DOCUMENT RESUME

ED 138 032 EC 100 666

AUTHOR Macy, Daniel J.
TITLE Evaluating an Individualized Mainstream Special Education Program in a Large Urban School District.
PUB DATE Apr 76
NOTE 26p.; Paper presented at the Annual International Convention, The Council for Exceptional Children (54th, Chicago, Illinois, April 4-9, 1976)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Conceptual Schemes; Elementary Secondary Education; *Evaluation Methods; Exceptional Child Education; *Handicapped Children; Individualized Programs; *Models; *Program Evaluation; *State Programs; Student Placement
IDENTIFIERS Least Restrictive Placement Alternative; *Texas (Dallas)

ABSTRACT

Described is a model for the evaluation of the Texas Comprehensive Special Education Program (Plan A) based on providing each child with an individualized educational plan and the least restrictive educational placement alternative. The plan is reported to have been implemented on a pilot basis over a 3-year period in the Dallas school district. It is explained that the focus of the evaluation model is a set of 28 questions which solicited context, process, and product evaluation information from multiple sources. The kinds of information yielded by the model are said to include the degree of program implementation operating efficiency, parent and staff reactions, teacher reaction to mainstreaming, and student progress. (Author/DB)

Evaluating an Individualized Mainstream Special Education Program in a Large Urban School District

Daniel J. Macy, Ph.D
Senior Evaluator - Special Education
Department of Research, Evaluation, and Information Systems
Dallas Independent School District
Dallas, Texas

Paper presented to the annual meeting of the Council for Exceptional Children, Los Angeles, 1975.

Abstract

This paper describes a model for the evaluation of the Texas Comprehensive Special Education program (Plan A) implemented on a pilot basis in a large urban school district. Plan A attempts to provide each child with an individualized educational plan using one or more instructional arrangements to integrate the child, whenever feasible, into the regular curriculum.

The focus of the evaluation model is a set of 28 questions which solicit context, process, and product evaluation information from multiple sources. The kinds of information yielded by the model include the degree of program implementation, operating efficiency, parent and staff reactions, teacher reaction to mainstreaming, and student progress. The trend toward mainstreaming and accountability attest to the value of the evaluation model, and selected concepts and procedures within the model will be useful in most situations.

Introduction

The Texas Education Agency has directed that the new Comprehensive Special Education Program (Plan A) be implemented in every Texas school district by September, 1976. Plan A attempts to provide each exceptional child with an individualized educational plan using one or more instructional arrangements to integrate the student, whenever feasible, into the regular curriculum. This paper describes the Plan A evaluation model as it has evolved during the three years of Plan A pilot implementation in the Dallas Independent School District, Dallas, Texas. The Plan A evaluation model is expressed largely in terms of context, process, and product evaluation as developed by Stufflebeam (1968). The Plan A evaluation model incorporates three newly-specified principles of program evaluation and primarily demonstrates the roles and contributions of process and product evaluation in the implementation of a new program.

Plan A Program Structure

Plan A is one of the most promising programs of today's innovative efforts in special education. Plan A is innovative because it is a radical departure from the traditional special education model of self-contained classes, and Plan A is promising because it individualizes instruction to suit the peculiar characteristics of each child. The latter point, individualized instruction, seems exceptionally critical when we consider the heterogeneity within the special education population.

Plan A grew out of a statewide needs assessment study conducted in the state of Texas. The study revealed that less than 50% of all handicapped children in Texas received special education and, perhaps more importantly, that many of the existing special education programs did not meet the needs of the children served.

The Texas state legislature passed special legislation to create Plan A, and one of the most significant features of that legislation is the allocation of state monies on the basis of a school district's average daily attendance (Texas Education Agency, 1973). Consequently, the identification of X number of retarded or disturbed children does not affect the amount of money received by a school district. Under the traditional special education program, it was necessary to identify and to label at least eight children as being retarded before a district could receive state funds for a teacher for one class of retarded children.

A second significant feature of Plan A is that state monies may be used to employ not only instructional personnel but also teacher aides and professional supportive personnel which may include counselors, psychologists, educational diagnosticians, and liaison teachers. Consequently, a school district may provide a continuum of instructional units ranging from self-contained classes through part-time resource room attendance to full-time attendance in the regular classroom.

A third significant feature of Plan A is that services are available to many more children than before. Children aged 3 years through 21 years are eligible to receive Plan A services, and learning disabled children are also eligible to receive services. We expect that District-wide implementation of Plan A in Dallas will at least triple the number of children in special education.

According to state policy, the admittance of a child into Plan A must be the decision of an Admission, Review, and Dismissal Committee. State policy directs that Committee membership must consist of representatives from the following three areas: administration,

instruction, and appraisal or special education. Thus, an Admission, Review, and Dismissal Committee must be composed of at least three people. An individualized educational plan must also be filed for each Plan A student, and the progress of each student must be reviewed periodically.

The foregoing comments have outlined the major legislative directives and state policie of Plan A, but, as one may have sensed from these directives and policies, there is considerable latitude and flexibility for implementing Plan A at the local school level. There is no reason to expect Plan A in Dallas, Houston or El Paso, for example, to be similar and, in fact, they are quite dissimilar.

The current year (1974-75) is the third year of operation for Plan A in the Dallas school district, and Plan A is operating as a pilot program in two high schools, three junior high schools, and sixteen elementary schools. The Plan A program relies primarily on resource rooms for the delivery of instructional services and places the responsibility for appraisal, eligibility determination, and instruction on local school personnel. In the past, placement has been the responsibility of a central administrative committee, and instruction has been the sole responsibility of a special education teacher.

The Dallas Plan A pilot program uses four levels of appraisal to individualize instruction and services. Each progressive appraisal level becomes more complex in order to provide expertise to meet the appraisal needs of more severely handicapping conditions. In other words, a child with mild learning problems would receive only the first one or two levels of appraisal, but a child with more extreme problems would receive the third and fourth appraisal levels.

At the first level of appraisal, a committee of local school personnel reviews children who have been referred for learning problems. This review committee then refers the child to whatever supporting or alternative program that appears to be appropriate. As you can see from figure 1, Plan A is one of the options available to the level one reviewing committee.

Level two appraisal marks the beginning of the delivery of Plan A services by one or more Plan A personnel. At level two, for example, the student may attend the resource room on a part-time basis for reading instruction, the counselor may provide a positive role image through a schedule of periodic interviews, and/or the educational diagnostician may assist the child's regular classroom teacher in the use of a particular instructional strategy (see figure 2).

In practice, most children served at level two possess a learning disability or speech problem. The Plan A model specifies that problems of retardation, emotional disturbance, brain injury, and so forth must receive appraisal levels three and four. These levels directly involve the local Admission, Review, and Dismissal Committees. ARD Committees, that is Admission, Review, and Dismissal, are composed of the local school principal, resource room teachers, regular teachers, parents (occasionally), and the following intinerant personnel: educational diagnostician, psychological associate, liaison teacher, counselor, and speech clinician. These itinerant personnel are known collectively as itinerant appraisal teams. Each of the six Plan A teams serve two to five schools on a rotating basis. The purpose of the itinerant appraisal teams is to provide high-level expertise to local schools on a part-time basis.

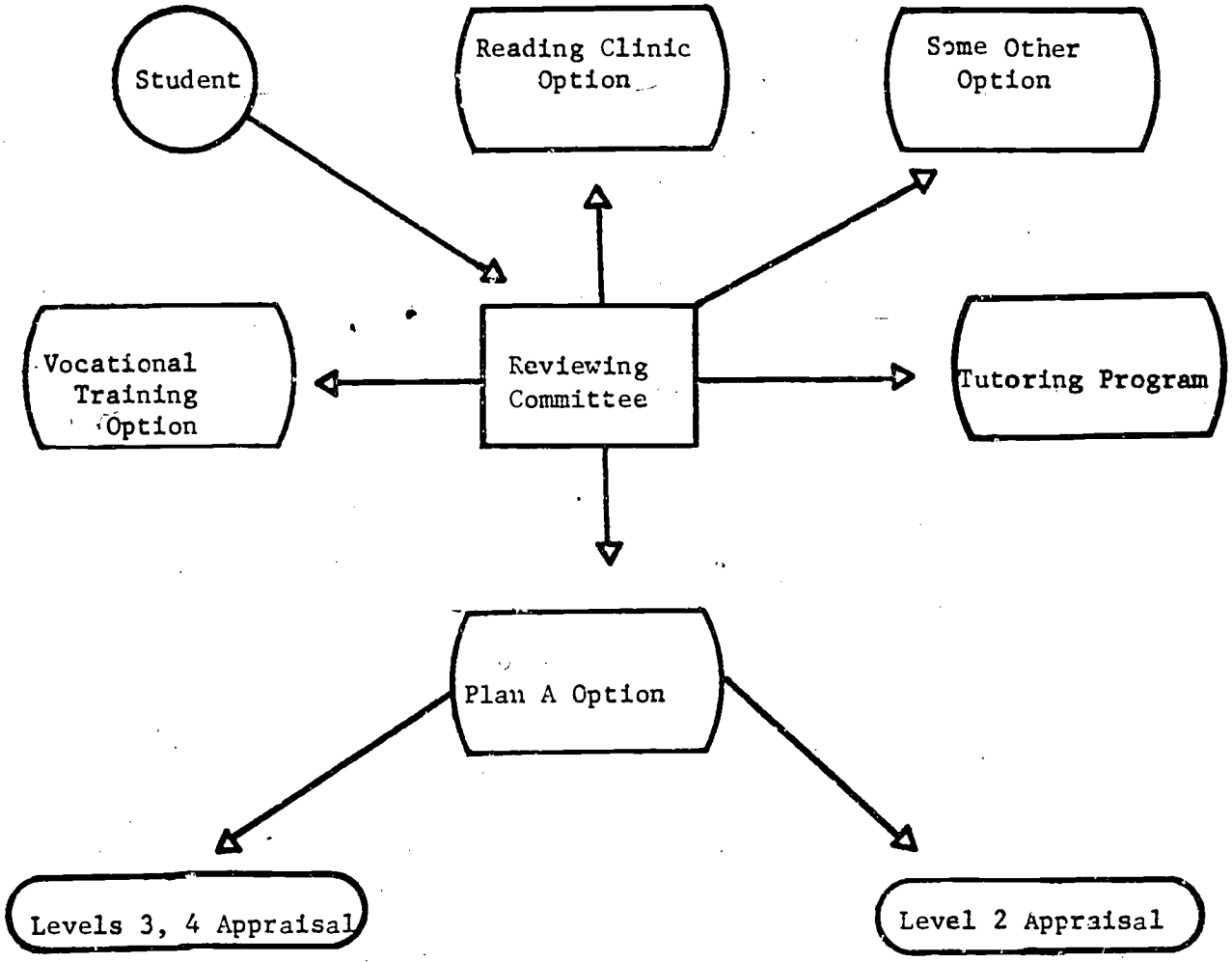

Figure 1

Level One Appraisal

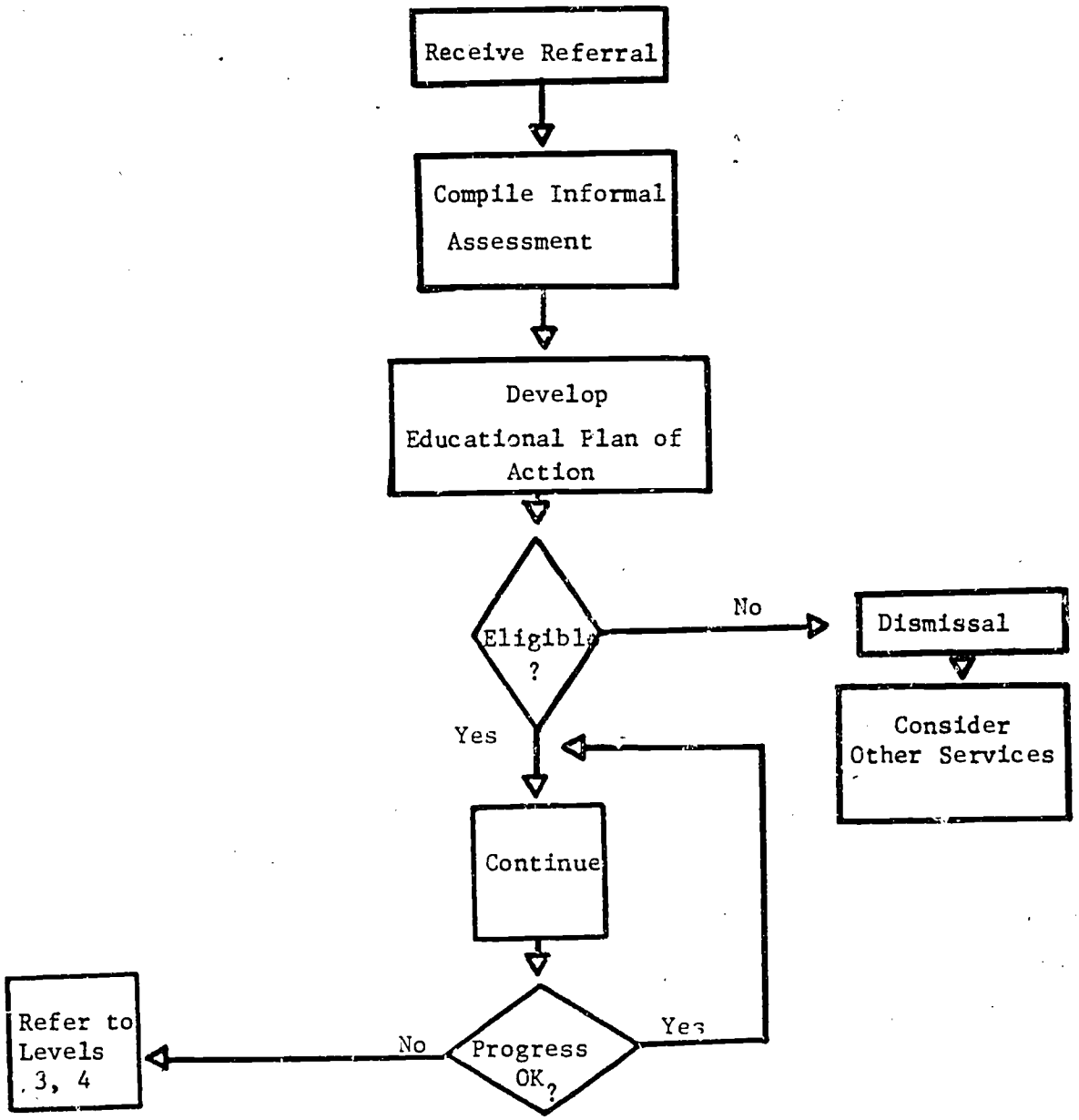

Figure 2

Level Two Appraisal

At the third level of appraisal, an ARD Committee meets in what is called a prestaffing in order to review a student's existing information profile, which will, of course contain information compiled at previous appraisal levels. The committee determines what additional information, if any, is needed to verify Plan A eligibility and to prescribe the individualized educational plan. Additional information specified in prestaffing may include intelligence testing, formal achievement testing, skills diagnosis, medical examination, classroom observations, and so forth (figure 3).

At the fourth level of appraisal, the ARD committee meets in what is called a staffing in order to review the information compiled from the prestaffing. The Committee determines if the student is eligible and, if so, prescribes the individualized educational plan. Such a plan may, for example, consist of attendance in a resource room for basic skills instruction and weekly conferences with the child's parent. Regardless of the level of appraisal, all educational plans specify behavioral objectives, enabling activities, and instructional materials (figure 4).

Educational plans developed at level four are based on in-depth appraisal compiled by a team of experts. Level two educational plans are more typically based on informal educational assessment and are developed by only one or two persons.

Future implementation of Plan A on a District-wide basis will probably involve a modified program structure, since operating policies and procedures have undergone considerable change during the past three-year pilot implementation of Plan A. However, it is felt that the basic procedures of referral and appraisal, as outlined in current program, will remain basically intact.

Principles of Program Evaluation

Having discussed the general programmatic structure of Plan A, let us turn our attention to ways and means of evaluating Plan A. The Plan A evaluation model emphasizes process evaluation in preference to product evaluation during the early phases of program implementation. Our experience continues to demonstrate the value of this approach in terms of information needs and program development.

The rationale underlying the Plan A process evaluation is best described in the words of Stufflebean (1968), who said:

> ....process evaluation is needed to provide periodic feedback to project managers and others responsible for continuous control and refinement of plans and procedures. The objective of process evaluation is to detect or predict, during the implementation stages, defects in the procedural design or its implementation. The overall strategy is to identify and monitor, on a continuous basis, the potential sources of failure in a project.

Thus, one important function of process evaluation is to identify problem areas in program implementation in order that management may possess the information, or at least part of the information, necessary for informed decision-making relative to program implementation. Problem areas in implementation may arise from unsound program structure or from failure of personnel to operate within the specified program structure. The function of process evaluation then is to identify or flag those problem areas in program implementation which deserve attention and may require intervention. However, the flag function of process evaluation, i.e., the identification of areas deserving the attention of management, should also include those outstanding positive features about program implementation in addition to the problem areas.

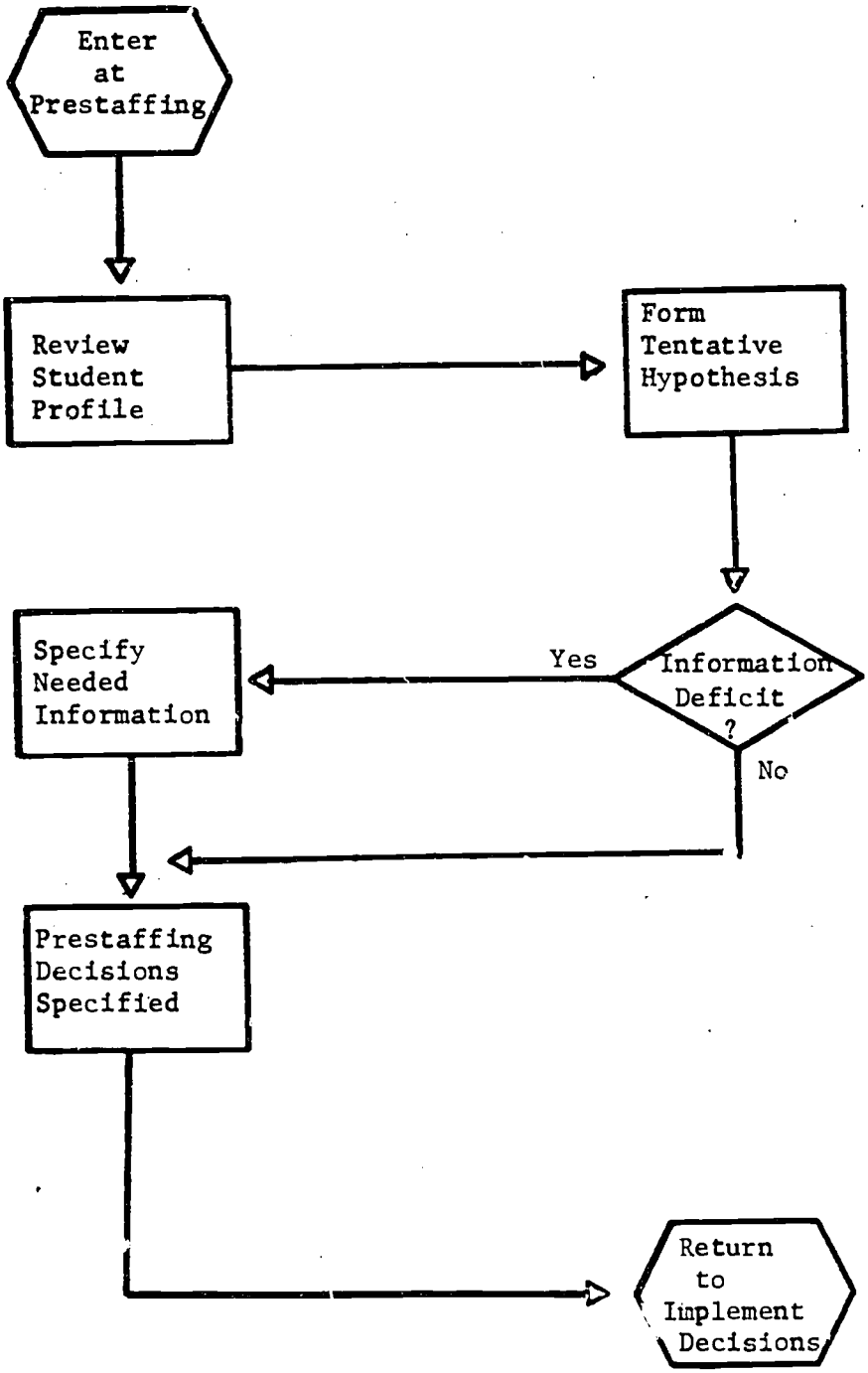

Figure 3

Level Three Appraisal

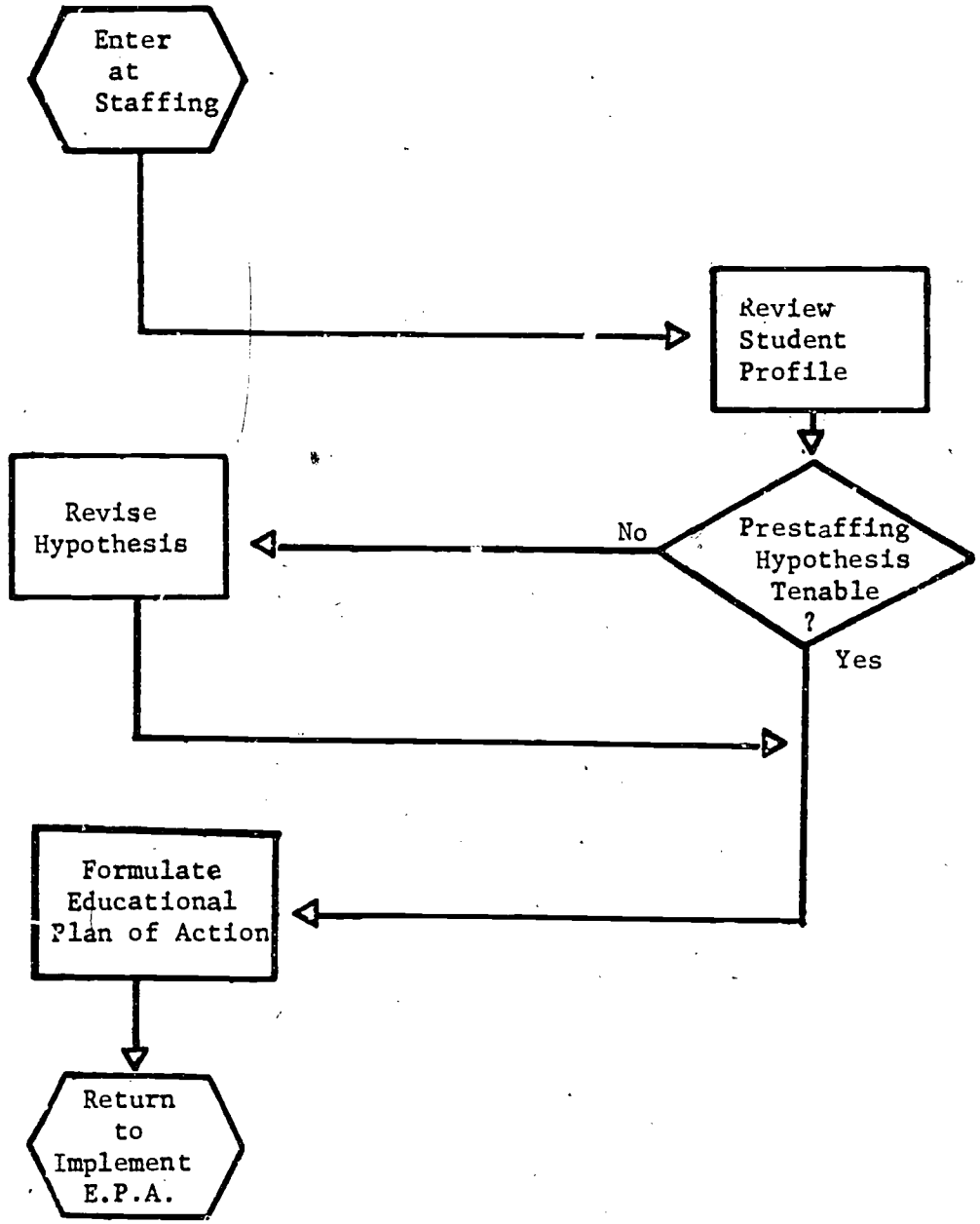

Figure 4

Level Four Appraisal

More definitively, then, the flag function of process evaluation provides management and personnel with the information necessary to determine any disparities between actual program implementation and program structure while enabling management to work toward the development of an optimal program structure through identification of desirable programmatic features. It is obvious that the very nature of the flag function dictates that such information be communicated to appropriate personnel as quickly and as efficiently as possible. However, one should realize that the option to react to process evaluation information resides with the responsibilities of program personnel and management, who may choose to delay revision or intervention for any number of reasons.

A second important function of process evaluation is to give validity to interpretation of product evaluation information. Although Stufflebeam (1968) generally defined product evaluation as the assessment of a project's effectiveness, product evaluation in the Plan A model primarily pertains to the assessment of program outcomes in terms of student achievements and changes. Thus, the validity function of Plan A process evaluation ensures that the interpretation of observed student achievements and changes will be valid in relation to the presence or absence of a program, at least as theoretically structured by program management. In other words, the interpretation that Plan A does not improve student attendance, when product evaluation reveals no attendance improvement, is obviously invalid and meaningless if students did not experience Plan A as specified by program structure.

It is clear that the validity function of process evaluation does not guarantee that the observed outcomes are attributable to the program in a causative sense, since the determination of causation also depends upon experimental design considerations implicit in product evaluation. Campbell and Stanley (1966), as well as others, have ably treated these considerations, and there is no need to elaborate in this paper. Rather, the validity function provides a necessary, but not sufficient, condition for the proper interpretation of product evaluation outcomes. Hence, if process evaluation were to show that Plan A had been implemented according to the specified program structure, and if project evaluation were to reveal no improved student attendance, the interpretation that Plan A did not improve attendance would only be meaningful to the extent of validity implicit in the research design of the product evaluation.

Concern expressed recently by Charters and Jones (1973) attests to the importance of the validity function of process evaluation. Charters and Jones observed that in many cases the difference between experimental and control programs is more fictional than factual, and they noted the ensuing difficulties in interpretation of evaluation findings. Stufflebeam (1968) originally pointed to the validity function when he cited the need to make rational interpretations of product evaluation outcomes in relation to process evaluation information. (Stufflebeam also included context and input evaluation, but the Plan A model includes only very limited context evaluation and no input evaluation.)

The validity and flag functions of process evaluation lead to the formulation of three principles that should be fundamental in the evaluation of educational programs (Macy, in press). The first two principles follow from both functions, while the third principle follows from the validity function. The first principle is as follows:

> the evaluation of new or developing programs should emphasize process evaluation in preference to product evaluation.

The assumption underlying the first principle is that new programs will encounter implementation difficulties. Hess and Buckholdt (1974) pointed out that the fragile status of new programs has been amply documented in the literature and that one is naive to assume the effective implementation and use of innovative programs. The author's experience supports Hess and Buckholdt on this point, since many of the programs implemented in the Dallas Independent School District experience implementation difficulty. Furthermore, the extent of difficulty is typically moderate and in some cases severe. The foregoing statement is not an indictment of these educators, but it is an indication of the difficulties inherent in establishing innovation and change in the real world of public education.

The second principle is as follows:

> the emphasis of the evaluation should change from process to product in accordance with the degree to which the program approaches optimal implementation.

This principle reflects the expectation that disparities between actual implementation and program structure, as well as the presence of unsound structural features, are likely to be more pronounced during the initial stages of implementation than during later stages. In chronological sense, one can view implementation in terms of two phases, the first being one of development and the second being one of maintenance. The development phase is that time from initial implementation in a school setting until the program achieves optimal implementation. The condition of optimal implementation is said to occur when program personnel function within the specified program structure. It is during the development phase of implementation that a program encounters most implementation difficulties, and it is during this time that the need for process evaluation information is most acute. The maintenance phase is that time after the attainment of optimal implementation and the time during which the program functions as a routine component of typical school operations. Of course, limited development may occur at various points throughout the maintenance phase.

The extent of formative activity which takes place within the development phase of implementation will obviously vary greatly across programs. The development phase of programs devised and piloted within the local school setting will necessarily contain a great deal of formative activity, whereas those programs devised and piloted elsewhere may experience little or no formative activity during implementation in the local school setting. This will be especially true for those programs used excensively by outside school systems and adopted for use by a local school system. It is important to realize that the implementation of all programs new to a school district or setting, regardless of genesis or previous success, will experience an identifiable development phase, since a given program structure must fitted to a given school setting, and since program personnel must operate within the specifications of program structure.

The third principle of program evaluation is as follows:

> process evaluation outcomes should determine the extent of product evaluation conducted in the evaluation of new or developing programs.

This third principle is a direct outgrowth of the validity function, since, to use research terminology, it is unreasonable to test for treatment effects if the subjects have not received the treatment. The third principle says that, as the degree of program

implementation varies from incomplete to complete, the emphasis of product evaluation varies from little or none to the maximum specified to meet information needs. The third principle implies that the proper justification for product evaluation, at least in a theoretical sense, is the degree of optimal implementation attained, since the validity function of process evaluation provides a necessary condition for the proper interpretation of product evaluation outcomes. Consequently, the collection of product evaluation data in the absence of process evaluation or the collection of extensive product data when process evaluation shows considerable deficiencies in implementation is most unreasonable, at least if one wishes to interpret observed product outcomes. Since process evaluation outcomes typically become available after the collection of product evaluation pretest data, adherence to the third principle may result in a decision to delete or reduce the collection of posttest data originally specified by the product evaluation design.

Before turning our attention to the Plan A evaluation model, it will be helpful to make a final observation regarding the operation of the principles of evaluation. It is these principles which determine the pattern of emphasis between process and product evaluation throughout the implementation of a program and especially throughout the initial developing phases of implementation. The principles of program evaluation would be most operative during the development phase, and the most pronounced changes in the pattern of emphasis between process and product evaluation would occur during that time. At the beginning of the development phase, the model should give process evaluation high emphasis and product evaluation low emphasis. By the end of the development phase, process evaluation should have low emphasis, and product evaluation should have high emphasis.

Figure 5 will be useful in summarizing the pattern of emphasis between process and product evaluation within the development and maintenance phases of implementation. Each of the three diagrams in figure 5 represent a pattern of emphasis between process and product evaluation for different programs. Each of the three programs (A, B, and C) experienced different degrees of success in implementation. One should note that the diagrams represent a highly improbable situation in that the proportion of development time is equivalent for three different programs. In reality, the proportion of development time would vary greatly across programs and settings.

If the pattern of emphasis in the diagram for program A represents a typical implementation situation, the diagram for program B is a situation in which the program approached optimal implementation rather quickly and efficiently, and the diagram for program C is a situation in which the program experienced considerable difficulty during the development phase of implementation. Accordingly, product evaluation of program B attained maximum specified emphasis rather quickly, whereas product evaluation of program C reached maximum emphasis much later than either program A or B.

Figure 5 also shows that the pattern of emphasis during the maintenance phase reflects the success of implementation during the development phase, since the maintenance emphasis of process evaluation of program C (which experienced considerable difficulty) was much higher than that of either A or B. The bump in the line representing maintenance process evaluation of program A indicates that the emphasis of process evaluation would typically vary in accordance with perceived need and would not necessarily remain static throughout the maintenance phase.

The diagrams in figure 5 also present several points which are common to all three programs. First, at the beginning of implementation, there is always a large discrepancy between the emphasis given process evaluation and the emphasis given product evaluation. A second observation is that, by the end of the developmental phase, product evaluation

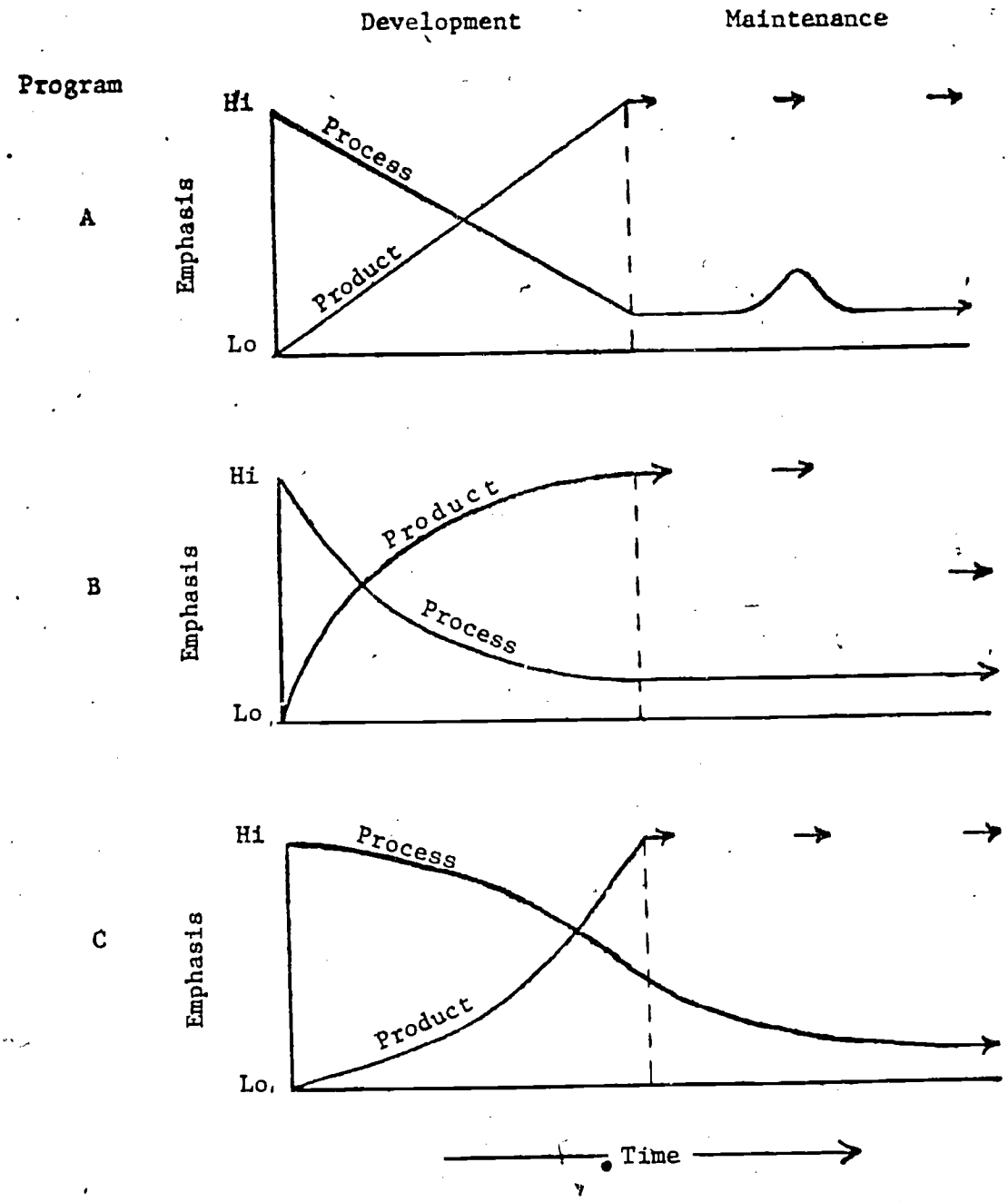

Figure 5

Three Possible Patterns of Emphasis
Between Process and Product
Evaluation Throughout Program
Implementation

has attained the maximum emphasis as dictated presumably by specified information needs, and the emphasis of process evaluation has diminished to the low level necessary to monitor continuing implementation adequately. A final observation is that process evaluation is continuous throughout the maintenance phase but that product evaluation is discontinuous and occurs only periodically throughout the maintenance phase. One should also observe that the emphasis of maintenance product evaluation need not remain fixed at a high level as is demonstrated for program B.

Plan A Evaluation Model

The foregoing development of the three principles of program evaluation provides a conceptual framework for the rationale underlying the Plan A evaluation model. As previously stated, 1974-75 is the third year in the development phase of Plan A implementation. The evaluation model placed primary emphasis on process evaluation and secondary emphasis on product evaluation during the first two years of implementation. The pattern of emphasis between process and product evaluation in the Plan A model was somewhat similar to that of program C in figure 5, as Plan A experienced considerable implementation difficulties during its first year.

Evaluation Questions

The evaluation questions provide the basic framework for the model and determine all ensuing evaluation activity. Determination of the evaluation questions must be a cooperative effort between management and evaluators in order for the questions to represent program objectives and related information needs. In determining the questions, managers and evaluators must weigh the importance of desired information against the cost of obtaining the information. Hence, the evaluation questions reflect priorities in identified information needs and specify the nature of the information sought.

During the first year of Plan A implementation (1972-73), the evaluation model specified 17 questions (3 context, 12 process, and 2 product). In the second year of implementation, the model specified 30 evaluation questions (2 context, 26 process, and 2 product). There were 14 evaluation questions specified in the third year of implementation (1 context, 4 process, and 9 product).

The following presents the complete set of Plan A evaluation questions specified in the model. These questions have been classified according to the kind of evaluation information solicited (context, process, or product). In most cases, the classification of questions is obvious, but the classification of some questions may appear arbitrary and open to question. However, the reported classifications were reasonable relative to the intent of each question and to anticipated use of information yielded by each question. For example, question number 39 (classified as product) might appear to solicit process information, but the intent and use of yielded information center about a judgment relative to the outcomes of Plan A. The "x" to the left of each question indicates the year in which the evaluation included that question.

1972-73	1973-74	1974-75	Context Questions
X	X	X	1. What demographic characteristics described Plan A students
X	X		2. What relationships, if any, exist between exceptional learning characteristics and selected student demographic characteristics?
X			3. What were the attitudes and opinions expressed by faculty members toward Plan A in schools scheduled for 1974-75 implementation?
$\overline{3}$	$\overline{2}$	$\overline{1}$	Total Entries

1972-73	1973-74	1974-75	Process Questions
	X		4. What was the extent of documentation in the Plan A records system?
X	X	X	5. How many students were referred to Plan A?
X			6. What were the sources of referrals?
	X		7. How many students were eligible to receive Plan A?
	X		8. Was there sufficient documentation to determine eligibilit
X	X		9. What were the reasons for referrals?
X	X	X	10. What was the frequency of occurrence of various exceptiona learning characteristics?
	X		11. How many students received level 3, 4 appraisal?
	X		12. Did the quality of appraisal meet minimum expectations?
X	X		13. What were the kinds of decisions made at prestaffings, and what was the frequency of each kind?
X	X		14. What were the kinds of decisions made at staffings, and what was the frequency of each kind?
X			15. How many people participated in prestaffings and staffings
	X		16. What services did students receive?
	X		17. What instructional arrangements were used in the delivery of services?
	X		18. How many students received educational plans?

1972-73	1973-74	1974-75		
	X		19.	Did the structure of educational plans meet program specifications?
	X		20.	Were the instructional objectives of the educational plan appropriate in relation to the student profiles?
X	X		21.	What was the extent of use of special instructional materials?
	X	X	22.	What was the extent of mainstream education received by former self-contained students?
X	X		23.	How much time did students (other than former self-contained students) spend outside the regular classroom in order to receive services?
	X		24.	What timelines described the progression of students through the appraisal process?
	X		25.	What proportion of the team populations (i.e., combined student population in those schools served by the appraisal team) was served at level 3, 4?
	X		26.	What proportion of staffing decisions (level 3, 4) were completed or initiated?
X	X		27.	How many children were dismissed from Plan A?
	X		28.	What were the reasons for dismissal?
X	X	X	29.	What was the reaction of appraisal team members and resource teachers to Plan A?
	X		30.	What steps were taken to prepare personnel for the implementation of Plan A?
X	X		31.	What steps were taken to educate the community about Plan A?
12	26	4		Total Entries

1972-73	1973-74	1974-75	Product Questions	
X	X	X	32.	What were the reading and math pre-post gains of Plan A students?
		X	33.	Given a three-year baseline profile of standardized test scores (reading and math), what changes were evident in score profiles after students entered Plan A?

1972-73	1973-74	1974-75		
		X	34.	What were the patterns in standardized test scores (reading and math) throughout two years of baseline experience and three years of Plan A experience?
X	X		35.	How did school attendance of students during the first year of Plan A compare to that of the previous year?
		X	36.	What were the patterns of school attendance throughout two years of baseline experience and three years of Plan A experience?
		X	37.	What was the school drop-out rate of Plan A students?
		X	38.	Were there differences between self-concept, attitude attitude toward school, and attitude toward reading of Plan A students and students in a comparison group?
		X	39.	What has been the three-year instructional program history of Plan A students in terms of Plan A, other special programs, and school transfers?
		X	40.	What has been the impact of Plan A on the total school in terms of the regular instructional program?
		X	41.	What are the attitudes of regular classroom teachers in Plan A schools toward Plan A, mainstream education, and individualized instruction?
2	2	9		Total Entries

Having considered the questions specified in the Plan A model, it will be informative to study the pattern of emphasis between process and product evaluation during the three-year implementation history of Plan A. Perhaps the most reasonable means of determing the degree of emphasis between process and product evaluation, as developed in the concept of pattern of emphasis, is in terms of the questions specified by the evaluation model. It seems reasonable therefore to argue that an evaluation model which contains one process question and ten product questions primarily emphasized product evaluation. Of course, the determination of emphasis in this manner is not precise since equal weithting of all questions is unreasonable. Presumably, one could devise a schema for more precise measurement of emphasis, but the need for such a schema does not appear to be acute, at this point in time.

The pattern of emphasis between process and product evaluation in Plan A could be represented as in Figure 6. One should note that there was insufficient emphasis on process evaluation during the initial year of implementation and that the third year emphasis of process evaluation reached a minimum level. Figure 6 also shows that the low level of product evaluation remained nearly constant during the first two years of implementation and then increased dramatically during the third year. The pattern of emphasis depicted for Plan A in Figure 6 implies that Plan A has reached optimal implementation, since the third year appears to terminate the development phase. While it is generally felt that the development phase of the Plan A pilot is complete, it would be extremely difficult to determine precisely the time when optimal implementation occurred.

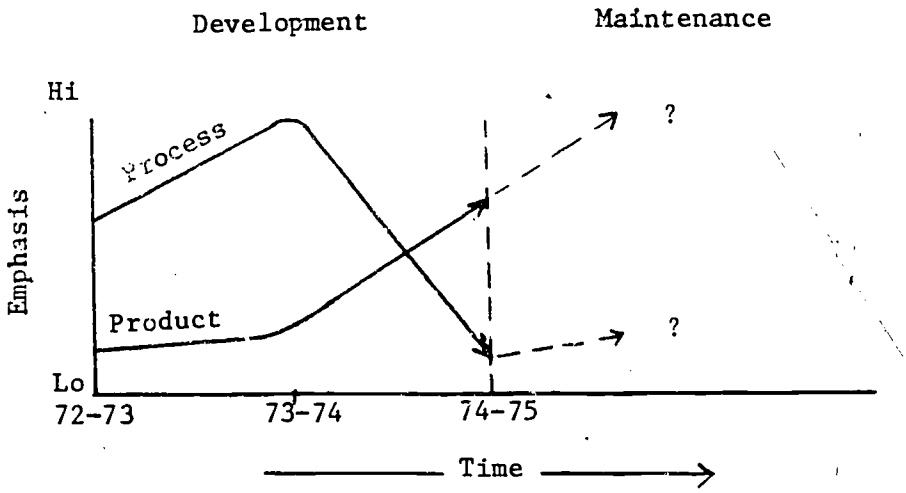

Figure 6

Pattern of Emphasis between Process and Product Evaluation throughout the Development Phase of Plan A Implementation

One can easily see that District-wide expansion of the Plan A pilot will result in the creation of a new development phase in the District-wide implementation of Plan A. Consequently, the evaluation model used in the District-wide Plan A implementation will give primary emphasis to process evaluation. However, the degree and duration of emphasis on process evaluation will likely not be as great as that in the Plan A pilot, as the District-wide implementation should encounter vastly fewer implementation difficulties. If so, more evaluation resources will be available for other kinds of evaluation.

Perhaps a final comment on the place of context evaluation in the Plan A model will be helpful. Stufflebeam (1968) indicated that the purpose of context evaluation was to define the environment in which change was to occur and that context evaluation would occur during the planning stages of the project. Inspection of the Plan A context questions shows that the Plan A model tends to define context evaluation as the description of those features that do not directly pertain to program operation or outcome (admittedly, this definition may lead to cases involving no clear distinction between context questions and other kinds of questions.). The The Plan A model also utilizes context evaluation during program implementation rather than during the planning stages.

Methodology

Many of the methodological concerns in the Plan A model obviously center about process evaluation. Again, Stufflebeam (1968) has provided a realistic basis for Plan A process methodology. Stufflebeam said:

> .... under process evaluation, the evaluator accepts the program as it is and as it evolves, and monitors the total situation as best he can by focusing the most sensitive and non-intervening data collection devices and techniques that he can obtain on the most crucial aspects of the project. Such evaluation is multivariate, and not all of the important variables can be specified before a project is initiated. The process evaluator focuses his attention on theoretically important variates, but he also remains alert to any unanticipated but significant events.

According to Stufflebeam, then, process data collection should be sensitive, non-intervening, and multivariate. The Plan A model relies on three sources of data for process evaluation. The major source is the Plan A reporting system, which consists of an extensive set of records designed to record key processes in the operation of Plan A. The second and third sources of process data are a mid-year survey of Plan A staff and school visits.

The "final" version of the Plan A reporting system was activated during the second year of Plan A implementation. The reporting system very closely parallels the levels of appraisal of Plan A program structure as described in the first portion of this paper. Each document within the system has been designed to facilitate and record a particular procedure or procedures in the Plan A structure. Relative to process evaluation, an outstanding positive feature of the reporting system is that process data collection is part of routine program procedure and is non-interventional. The reporting system was designed cooperatively by evaluation and management personnel in order to provide a unified reporting system which would meet identified information needs without duplicating paperwork. As such, the system was designed to provide an operational guide for Plan A program structure, to meet auditing requirements of the Texas Education Agency, and to provide management with desired information.

The reporting system documents are printed on three-part NCR paper so that a complete Plan A student folder can be available in the student's school, the itinerant appraisal team office, and the central administration building. A student's folder may contain from 4 to 14 documents depending on the levels of appraisal received by the student.

A student's Plan A folder may contain demographic information, comments from classroom teachers, results of appraisal, records of level three and four appraisal and service, educational plans, a medical examination report, sociological and psychological reports, eligibility verification, and a report from the year-end review.

During the second year of Plan A implementation, evaluation personnel routinely transferred information from the documents to coding forms as the documents fed into the central administration building. Key punching of data occurred in the first week of June, and a set of ALGOL computer routines provided analyses of each document as well as overall summaries of the contents of the reporting system. The computer routines generated numerous tables needed to respond to the process questions. The final evaluation report, which was available to the Board of Education in July, contained this information. (Plan A management received process evaluation results on an interim basis.)

Evaluation personnel also maintained a manual tabulation of the number and kinds of documents received from each school, and this procedure was most valuable in the generation of interim reports (October and January) to management. Table 1 presents the extent of documentation from three elementary schools as contained in one interim report. Such reporting has enabled management to assist better personnel in the field in implementing Plan A, since the documents paralleled the operational structure of Plan A.

Table 1

Extent of Plan A Documentation

School	Number of Plan A Students	Documents											
		1.0	1.1	2.0	3.0	3.1	4.0	4.1	4.2	4.3	4.4	4.5	234.0
A	107	55	41	67	5	2	6	14	10	4	14		2
B	68	30	31	32	6	5	6	7	2		1	1	
C	54	21	12	40		3	2	4		2	8		5

A brief inspection of the information in Table 1 can show one of the ways in which the reporting system was helpful. School A had 107 children in Plan A, but only 67 had an educational plan (2.0) on file, and state Plan A policy requires that all children have a written, individualized educational plan. The process information presented in Table 1, then, demonstrates the flag function of process evaluation, since the absence of individualized educational plans is inconsistent with Plan A program structure as well as with state policy.

One should note that the flag function of Plan A process evaluation, as realized in the reporting system, possessed some degree of error because of time lags in completing and transfering the paper work. In fact, it would be unfair to imply that the reporting system always functions as intended, since the complete adoption of the system was difficult for some personnel. However, the complaint that the paper work is too cumbersome or "gets in the way" was one indication that the complaining personnel may have been using a program structure other than that specified as Plan A. On the other hand, the paper work may be too cumbersome, at least at the central building.

It is evident from the evaluation questions that the third-year Plan A evaluation solicited only a small proportion of the kinds of information contained in the Plan A reporting system. However, the system was still used to meet state auditing requirements and to maintain continuity in the District's special education files. The reporting system is undergoing some revision to fit modifications in Plan A program structure in anticipation of District-wide implementation of Plan A.

The two remaining sources of process data are a mid-year survey of Plan A staff and school visits. The survey gives resource room teachers, itinerant appraisal team members, and school principals an opportunity to give management anonymous feedback. The survey of resource room teachers solicits open-ended responses to the following questions.

1. Do you feel that you are a contributing member of the ARD team effort? Why?

2. How much time are you usually able to devote to each child while he or she is in the resource room? Is this adequate?

3. Do you feel you have adequate interaction with the student's classroom teacher?

4. Are the educational plans appropriately individualized to suit the student?

5. Has the educational diagnostician been helpful to you in preparing plans and/or instructional strategies?

6. Do the appraisal team members, in your mind, make a positive contribution to the educational process in your school?

7. Are the physical arrangements associated with the resource room suitable?

8. Is your workload suitable?

The survey of team members solicits open-ended responses to the following items:

regular classroom teachers	educational plans
resource room teachers	instructional materials
team members	materials center
school principals	student testing
program management	professional workload
Plan A program structure	Plan A reporting system
program evaluation	

The survey of school principals contains the following questions:

1. Do your Plan A teachers have suitable instructional materials?

2. With current Plan A staff and budget, are you able to meet the needs of your students?

3. Do your Plan A teachers recieve adequate staff development and in-service training

4. What suggestions would you make regarding the Plan A reporting system?

5. Do you encounter any particular problems in verifying student eligibility for Plan A?

6. How do you percieve the role of the itinerant appraisal team?

7. What suggestions would you make regarding the use of appraisal personnel in future expansion of Plan A?

8. Do you encounter any problems in Plan A operating procedures?

9. What areas, if any, currently need administrative attention or clarification?

10. What kinds of evaluation information would be of personal value to you?

The purpose of open-ended survey items is to give respondents ample opportunity to express whatever issues and concerns, both positive and negative, that seem relevant to them. Experience has shown that the preferred means of reporting this information is a verbatim transcript of all responses grouped by survey item. The heterogeneity of responses makes summarization difficult, and management can best hear the survey by studying the verbatim responses. It is probable that open-ended survey items are most appropriate in those cases involving implementation difficulties.

The Plan A evaluation model does not specify a formal schedule of school visits. Rather, evaluators visit Plan A schools intermittently throughout the year (although most visits occur during the fall and spring) for the purpose of talking informally with Plan A personnel. Besides being an excellent source of process evaluation data, the visits are invaluable in terms of increasing knowledge and understanding of Plan A. The personal relationship with teachers and principals during these visits also does a great deal to dispel myths and fears about evaluation and to promote cooperation with program staff. One can also easily see how school visits are compatible with the flag and validity functions of process evaluation. In short, it is difficult to underestimate the importance of school visits, and process evaluation which does not make considerable use of school visits, either formal or informal, is inadequate.

It is obvious that all three sources of process data are subject to error and consequent misrepresentation of reality. Problems with the Plan A reporting system have already been discussed, and the survey of Plan A staff is plagued by the usual problems associated with voluntary-return surveys. Although error is unavoidable, the use of several data sources provides for possible replication of observations and may contribute to confidence in process evaluation findings. An example is that all three process data sources have identified operational problems surrounding the implementation of the itinerant appraisal teams. Consequently, one can be reasonably confident in the reliability of the process evaluation data regarding appraisal team implementation.

Recall that the third principle of program evaluation indicates that process evaluation should determine the extent of product evaluation. There is a critical need for evaluation studies to include an implementation score in order to identify the range of implementation disparities among educational units. Hess and Buckholdt (1974) have suggested the use of an implementation score and have also cited the need for identification of those variables related to program implementation. Accordingly, the Plan A evaluation model specifies product data collection in only those schools in which the implementation of Plan A has been within the specifications of program structure. The identification of those schools requires an operational definition of acceptable program implementation, and the formulation of the definition has been no easy task. In fact, initial efforts were only moderately successful. It is hoped that future efforts in this area can provide a schema for the determination of the extent of individualized instruction delivered within the Plan A resource rooms. Of course, a general operational definition of acceptable program implementation must address many additional factors such as decision-making processes, dissemination of materials, and so forth.

With respect to product evaluation methodology, the Plan A model relied primarily on pre-post comparisons of group measures during the first two years of implementation. After the second year, the shift of emphasis to product evaluation permitted the development of more refined product evaluation designs and limited use of comparison student groups. The use of the term comparison group rather than control group is significant, in that a contro in the sense of experimental design is almost nonexistent in the settings of program evaluation. Bosco (1971) has pointed out that control groups in evaluation studies do not provide a rigorous control and has also suggested the term comparison group rather than control group. Questions of design in product evaluation studies become critical when one wishes to attribute an observed outcome, such as improved attendance, to Plan A. The validity function of process evaluation can tell one that improved attendance might be attributable to Plan A, but the confidence one places in that conclusion rests upon the sound ness of the product evaluation design. In the case of Plan A, much of the product evaluation design would only support limited confidence in conclusions about program effects, but the possible replication of product outcomes across time will provide increased confidence in such conclusions.

A major goal in the increased emphasis of future product evaluation is the individualization of product evaluation. Although reading and math are the major academic needs of Plan A students, and although group test scores yield valuable data, Plan A is an individualized program. It seems reasonable that the assessment of student changes should also be individualized. A number of Plan A students, according to Plan A staff, have made marke and profound improvement in their abilities to function and learn within the social setting of the school, and Plan A product evaluation currently cannot document these outcomes. At this point, one might argue that the model's emphasis of process evaluation in preferenc to product evaluation (first principle of program evaluation) has been unwise, but the response is obvious. Without Plan A process evaluation, the reported student improvements might not have been so great or might not have occurred.

In progressing toward more individualized product evaluation, the current Plan A model includes the mastery of instructional objectives by individual students as a variable in product evaluation. This procedure looks promising, even though the wide range in the objective-writing skill of Plan A staff has made it difficult to determine mastery of many objectives. The inclusion of mastery of objectives in product evaluation should not only measure success of instruction but ultimately permit some degree of cost-benefit study. As yet there has been no formal documentation of changes in those students who have made marked and profound improvement, but this would be highly desirable.

In addition to being individualized, Plan A is also a mainstream special education program, and one of the context evaluation questions asked about the attitudes and opinions of regular classroom teachers in schools scheduled for 1974-75 Plan A implementation, Stern's Classroom Integration Inventory,[1] has been helpful in assessing opinions toward the acceptability of exceptional children in the regular classroom.

The Plan A evaluation model gives context evaluation only limited emphasis, but this should not indicate that context evaluation is relatively unimportant. Rather, the Plan A model is an outgrowth of needs and pressures as perceived in the real world. School boards and program managers typically desire product evaluation outcomes and the principles of program evaluation, as developed in this paper, assert the necessity of process evaluation.

[1]Haring, N. G., Stern, G. G., and Cruickshank, W. M. Attitudes of Educators Toward Exceptional Children. Syracuse: Syracuse University Press, 1958, pp. 143-146.

In the case of Plan A, process and product evaluation consume the bulk of evaluation resources, and there is little opportunity for context evaluation.

Many program evaluations do not typically include process evaluation, even though its inclusion could greatly increase the utility of evaluation studies. Bosco (1971) stated that evaluations of educational programs have been notoriously futile, since evaluation results could rarely help modify the program. The literature also contains numerous complaints of useless evaluations. The Plan A model, which incorporates the principles of program evaluation, does not generate useless information, and program management has found the evaluation to be most useful in program development as well as in meeting information needs. It is important to realize that the Plan A model consumes considerable resources and requires more than just one or two years for many information pay-offs. However, it should be obvious, and especially so to practitioners, that there are no short cuts to useful evaluation.

The principles of program evaluation outlined in this paper and the Plan A model demonstrate the importance of process evaluation. Unfortunately, the omission of process evaluation is one of the short cuts frequently employed. Evaluators may feel pressed by administrators and government officials for product evaluation infirmation and fail to give process evaluation proper consideration. Another explanation may be the potential threat that process evaluation can pose not only to classroom teachers but to administrators as well. Process information which reveals serious shortcomings in program implementation can be quite threatening to program management.

Other possible explanations for omission or minimal emphasis of process evaluation center about process methodology. Process data are generally more difficult to collect than the usual pre-post product measurements, since process data are much more heterogeneous in the kinds of observations taken and require much more epetitive observation. Additionally, evaluators with strong research backgrounds may tend to prefer product data over process data. At the present time, there appears to be little basis for developing a general linear model for the Plan A reporting system or even for an orthogonal rotation.

Finally, a general absence of process evaluation methodology may account for the infrequent use of process evaluation. Denton (1973) has presented a rationale for determining the length and frequency of observations in a schedule of classroom visits, but process evaluation methodology is generally in a primitive stage of development. It is hoped that the ideas expressed in this paper will be helpful, evan though there is ample room for further refinement and development.

REFERENCES

Administrative Guide and Handbook for Special Education: Bulletin 711, Texas Education Agency, Austin, Texas, 1973.

Bosco, James. The role of the administrator in the improvement of evaluation studies. Education, 1971, 92 (2), 70-74.

Campbell, Donald T. and Stanley, Julian C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1966.

Charters, W. W., Jr. and Jones, John E. On the risk of appraising non-events in program evaluation. Educational Researcher, 1973, 2 (11), 5-7.

Denton, William T. Program evaluation in vocational and technical education. ERIC Clearinghouse on Vocational and Technical Education, Columbus, 1973.

Hess, Robert J. and Buckholdt, David R. Degree of implementation as a critical variable in program evaluation. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April, 1974

Macy, Daniel J. Design for the evaluation of an individualized special education program. Paper presented to the third annual meeting of the Rocky Mountain Educational Research Association, Las Cruces, New Mexico, 1972.

Macy, Daniel J. Evaluation of an individualized special education program in a large urban school district. Paper presented at the fourth annual meeting of the Rocky Mountain Educational Research Association, Tucson, Arizona, 1973.

Macy, Daniel J. The role of process evaluation in program development and implementation. Educational Technology, 1975, 15(4), 42-47.

Stufflebeam, Daniel L. Evaluation as enlightenment for decision-making. An address delivered at the Working Conference on Assessment Theory, Sarosota, Florida, January, 19