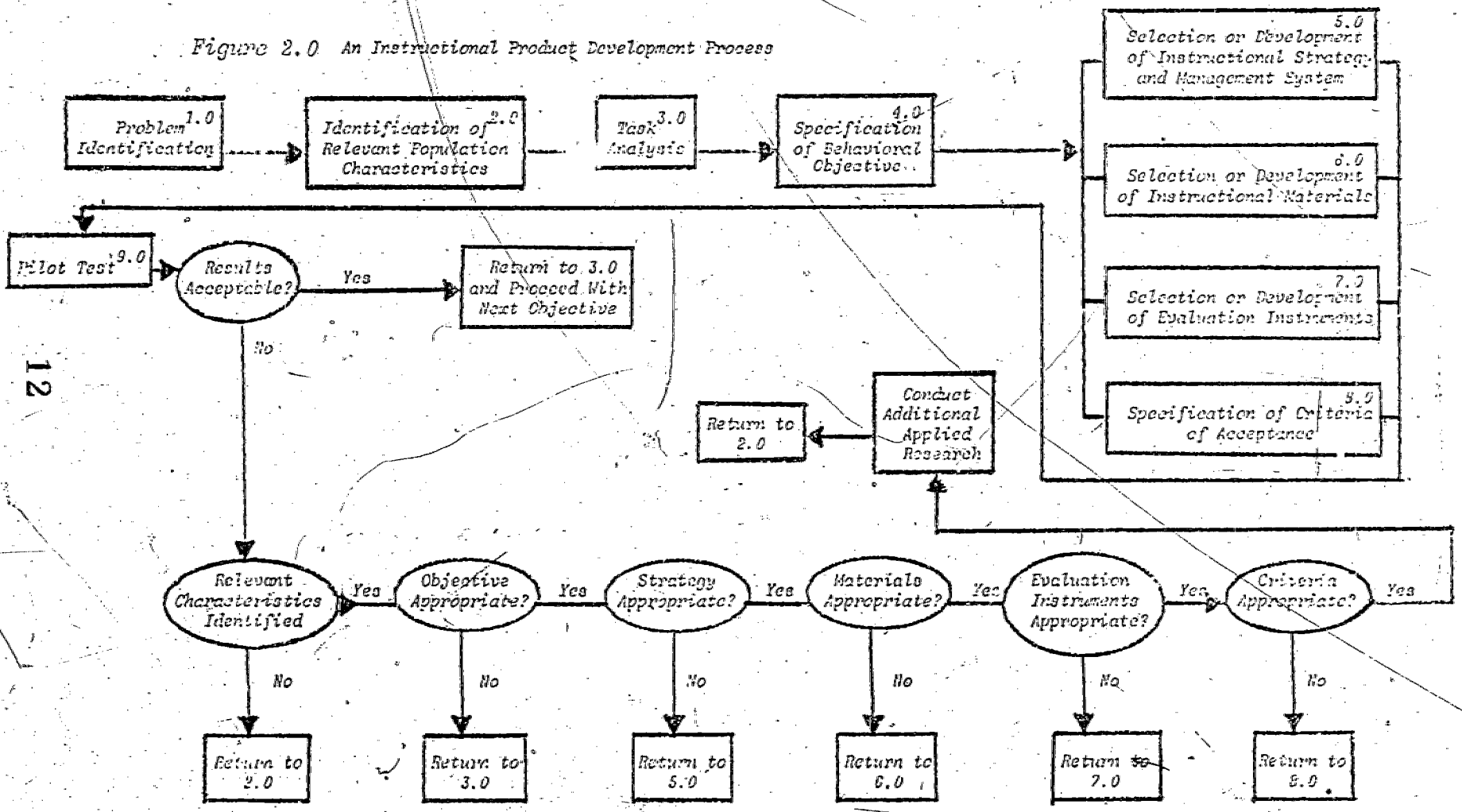
If sufficient information resources are not available, programs often are implemented without sufficient support data and an information base must be built through a series of systematic evaluation and applied research studies. In some cases, development centers are established to cope with the problem of insufficient information resources. These centers are responsible for developing instructional systems to meet the needs outlined by context evaluation. Materials and instructional systems are developed at the local level only if no potentially useful commercial materials are available, since the development of instructional systems is an extremely costly proposition. Figure 2.0 outlines a developmental process that theoretically could be followed in developing instructional systems. An excellent introductory reference on instructional product development was edited by Baker and Schutz (1971).

Figure 2.0

Program Planning

If sufficient resources are available, the extended program planning phase is entered. This phase is implemented using information gleaned from input evaluation. The program planning stage includes the development of detailed program objectives and of a management plan. One product of the extended program planning phase is a detailed program evaluation design specifying the criteria by which the instructional materials will be judged. This design and its implementation should be done by an individual who works closely with program management yet is independent of them to allow for maximum flexibility. Webster (1975) outlined an operational procedure

Figure 2.0 An Instructional Product Development Process



for assuring the independence of project evaluators.

Program Implementation

Once the program implementation phase is entered, the evaluator provides continuous formative evaluation reports on program implementation. These reports fall primarily into two categories, process evaluation and interim product evaluation. Process evaluation has three major objectives: (1) The detection or prediction of defects in procedural design or its implementation during program implementation stages; (2) the provision of information for programmed decisions; and (3) the maintenance of a record of the implementation procedure as it occurs (Stufflebeam, et al., 1971). Thus, process evaluation information keeps program management informed of the extent to which program implementation conforms to specifications and, from an evaluation standpoint, guards against the evaluation of a fictitious event. Methods of accomplishing this have been discussed by Ashburn (1975), Denton (1976), Macy (1975), Niedermeyer (1972), and Rosenshine and Furst (1973).

Interim product evaluation provides periodic feedback to program management on the attainment of specific sub-objectives during the implementation phase. Thus, process and interim product evaluation reports inform program management as to implementation and goal attainment levels while program adjustments are still feasible. Formative evaluation methodology has been discussed by Bloom, Hastings, and Madaus (1971), Rossi (1969), and Scriven (1967).

Product Evaluation

Upon completion of a given cycle of program implementation, a summative product evaluation report is prepared. This report generally addresses three areas of concern: (1) The extent to which program objectives were achieved relative to some specific set of criteria; (2) the extent to which system objectives were achieved relative to alternative instructional strategies; and (3) the cost-effectiveness of the program relative to alternative instructional strategies. It should be obvious that information on these three areas must be interpreted in light of process and interim product evaluation information. Without information about program implementation, product evaluation information is of little use.

Figure 3.0 outlines the necessary steps in product evaluation. The reader should bear in mind that product evaluation depends on strong experimental design and commences at that point that the need has been identified, resources have been allocated, and the program has been tentatively planned. Product evaluation methodology was the primary focus of many of the evaluation models discussed earlier. Useful articles on various aspects of product evaluation include those by Alkin (1970), Baker and Schutz (1972), and Webster (1971). In addition, some good texts on experimental design and applied statistics include those by Campbell and Stanley (1966), Edwards (1972), Finn (1974), Glass and Stanley (1970), and Tatsuoka (1971).

Figure 3.0

Applied Research

Many evaluation systems stop at the provision of product evaluation information, which generally bears upon the performance of different groups of students under

Figure 3.0 Necessary Steps in Product Evaluation

1. Program Objectives Determination

   1.1. Meet with decision makers and program managers to determine the program objectives.
   1.2. Refine objectives through thorough analysis, review of literature, questioning decision makers, analysis of input data, etc.

2. Information Regarding Program Decisions

   2.1. Using the objectives, meet with decision makers, etc. to generate a list of the critical decisions to be made concerning the objectives and the program.
   2.2. Determine the types of information necessary to make the various decisions.
   2.3. Estimate the critical decisions and plan the information sources so critical decisions receive the most information.

3. Definition of Measurable Objectives and Related Decisions

   3.1. Work with project personnel to mold objectives so that they may be measured.
   3.2. Put into operation basis for decision-making to relate to measured achievement of objectives.

4. Plan of Evaluation Dissemination

   4.1. Identify the various audiences of the evaluation report and estimate the level of sophistication of the intended audience.

5. Identification of Measuring Instruments

   5.1. Review objectives and decisions and evaluate existing instruments to determine those which can be employed in the evaluation.
   5.2. Determine areas where no satisfactory instruments are available and develop complete specifications of instruments that are to be constructed.

6. Instrument Development and Testing

   6.1. Develop needed instruments.
   6.2. Test new instruments, if necessary, on a sample of subjects.
   6.3. Refine new instruments on the basis of these tests.
   6.4. Test administration of any unconventional instruments or observation procedures.

7. Information Collection Scheduling

   7.1. Specify sampling procedures to be employed.
   7.2. Determine the schedule of observations and the instruments to be administered at each observation point.
   7.3. Schedule the personnel needed to administer instruments.

8. Organization of Data Analysis

   8.1. Determine various formats of data including card and tape format specifications at various stages of collection and analysis. Specify processing necessary to put data into correct format at each stage of analysis.
   8.2. Plan nonstatistical analysis of data and resources necessary to perform analysis.
   8.3. Plan statistical analysis of data and programs necessary to analyze data.
   8.4. Determine which programs are already written and are ready to use, which programs are written but need modifications to handle data in their intended formats, and which programs need to be written with specifications of these programs.

9. Formal Evaluation/Research Design

   9.1. Prepare design including specification of

      9.1.1. Objectives
      9.1.2. Instrumentation
      9.1.3. Analysis methodology
      9.1.4. Data collection and reporting schedules
      9.1.5. Sampling procedures
      9.1.6. Data analyses schedules
      9.1.7. Final reporting schedules

   9.2. Type, print, and collate design.
   9.3. Disseminate formal design.

10. Computer Program Development

   10.1. Develop necessary programs for analysis.
   10.2. Make necessary modification of existing programs.
   10.3. Run all programs to be used on sample data in the proper medium and format. Construct sample data to simulate problems in actual data (mispunching, missing data, etc.).

11. Process Evaluation

11.1. Collect or supervise and coordinate collection of process evaluation information.
11.2. Prepare process evaluation information for analysis.

12. Product Evaluation

12.1. Collect or supervise and coordinate the collection of product evaluation information.
12.2. Prepare product evaluation information for analysis.

13. Interim Data Analysis

13.1. Organize interim data.
13.2. Perform analysis of interim data.

14. Formative Evaluation Reports

14.1. Prepare formative evaluation reports.
14.2. Type, print, and collate formative evaluation reports.
14.3. Disseminate formative evaluation reports to project management and staff.

15. Summative Data Analysis

15.1. Organize summative data.
15.2. Perform analysis of summative data.

16. Summative Evaluation/Research Reports

16.1. Prepare the various summative evaluation/research reports for each audience including objectives, findings, and recommendations expressed in an appropriate manner for the intended audience. This preparation includes the abstract of the report.
16.2. Have report carefully proofread and corrected.
16.3. Type, print, and collate the summative evaluation reports.
16.4. Disseminate the summative evaluation/research reports to project personnel, district management, and the Board of Education.

17. Interpretation of Reports

17.1. Meet with project personnel to interpret reports.
17.2. Meet with district management and the Board of Education to aid in report interpretation.

18. Further Report Dissemination

18.1. Disseminate summative evaluation/research reports to all affected district administrators and to interested professional staff.
18.2. Prepare and disseminate a book of evaluation and research abstracts to all professional staff.

19. Report Feedback

19.1. Meet with decision makers to obtain feedback regarding the report with the purpose of improving reporting activities.

varying treatment configurations. Unfortunately, as a result, most product evaluation reports generally have focused on the search for single best treatments for all learners, i.e., main effects. In order to provide needed information for educational decision makers, applied research studies involving the systematic investigation of aptitude-treatment and trait-trait interactions must be undertaken. Such studies would be expected to provide important information on replicable relationships among student, teacher, and program characteristics (Bracht, 1969; Cronbach and Snow, 1969; Solomon, 1972; Webster and Mendro, 1974).

The basic assumption of aptitude-treatment interaction research is that learners possess characteristics or traits that interact positively or negatively with specific treatments or program characteristics. Messick (1970) outlined some cognitive-style dimensions that represent a person's typical modes of perceiving, remembering, thinking, and problem solving, and, as such, would provide excellent variables for aptitude-treatment interaction studies. In addition, many affective variables warrant investigation when attempting to validate replicable teacher-student (trait-trait) interactions. The basic assumption of trait-trait interaction research is that teachers possess characteristics or traits that, independent of program, interact positively or negatively with specific characteristics or traits possessed by learners. Such interactions may involve variables such as arithmetic reasoning, language usage, vocabulary, abstract reasoning, mechanical reasoning, creativity, anxiety, affiliation, aggressiveness, compulsiveness, dogmatism, paranoia, and status variables such as sex, age, or ethnicity.

Nonempirical Information

Once context, input, process, and product evaluation information, as well as applied research data, are available, nonempirical information once more is brought to bear upon the decision-making process. It would be naive to expect educational decisions to be made purely on the basis of empirical data. Once again, information such as the absolute program costs, capabilities of program staff members, political feasibility of program implementation, and existing facilities and resources must be considered by decision makers.

Determination of Program Fate

In determining the fate of a given program, four primary choices are available to decision makers: to continue, discontinue, expand, or revise the program. First, they can choose to continue the program in its current setting. If this alternative is chosen, the summative product evaluation report and the applied research data become the context evaluation information for the next implementation phase, and program implementation commences. This alternative generally occurs when decisions are to be made on the basis of longitudinal studies, i.e., where it is expected that results will not be in evidence after a relatively short implementation period.

A second alternative is to discontinue the program. This is usually done after product evaluation studies demonstrate the failure of the program to meet its objectives or in those cases where the program is simply not cost-effective. (Failure to meet objectives is often a necessary, but not sufficient, condition for program discontinuation.) Once a program is discontinued, the system returns to the context evaluation phase and once again applies the needs assessment and orientation phases.

Third, if the product evaluation and applied research information are favorable and if it is practically and politically feasible, the program may be expanded to serve additional students. Prior to expanding the program, additional context evaluation information must be examined to determine if similar needs exist in other settings. If such needs are demonstrated, then the program planning stage is entered to extend the program implementation. Other settings eligible for program expansion may include entire schools or specific subpopulations (e.g., highly motivated students) as indicated by applied research data. If such needs do not exist, the program is continued with the original target population or a reduced target population based on the results of the applied research studies. The extent of continued evaluation under either the expansion or continuation alternative is determined by decision makers with advice from evaluation personnel.

A fourth alternative involves program revision. Much program revision should be accomplished on the basis of process and interim product evaluation reports. Often, however, summative product evaluation and applied research reports reveal weaknesses in portions of programs that would otherwise appear to be functional. In this instance, the summative product evaluation and applied research reports become the context evaluation information for the next program-planning cycle.

A PRACTICAL EXAMPLE

This evaluation model was used in the Dallas Independent School District's Targeted Achievement in Reading Program (TARP). Context evaluation data, compiled in 1971, demonstrated a clear need for reading intervention in a number of the District's schools. Among the data identifying such a need were:

1. Extremely low scores in reading and language on standardized tests, usually at the chance level,
2. Extremely low scores in reading and language on the District's criterion-referenced tests,
3. High student mobility,
4. Low student attendance,
5. High teacher turnover,
6. Low student motivation,
7. Parent apathy.

Thus the need was established for programs that would: (1) Teach basic reading and language skills to a highly mobile population; (2) be sufficiently interesting to increase student attendance and motivation; (3) have a parent-training component to increase parent interest; and (4) have a teacher-development component to increase teacher skills and motivation in dealing with disadvantaged students, thereby reducing teacher turnover.

A search was begun for programs designed to enhance individualized instruction that possessed the four characteristics listed above. Program planners used information that had worked in other areas [previous product evaluation], suggestions that should work [basic research], and, where available, the characteristics of environments in which specific programs had met their objectives [applied research]. Community and staff opinions also were considered [non-research and evaluation information]. Four programs were chosen--three on the basis of all four considerations [previous product evaluation, basic research information, applied research information, and non-research and evaluation information], and one solely on the basis of opinion [non-research and evaluation information], without reference to previous product evaluation information that suggested it would not work in the desired environment. Cost considerations were not foremost in the decision since the District was primarily concerned with finding material resources necessary to meet the demonstrated need.

Once the four programs were chosen, managers were appointed, and operational program planning commenced. Program managers, with technical assistance from evaluation staff, developed specific objectives and management plans to accomplish those objectives. Evaluators, responsible to the Department of Research and Evaluation, and program facilitators, responsible to the program managers, were assigned and began implementing their respective tasks. The program facilitators implemented their programs according to the management plans, and the evaluators reported on how close project implementation was to planned implementation [process evaluation] and on how well the respective programs were doing relative to specific sub-objectives [interim product evaluation]. These reports were produced at regular intervals and were used by project management to make in-course adjustments in the projects. Many of the interim process and product evaluation reports were in the form of memoranda.

Since it was predicted that several years would be required to implement the programs in the schools, the TARP project involved a three-year longitudinal evaluation design. That is, decisions about the fate of the programs would be made on the basis of three years of evaluative data. In-course adjustments would be made on the basis of interim and annual evaluation reports.

The product evaluation reports examined four basic questions about each of the four programs. These questions were:

1. Did each of the programs meet its objectives?

2. How did the programs compare to each other in terms of meeting District and national objectives in reading and language?

3. How did the programs compare to each other in terms of increasing student attendance and motivation, decreasing teacher turnover, and increasing parental interest?

4. How cost-effective were the programs?

In addition, applied research information on the types of teacher-student combinations that produced the best results in each of the programs was examined. Variables such as cognitive ability, anxiety, sex, ethnic background, age, experience,

training, attitude toward instruction, racial attitudes, creativity, social class, mobility, and economic level of both teachers and students were investigated, with some very enlightening results. For instance, it was discovered that certain types of students do better in some programs when taught by specific types of teachers.

Although designed to run for three years, one of the four programs was judged so clearly inferior on the basis of the established criteria that it was eliminated after two years. As it turned out, the program eliminated was the one that was chosen originally without consideration of evaluative data. Thus two years of resources were wasted because of insufficient or inadequate input information.

Three programs remained after three years of implementation, with each program in about ten schools. The TARP program then was to be expanded to 52 additional schools of similar context to the 30 schools where the programs first were tried. The programs were found similarly successful in meeting their objectives and those of the District. They produced similar patterns of performance on standardized tests and similar outcomes on concomitant variables (attendance, motivation, teacher turnover, parental interest). The questions of interest were: (1) Which programs were to be eliminated; (2) which were to be expanded; and (3) which, if any, were to continue operating in the context in which they had originally been implemented.

Since all the programs achieved similar results, none was to be eliminated on the basis of outcomes. The next question was that of input, or cost. All three programs cost about the same to operate, once they were established. However, one program cost about three times as much as the other two to implement. Thus, the cost-effectiveness question yielded additional information for decision-making. It would not be particularly useful to eliminate the expensive program, since once implemented it cost approximately the same as the other two (and applied research information suggested that its more structured approach was more effective with high anxiety students). Yet its additional implementation cost would not warrant its expansion. Therefore, it was decided to keep that program in the schools where it had been originally implemented, and, when possible, to assign high anxiety students to it.

The other two programs achieved about the same results and cost approximately the same. Therefore, those two programs were expanded to the 52 additional schools, with school faculties and communities choosing the program they preferred. Thus, school communities were permitted freedom of choice within a restricted set of two alternatives, either of which had a high probability of success.

Once the programs passed out of the project evaluation phase they were continuously monitored through system-wide context evaluation. Currently, if achievement results or other variables in the participating schools display decrements, the evaluation process is re-entered, and replacement programs or strategies are initiated on a sampling basis.

SUMMARY

This paper has attempted to summarize the state of the art in the evaluation of instructional materials. The position has been that such evaluation should be implemented in a systematic manner and should include context, input, process, and product evaluation with applied and basic research input. The relevant evaluation literature has been summarized, and an operational model for evaluating instructional materials that could be implemented in a day or a lifetime has been presented, along with an example of its usefulness. For further information, the reader is encouraged to pursue readings from the following list of references. Two very good general references on evaluation are Anderson, Ball, and Murphy (1975) and Worthen and Sanders (1973), and some excellent evaluation units are operating in the Cincinnati, Dallas, and Philadelphia public school systems.

REFERENCES

Alkin, M. C. "Evaluating Net Cost-Effectiveness of Instructional Programs." In *The Evaluation of Instruction: Issues and Problems*, edited by M. C. Wittrock and D. E. Wiley, pp. 221-238. New York: Holt, Rinehart and Winston, 1970.

Draws the distinction between cost-benefit and cost-effectiveness analysis and outlines components of a model for conducting cost-effectiveness evaluations in education. The model considers student inputs, educational outputs, financial inputs, external systems, and manipulatable characteristics. Potential applications of the model are noted.

Alkin, M. C., and Wooley, D. C. *A Model for Educational Evaluation.* Paper presented at PLEDGE Conference, San Dimas, California, October 8-11, 1969. 12pp. ED 036 898.

The primary source for the Decision Oriented Classification Model. Defines evaluation as the process of ascertaining decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision-makers in selecting among alternatives.

Anderson, S. B.; Ball, S.; and Murphy, R. T. *Encyclopedia of Educational Evaluation: Concepts and Techniques for Evaluating Education and Training Programs.* San Francisco: Jossey-Bass, 1975.

Provides an excellent overview of the field of educational evaluation in terms that are generally comprehensible to non-specialists.

Armstrong, J. R., ed. *A Sourcebook for the Evaluation of Instructional Materials and Media.* Arlington, Virginia: IMC/RMC Network; Madison: University of Wisconsin, Special Education Instructional Materials Center, August, 1973. 520pp. ED 107 050.

Describes a generalized model for the evaluation of instructional materials and media designed to be usable by the Instructional Materials Centers/Regional Media Centers/Regional Resource Centers Network. The author discusses general procedures for selecting instructional materials at various stages of the evaluation process. Appendixes include a bibliography and a set of forms for evaluating instructional materials and media.

Ashburn, A. G. "Analysis of Activity-Decision Flow in Process Evaluation Design." *Educational Technology*, 15 (June, 1975), 36-39.

Presents a generalized activity-decision flow diagram delineating the types of process evaluation information needs that are likely to arise in the course of project implementation.

Baker, R. L., and Schutz, R. E., eds. *Instructional Product Development.* 7 vols. Southwest Regional Laboratory for Educational Research and Development. New York: Van Nostrand Reinhold, 1971.

An introductory text. Includes instructional sequences dealing with stating educational objectives, constructing objectives of cognitive behavior, developing instructional specifications, educational criterion

measures, rules for the development of instructional products, preparing instructional products, and managing classroom contingencies.

Baker, R. L., and Schutz, R. E., eds. *Instructional Product Research.* 8 vols. Southwest Regional Laboratory for Educational Research and Development. New York: Van Nostrand Reinhold, 1972.
An introductory text. Includes instructional sequences dealing with classifying and interpreting research studies, selecting variables for research, proposal writing, simplified research design, simplified analysis, the use of library computer programs for statistical analysis, and writing research reports.

Bloom, B. S.; Hastings, J. T.; and Madaus, G. F. *Handbook on Formative and Summative Evaluation of Student Learning.* New York: McGraw-Hill, 1971.
Provides an introductory overview of the use of evaluation in education. The final 11 chapters deal with the application of evaluation to 11 different subject areas on levels ranging from preschool education to industrial education.

Bolin, J. G. "Six Criteria for Better Goals." *Improving College and University Teaching,* 21 (August, 1973), 245-247.
Posits six criteria for developing adequate goals. Criteria include: compatibility, attainability, intelligibility, acceptability, measurability, and accountability.

Bracht, G. H. *The Relationship of Treatment Tasks, Personalogical Variables, and Dependent Variables to Aptitude-Treatment Interactions.* Doctoral dissertation, University of Colorado, Boulder, Colorado, 1969. Available from: University Microfilms, 300 N. Zeeb Road, Ann Arbor, Michigan 48106 (Order No. 70-05820; $15.00 papercopy, $7.50 microfiche).
Provides an excellent review of the research on aptitude-treatment interactions.

Caldwell, M. S. *Input Evaluation and Educational Planning.* Columbus: The Ohio State University, Evaluation Center, January, 1968. 23pp. ED 025 043.
Suggests eight criteria that might be used as a screening system for instructional materials. Criteria include: relevance, legality, congruence, legitimacy, compatibility, balance, practicability, and cost-effectiveness.

Campbell, D. T., and Stanley, J. C. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally, 1966.
The classic work on applied experimental design. Applies the principles of experimental design to field settings.

Carpenter, C. R., and Froke, M. *Description of a Practical Procedure for Assessing Instructional Film and Television Programs.* University Park: The Pennsylvania State University, Department of Psychology, August, 1968. 43pp. ED 037 102.
Provides a system of checklists for rating instructional materials. Although designed for use with instructional film and television programs,

it would need only slight modifications to be used with other instructional materials.

Center for the Study of Evaluation. *CSE Elementary School Evaluation Kit: Needs Assessment.* Boston: Allyn and Bacon, 1973.
Introductory level training materials dealing with the goal selection process, test selection, data collection, and the identification of critical need areas.

*Cook, J. M., and Walbesser, H. M. *How to Meet Accountability with Behavioral Objectives and Learning Hierarchies.* Columbia, Maryland: Advance Educational Press, 1973.
Discusses the use of behavioral objectives and learning hierarchies in evaluation.

Crane, P., and Abt, C. C. "A Model for Curriculum Evaluation." *Educational Technology,* 9 (October, 1969), 17-25.
The primary source for the Weighted Criteria Approach to curriculum evaluation. Based on cost-effectiveness techniques applied to curriculum materials in order to analyze components, quality, and cost of the materials being considered.

Cronbach, L. J., and Snow, R. E. *Individual Differences in Learning Ability as a Function of Instructional Variables. Final Report.* Stanford, California: Stanford University, School of Education, March, 1969. 221pp. ED 029 001.
A careful review of relevant literature in the area of research on learning and individual differences. Provides methodological suggestions for better aptitude-treatment in interaction studies.

*Denton, W. T. *Context Evaluation, Needs Assessment, and Goal Setting: The Infinite Loop.* Paper presented at Evaluation of Instructional Materials Workshop, AECT Convention, Anaheim, California, March, 1976.
Discusses the relationship and methodology of context evaluation, needs assessment, and goal setting.

*Denton, W. T. *The Design and Implementation of Process and Interim Product Evaluation.* Paper presented at Evaluation of Instructional Materials Workshop, AECT Convention, Anaheim, California, March, 1976.
Discusses the role and methodology of process and interim product evaluation. Contends that formative evaluation of this type is the most important form of evaluation. Provides applied examples.

Educational Products Information Exchange Institute. "Looking Ahead: Selection Kits Will Blunt Buying Blunders." *Educational Product "In Brief" Report Number 32,* 4 (February, 1971), 5-6.
Discusses EPIE Kits that include techniques for evaluating and sources of information for selecting instructional materials and equipment.

Edwards, A. L. *Experimental Design in Psychological Research.* 4th ed. New York: Holt, Rinehart and Winston, 1972.
An introductory text on applied statistics and experimental design

*Contact William J. Webster, Deputy Associate Superintendent, Research, Evaluation, and Information Systems, S. J. Hay Building, 3801 Herschel Avenue, Dallas, Texas 75220.

for behavioral scientists. Concentrates on inferential statistics through univariate analysis of variance, with careful attention to experimental design. Particularly thorough coverage of analysis of variance for an introductory text.

Feldmesser, R. A. *Educational Goal Indicators for New Jersey.* Princeton, New Jersey: Educational Testing Service, February, 1973. 68pp. ED 077 945.
Description of a rationale for the needs assessment program carried out by the State Department of New Jersey.

Finn, J. D. *A General Model for Multivariate Analysis.* New York: Holt, Rinehart and Winston, 1974.
An intermediate level applied statistics text on multivariate analysis. The book relies entirely on the statement and formulation of linear models to solve multivariate problems. Closely aligned to the MULTIVARIANCE computer program, it uses sample problems to illustrate various analysis techniques--from multiple regression analysis through factor analysis, to multivariate extension of analysis of variance, analysis of covariance, and regression analysis.

Florida Educational Research and Development Council. *Plan for Study of the Educational Needs of Florida.* Gainesville: University of Florida, 1968.
Outlines a plan for a statewide needs assessment in Florida.

Glass, G. V. "Educational Product Evaluation: A Prototype Format Applied." *Educational Researcher,* 1 (1972), 1-4.
The primary source for the Trade-off and Comparative Cost approach to evaluation. Presents a model that could be used in examining instructional materials for possible adoption.

Glass, G. V., and Stanley, J. C. *Statistical Methods in Education and Psychology.* Englewood Cliffs, New Jersey: Prentice-Hall, 1970.
A particularly thorough introductory text on applied univariate statistics and experimental design. Discusses descriptive and inferential statistics through analysis of variance.

*Grobe, R. *Input Evaluation: Planned Development of Educational Strategies and Programs.* Paper presented at Evaluation of Instructional Materials Workshop, AECT Convention, Anaheim, California, March, 1976.
Discusses the methodology and purpose of input evaluation. Provides applied examples of the process.

Hunter, M. G., and Schooley, D. E. *The Synergistic Evaluation Model.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, February 25-March 1, 1973. 10pp. ED 075 503.
The primary source for the Synergistic Evaluation Model. Defines evaluation as a process for gathering and providing information for decision-making.

---

*Contact William J. Webster, Deputy Associate Superintendent, Research, Evaluation, and Information Systems, S. J. Hay Building, 3801 Herschel Avenue, Dallas, Texas 75220.

Kaufman, R. *Educational System Planning.* Englewood Cliffs, New Jersey: Prentice-Hall, 1972.
An introductory text on educational planning.

Komoski, K. "Instructional Materials: Do They or Don't They?" *Learning,* 3 (March, 1975), 92-93.
Describes a means for providing a systematic flow of information to help teachers and schools evaluate and select those teaching/learning materials that are most likely to work well with specific programs and students.

Komoski, P. K. *Development of a System for an Educational Products Information Exchange. Final Report.* New York: Educational Products Information Exchange Institute, November, 1967. 187pp. ED 020 566.
Outlines an easily accessible nationwide system for exchanging descriptive, evaluative product information among all sectors of the educational community on a cooperative cost-sharing basis. The feasibility of such an arrangement among educational professionals is investigated as a selection base for curriculum use or design improvement.

Kovac, R. Contact person for information on selecting and evaluating non-print materials (1976). Head of Selection, Catalog and Information Services, Purdue Audio-Visual Center, West Lafayette, Indiana 47907.

*Lamrock, A. L.; Smith, A. D.; and Warren, P. W. *Evaluation: Its Scope and System for Evaluating Development.* Ottawa: New Start Corporation, March, 1971.
The primary source for the New Start Evaluation Model. This model focuses on providing information to course developers so that they might make appropriate decisions to accept, modify, or reject developed components; obtain information for dissemination activities; and obtain information to aid in product installation.

Lindvall, C. M., and Cox, R. *The IPI Evaluation Program.* AERA Monograph Series on Curriculum Evaluation. Chicago: Rand McNally, 1970.
A primary source for the Individually Prescribed Instruction (IPI) Evaluation Model. This model focuses on providing reliable information for decision-making during the development stages of innovative programs.

Macy, D. J. "The Role of Process Evaluation in Program Development and Implementation." *Educational Technology,* 15 (April, 1975), 42-47.
An applied discussion of the role of process evaluation in program development and implementation.

Messick, S. "The Criterion Problem in the Evaluation of Instruction: Assessing Possible, Not Just Intended, Outcomes." In *The Evaluation of Instruction: Issues and Problems,* edited by M. C. Wittrock and D. E. Wiley, pp. 183-204. New York: Holt, Rinehart and Winston, 1970.
Discusses cognitive styles and affective reactions as two major classes of criterion variables that should be considered in the evaluation of instruction.

---

*Contact William J. Webster, Deputy Associate Superintendent, Research, Evaluation, and Information Systems, S. J. Hay Building, 3801 Herschel Avenue, Dallas, Texas 75220.

Miller, R. I. *Criteria for Assessing Instructional Units.* Lexington: University of Kentucky, Department of Social and Philosophical Studies in Education, October, 1969. Project funded by the Fund for Media Research, General Learning Corporation, and the Xerox Corporation.
Suggests a plan for assessing instructional strategies using 11 specific criteria. For each of the criteria a number of questions are listed for consideration in the overall analysis.

Morgan, R. L. *An Approach to Evaluation: A Model for Evaluating the North Carolina Exemplary Program.* Raleigh: North Carolina State University, Center for Occupational Education, July, 1970. Paper presented at the National Institute on Exemplary Projects in Vocational Education, Squaw Valley, California, July, 1970. 26pp. ED 042 910.
The primary source for the Apex Evaluation Model. It gives no explicit definition of evaluation, but implies that evaluation exists to serve the decision-making process.

National Institute of Education. *NIE Product Rating Form.* Washington, D.C.: National Institute of Education, 1974. 14pp. To be available through ERIC.
The form used by the National Institute of Education to facilitate review of the quality of educational products. This form was developed to indicate dimensions of apparent quality even when highly credible evidence is not available to confirm the ratings. The form is quite generalizable.

Niedermeyer, F. C. *Observation Procedures for Classroom Tryout of Instructional Materials and Procedures.* Los Alamitos, California: Southwest Regional Laboratory for Educational Research and Development, June, 1972. 15pp. ED 106 887.
Suggests certain procedures for reporting classroom observation of a program lesson. These procedures are illustrated in the observation of two laboratory programs, "Composition Skills" and "Drama and Public Speaking."

Ott, J. M.; Fletcher, S.; and Turner, D. G. "Taxonomy of Administrative Information Needs: An Aid to Educational Planning and Evaluation." *Educational Technology*, 8, 5 (1966), 29-31.
The primary source for the Ott Evaluation Model. Presents a classification system of decision situations that must be served by evaluation.

Peper, J. B. *An Ontological Model of Evaluation: A Dynamic Model for Aiding Organizational Development.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, February 25-March 1, 1973. 28pp. ED 078 039.
The primary source for the Ontological Evaluation Model. This model focuses on the organization as a whole, rather than on individual programs within the organization.

Provus, M. M. *Discrepancy Evaluation: For Educational Program Improvement and Assessment.* Berkeley, California: McCutchan, 1971.
The primary source for the Discrepancy Evaluation Model. Defines evaluation as the process of defining program standards, determining

whether a discrepancy exists between some aspect of program performance and standards, and using discrepancy information to change performance or to change program standards.

Rosenshine, B., and Furst, N. "The Use of Direct Observation to Study Teaching." In *Second Handbook of Research on Teaching*, edited by R. Travers, pp. 122-183. Chicago: Rand McNally, 1973.
Provides a review of the literature on the use of direct observation schema to study teaching. An excellent source on classroom observation procedures and models.

Rossi, P. H. "Practice, Method and Theory in Evaluating Social Action Programs." In *On Fighting Poverty: Perspectives from Experience*, edited by J. L. Sundquist, pp. 219-234. New York: Basic Books, 1969.
Discusses the role of formative evaluation in evaluating social programs.

Scriven, M. "The Methodology of Evaluation." In *Perspectives of Curriculum Evaluation*, AERA Monograph #1, edited by R. Tyler, R. Gagne, and M. Scriven. Chicago: Rand McNally, 1967. Out of print.
Focuses on curricular evaluation. Discusses, among other topics, the roles of formative and summative evaluation in providing information for curriculum development and validation.

Solomon, G. S. "Heuristic Models for the Generation of Aptitude-Treatment Interaction Hypotheses." *Review of Educational Research*, 42 (1972), 327-343.
Presents three heuristic models for the design of aptitude-treatment interaction studies.

Stake, R. E. "The Countenance of Educational Evaluation." *Teachers College Record*, 68 (1967), 523-540.
The primary source for the Countenance of Educational Evaluation. Suggests that evaluation ought to be concerned with three classes of conditions: antecedents, transactions, and outcomes.

Stufflebeam, D. L., ed. *Educational Evaluation and Decision Making*. Itaska, Illinois: F. E. Peacock Publishers, 1971.
The primary source for the discussion of the CIPP (context, input, process, and product) model for educational evaluation. Defines evaluation as the acquisition of information through formal means, such as criteria, measurement, and statistics, to provide rational bases for arriving at judgments in decision-making.

Tatsuoka, M. M. *Multivariate Analysis: Techniques for Educational and Psychological Research*. New York: John Wiley and Sons, 1971.
An applied text on multivariate analysis. Deals clearly and concisely with such topics as multiple regression, multivariate analysis of variance and covariance, principal components analysis, discriminant analysis, and canonical correlation.

Tripodi, T.; Fellin, P.; and Epstein, I. *Social Program Evaluation: Guidelines for Health, Education, and Welfare Administrators*. Itaska, Illinois: F. E. Peacock Publishers, 1971.
The primary source for the discussion of the Differential Evaluation Model. Defines evaluation as the systematic accumulation of facts for providing information about the achievement of program requisites and goals relative to efforts, effectiveness, and efficiency within any stage of program development. Such facts are incorporated into some designated value system for making decisions about social programs.

Uhl, N. P. *Encouraging Convergence of Opinion, through the Use of Delphi Technique, in the Process of Identifying an Institution's Goals*. Princeton, New Jersey: Educational Testing Service, February, 1971. 126pp. ED 049 713.
Discusses the Delphi Technique as a method for gaining convergence of opinion.

Weatherman, R., and Swenson, K. "Delphi Technique." In *Futurism in Education: Methodologies*, edited by S. P. Hencley and J. R. Yates. Berkeley, California: McCutchan, 1974.
Presents the methodology of the Delphi Technique, which is designed to encourage convergence of opinion.

Webster, W. J. *The Organization and Functions of Research and Evaluation in a Large Urban School District*. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C., March 30-April 3, 1975. 72pp. ED 106 345.
Outlines the purpose, organization, and functions of research and evaluation in a large urban school district. Major issues discussed include methods of ensuring the credibility of evaluation information, the establishment and maintenance of meaningful reporting cycles, the recruitment and employment of qualified personnel, the release and dissemination of evaluation and research data, and the relationship between research, evaluation, planning, and development.

Webster, W. J. "Technical Auditing Procedures." In *Educational Product Audit Handbook*, pp. 38-103. Arlington, Virginia: Institute for the Development of Educational Auditing, 1971.
Provides an overview of procedures and issues in evaluation design and a taxonomy of statistical techniques for use in specific problem situations.

Webster, W. J., and Mendro, R. L. "The Investigation of Aptitude-Treatment Interactions as an Integral Part of Program Evaluation." *Journal of Experimental Education*, 43 (Fall, 1974); 86-91.
Presents a model for integrating aptitude-treatment and trait-trait interaction studies into program evaluation.

Webster, W. J., and Mendro, R. L. "A Pragmatic Model for a Comprehensive Public School Research and Evaluation System." *Journal of Educational Research*, 69 (December, 1975), 156-159.
Presents an applied model for planning, research, and evaluation. Incorporates context, input, process, and product evaluation, as well as basic and applied research.

Webster, W. J., and Schuhmacher, C. C. "A Unified Strategy for System-Wide Research and Evaluation." *Educational Technology*, 13 (May, 1973), 68-72.
Discusses an operational context evaluation system designed to provide decision-makers with important baseline information about student, staff, and community characteristics.

Wittrock, M. C., and Wiley, D. E. *The Evaluation of Instruction: Issues and Problems*. New York: Holt, Rinehart and Winston, 1970.
Reports on the outcomes of a symposium designed to bring new approaches to the problems and issues of instructional evaluation. Topics include: evaluation theory, instructional variables, contextual variables, instructional criteria, and evaluation methodology.

Worthen, B. R., and Sanders, J. R. *Educational Evaluation: Theory and Practice*. Belmont, California: Wadsworth, 1973.
A book of readings that includes many of the classic works in the field of educational evaluation.

The material in this publication was prepared pursuant to a contract with the National Institute of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgement in professional and technical matters. Prior to publication, this manuscript was submitted to the executive director of the Educational Products Information Exchange Institute for critical review and determination of professional competence. This publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of either the reviewer or the National Institute of Education. Contract No. NIE-C-74-0027.

ERIC ORDERING INFORMATION

Documents with ED numbers listed in this paper may be ordered from: ERIC Document Reproduction Service (EDRS), P.O. Box 190, Arlington, Virginia 22210. A check for the price of the document, plus postage, as figured by the following price charts, must be enclosed with orders. These documents cannot be ordered from individual Clearinghouses.

# of Pages	Microfiche Prices	# of Pages	Photocopy Prices
1- 480	$.83	1- 25	$1.67
481- 576	1.00	26- 50	2.06
577- 672	1.16	51- 75	3.50
673- 768	1.33	76-100	4.67
769- 864	1.50	101-125	6.01
865- 960	1.67	126-150	7.35
961-1056	1.83	151-175	8.69
1057-1152	2.00	176-200	10.03
1153-1248	2.17	201-225	11.37
1249-1344	2.33	226-250	12.71
1345-1440	2.50	251-275	14.05
Add $0.67 for each additional 96 pages or fraction thereof.		Add $1.34 for each additional 25 pages or fraction thereof.	

U.S. postage must be added to the above charges.

1- 60 pages	21c
61-120 pages	30c
121-180 pages	39c
181-240 pages	48c
241-300 pages	57c
301-360 pages	66c
361-420 pages	75c

* * *

Documents without ED numbers may be ordered from the source given in the citation.