An Overview of Curriculum Evaluation*

Kahsay Gebre**

Introduction

How do we know whether or not our educational goals and objectives are appropriate? How do we know whether or not there is a relationship among the prescribed goals and objectives, the selected learning experiences and assessment mechanisms? How do we know whether or not the designed curriculum is being implemented as planned? How do we know whether or not the prescribed goals and objectives are attained? Other than by intuition or other non-scientific, seat-of-the-pants method, it seems impossible to answer these and other related questions unless we conduct curriculum evaluation.

This paper attempts to provide a general view of curriculum evaluation. First, it highlights what curriculum evaluation is. Second, it presents some selected curriculum evaluation models. Then, it gives a short account of some curriculum evaluation techniques. Finally, it winds up with concluding remarks.

The Concept of Curriculum Evaluation

There is no one agreed upon definition of curriculum evaluation. Different curricularists and/or evaluators define it differently in light of the definition they give to curriculum and the respective purpose of evaluation. For instance, Brady (1987:160) reviews the following definitions:

- It measures the degree to which the performance of students meets behaviorally stated objectives.
- It compares the performance of students against certain standards
- It describes and judges the curriculum.
- It identifies areas for curriculum decision making, and selects and analyzes information relevant to those decision areas.

---

* Presented at a Workshop, On Improving the Quality of Curriculum Materials, Mekelle, Dec. 1998

** Kahsay Gebre, M.A in Curriculum and Instruction, Tigray Education Bureau,

An Overview of Curriculum Evaluation*

Kahsay Gebre**

Introduction

How do we know whether or not our educational goals and objectives are appropriate? How do we know whether or not there is a relationship among the prescribed goals and objectives, the selected learning experiences and assessment mechanisms? How do we know whether or not the designed curriculum is being implemented as planned? How do we know whether or not the prescribed goals and objectives are attained? Other than by intuition or other non-scientific, seat-of-the-pants method, it seems impossible to answer these and other related questions unless we conduct curriculum evaluation.

This paper attempts to provide a general view of curriculum evaluation. First, it highlights what curriculum evaluation is. Second, it presents some selected curriculum evaluation models. Then, it gives a short account of some curriculum evaluation techniques. Finally, it winds up with concluding remarks.

The Concept of Curriculum Evaluation

There is no one agreed upon definition of curriculum evaluation. Different curricularists and/or evaluators define it differently in light of the definition they give to curriculum and the respective purpose of evaluation. For instance, Brady (1987:160) reviews the following definitions:

- It measures the degree to which the performance of students meets behaviorally stated objectives.
- It compares the performance of students against certain standards
- It describes and judges the curriculum.
- It identifies areas for curriculum decision making, and selects and analyzes information relevant to those decision areas.

---

* Presented at a Workshop, On Improving the Quality of Curriculum Materials, Mekelle, Dec. 1998

** Kahsay Gebre, M.A in Curriculum and Instruction, Tigray Education Bureau,

. It uses professional knowledge to judge the ongoing processes involved in the curriculum implementation.

Nevertheless, what should be clear here is that though there are various definitions of the term, all have certain things in common. These are data collection, data analysis and/or interpretation, judgment and decision making. Without involving these steps, no evaluation will be carried out. Indeed, the way an evaluator defines evaluation will determine how he/she approaches it. For instance, if evaluation is defined as the degree to which student performance meets objectives, then behavioral objectives would be stated and the evaluator would measure the relevant students' behaviors. If evaluation is defined as identifying areas for curriculum decision making, the evaluator would select information related to the advantages and disadvantages of each decision alternative. Thus, it is recommendable that either evaluators should give operational definition of the term or select from the existing ones which best suit their purposes.

Evaluation may be carried out for one or more purposes. Anderson and Ball (1978) identified six purposes of curriculum evaluation (which are not necessarily mutually exclusive). These are:-

. to contribute to decision about curriculum introduction;
. to contribute to decisions about curriculum continuation and expansion;
. to contribute to decisions about curriculum modifications (improvement);
. to obtain evidence to rally support for a curriculum;
. to obtain evidence to rally opposition to a curriculum, and
. to contribute to the understanding of basic psychological, social, and other processes.

Curriculum Models

Evaluation models are simply guides or frameworks or designs. They refer to the conditions and procedures used by evaluators to collect data (Kaufman and Thomas, 1980:109). There are several designs or models, each having a slightly different purpose and is useful in slightly different situations. It is

important that evaluators become familiar with the range of models of evaluation. This is because a knowledge of several models of curriculum evaluation enables them to draw selectively from those models available in order to satisfy their own needs.

Actually, the basis for selection of specific model(s) is the type of question to be answered and the aspect of the curriculum one wants to evaluate, and it is the responsibility of the evaluator to specify exactly what questions to be answered and the aspect of curriculum to be evaluated.

It is beyond the scope of this paper, however, to examine all of the evaluation models. The selected models to be discussed here are: Scriven's Formative-Summative Evaluation Model, CIPP Model, and Tyler's Goal-Free Model.

Scriven's Formative - Summative Model

Formative - Summative Evaluation Model was coined by Scriven (1971). He suggested that evaluation is more than determining the extent to which goals have been met; it is also an assessment of the extrinsic worth of the goals themselves. He argues if the goals are not worth achieving in the first place, then why is it of interest to determine if these goals have been met? That is, if particular program goals are judged to be worthless, it is within the duties of the evaluation team to make recommendations regarding what appropriate goals might be (Kaufman and Thomas, 1980:110). Explicitly, Scriven in this model emphasizes the necessity for determining the worth of the goals and the provision of information for program designers and implementers.

Formative evaluation involves the assessment of progress toward the identified goals while the program is in progress. Under this model, *the midstream measurements* would be taken, and the extent to which the program appeared to be meeting these criteria could be assessed. It is a sign of success if the program works. If a particular method, such as an instructional strategy, or the use of media is not working as intended, this intermediate feedback could be used by the curriculum designers to make the necessary changes before any time or money is wasted (Kaufman and Thomas 1980:111).

In formative evaluation, the basic questions mostly raised are:

. Does the program seem to be working as originally planned?
. Are all components of the program functioning effectively or do some of them require revision?
. Are there serendipitous events that should be incorporated into the formal structure of the program?

Thus, formative evaluation is especially useful to curriculum developers. It can give them feedback before the entire instructional package is delivered so that modifications and adjustments can be made.

Summative evaluation, on the other hand, is useful in determining the extent to which the curriculum goals are actually met. It is carried out after the completion of the curriculum and addresses issues about the overall effectiveness of the curriculum.

Summative evaluation addresses questions such as:

. what do the clients know about the contents that were taught?
. To what extent have attitudes of the clients been changed as a result of this curriculum?
. Have the goals and objectives set forth by the curriculum developers been met?

In implementing the summative evaluation, care must be taken to select and/or design appropriate and valid instruments. That is, care must be taken to measure the actual effects or results of the program rather than some extraneous influences on the students.

CIPP Model

This model was coined by Stufflebeam and Guba (1971). CIPP is an abbreviation: C is the context; I, the input; P, the process, and the other P, product. This model requires that a series of decisions has to be made at each of the segments.

In this model, the emphasis is on the provision of information for decision makers. That is, data are collected and information presented to someone else who will determine its worth. And, the evaluator should be a person not directly connected with the program, but one who works with those involved in the program (the decision makers, or those who are ultimately responsible for the *go/no go* decision). Note that the evaluator works with the curriculum developers in specifying the information to be collected as well as in providing insights into the provision and interpretation of data used in the decision. Thus, this model requires team effort with emphasis on cooperation among those curriculum developers, as well as carrying out the evaluation.

Context evaluation is useful in the earliest phase of program development: the identification of needs and the designing of a rationale for the program. Typically, a thorough description of the existing program is developed, and then the present is compared with potential or possible programs. These objectives are then used to design an instructional program (Kaunfman and Thomas, 1980:116).

Input evaluation is useful in identifying what actually is required to meet the objectives defined in the context evaluation. The basic question here is can existing resources be reallocated or will additional resources be necessary? Again, the emphasis is on the gathering of information to be used by the decision makers. Here the decisions revolve around the issue of how to structure the instructional program to make the best use of resources in obtaining the identified program objectives (Kaufman and Thomas, 1980 :117)

Process evaluation has many things in common with formative evaluation. Both are used while the program is being delivered and are useful in determining if the program being delivered is as it was originally planned. The in-process, or ongoing evaluation is particularly useful in identifying strengths and weaknesses of the program that might not be identified after the conclusion of the program. An example would be an observation of the process of instruction and teacher-student interaction. Data gathered here can be used to improve the program before it is too late. This aspect of the CIPP model differs from Scriven's formative evaluation in that the emphasis here is on the gathering of information to provide a basis for informed decision making about

the progress of the program rather than making the actual program change (Kaufman and Thomas, 1980).

Product evaluation occurs during as well as after the program, with the emphasis on the gathering of the information necessary for decisions to be made regarding the program. Here the crucial questions are: should the program be continued, modified, or terminated? Should the objectives be redefined? What are the appropriate decisions to be made regarding the disposition of the program? Again, the emphasis is on provision of data necessary to make an informed decision (Kaufman and Thomas, 1980:117-118).

As can be seen from the description of the CIPP model, it is quite comprehensive and is useful during all phases of program design, development, and implementation and end of program assessment. Each is useful in particular settings or aspects of a total program. And, one may choose any one of the four phases if this best meets the requirements of the evaluation.

Tyler's Goal Attainment Model

The emphasis of this model, as implied by its title, is on the determination of the extent to which the goals defined for the program have been attained.

Briefly, Tyler's approach involves:

- Specification of goals or objectives,
- Stating the objectives in behavioral terms,
- Measuring aspects of student performance at the completion of the program,
- Comparing the test results with behavioral objectives (Brady, 1987: 182-183)

In this model, some of the questions to be raised would be:

- Are the objectives clearly stated?
- Is the content appropriate to attain the objectives?
- Are the methods appropriate to attain the objectives?

Are the assessment procedures appropriate to measure the objectives?
Is there an obvious link between the four curriculum elements?

For Tyler, evaluation is a continuous process. Feedback may lead to a redefinition of the objectives. Tyler believed that vaguely defined objectives have little value when evaluation is based on the degree to which the objectives are achieved. If, after the measurement of student performance, the objectives are attained, the curriculum is judged to be successful. And, it is for this reason that Tyler's approach to curriculum evaluation can be considered a 'goal attainment' model (Brady, 1987 183).

Tyler's model is less comprehensive than other models. Nevertheless, it has one important aspect not explicitly stated in the other models: the statement of the goals in terms of measurable behavioral objectives. Since results must be measured before decisions can be made, the concern with measurement early in the planning stages is very important.

Scriven's Goal Free Evaluation

There is an argument that paying attention to the prescribed goals limits the amount and sometimes the quality of the information available for decisions about the program. In those evaluation models in which the products or outputs of the program are compared with the goals set for the program, any unexpected important results or side effects would not be included in the main evaluation process and might be overlooked altogether. In other words, when evaluators confine themselves to looking only at prescribed goals of the intended effects, they may have a tunnel vision and miss some very important results.

Scriven proposed Goal Free evaluation in an attempt to identify all of the results of the program. The premise of Goal Free evaluation is that by not limiting oneself to or biasing the evaluation with the stated goals, the evaluation can be more open to the total impact-positive and negative, intended and unintended result of the program. Explicitly, Goal Free evaluation is an aproach to evaluation in which merit is determined by an examination of

program effects without reference to prescribed goals or objectives. It looks at the actual results of a program, whether planned or unplanned.

In this model, the evaluator must be able to specify a variety of ways in which a program could have potential impact on the clients, and then collect appropriate information (test scores, observation, informal interviews) to determine the actual impact. Here, the Goal Free evaluator must use his or her best professional skill to discover and document program effects.

In Goal Free evaluation model, the demonstrated needs are used as criteria for making favorable and unfavorable judgments. Advocates of Goal Free evaluation argue that the important criterion in evaluation is not the degree to which the program meets its goals, but the degree to which it meets demonstrated needs. Hence, the critical task for the evaluator is to determine the needs of the affected population, and these become the basis for judgments about program effects.

A Goal-Free evaluation may be used in conjunction with any of the goal-based models. By using both approaches the amount of gap filling (meeting identified goals) and the side effects, both positive and negative, may be identified. Using both models would address such questions as: To what extent were the goals identified actually met, and in so doing, were there any negative effects in the client group?

Techniques For Curriculum Evaluation

There are various data collecting techniques in curriculum evaluation. Each technique is appropriate for the collection of certain type of evidence or information. The evaluator has to select appropriate techniques which will provide the required data to answer his/her questions. And evaluators can construct their own instruments or can select or modify the existing ones.

Some of the data collecting techniques are the following:

Questionnaires

A questionnaire is a device consisting of a series of questions dealing with some psychological, social, educational problems is given or sent to an individual with the object to attain data. The questionnaire can be used to collect information from teachers, students principals and the school community in general. The essential steps in the development of questionnaire are:

- Determine the objectives of the questionnaire;
- Develop the questionnaire after determining the appropriate format;
- Check the validity of items (with an expert or panel of judges);
- Pilot the questionnaire using a small but representative sample;
- Refine the items in the light of the pilot study;
- Administer the refined questionnaire to the whole sample (but do not include those who participated in the pilot study).

Interview

The interview is a process of communication or interaction in which the subject or interviewee gives the needed information verbally in a face-to-face situation.

Rating Scale

Rating scale refers to a scale with a set of points which describe varying degrees of the dimension of attribute being observed. It can be used to judge the effectiveness of teachers and students performance or any aspects of school organization in a systematic way. For example: How effective was the presentation of material aids in the class by the teacher?

Very effective	slightly effective	Average	slightly ineffective	Very ineffective

Check list

A checklist is a simple device consisting of a prepared list of items which are thought by the evaluator to be relevant to the problem being evaluated. After each item a space is provided for the observer to indicate the presence or absence of the item by checking "yes" or "no", or a type or a number of items may be indicated by inserting the appropriate word or number.

Observation

Observation is the process in which one or more persons observe what is accruing in some real-life situation, and they classify and record pertinent happenings according to some planned scheme. It is used to evaluate the overt behavior of individuals in controlled and uncontrolled situations.

Anecdotal Records

Anecdotal records are descriptions of observed events. The observer records his or her observation usually in a few paragraphs of continuous prose. The record is made as soon as possible, after the event and an effort is made to keep fact and interpretation separate.

Pencil and paper tests of ability

These are used by the evaluator when a measure of student performance is required. They are used to measure student achievement or aptitude.

Teacher and student annotation of materials

Annotation of the materials and learning experiences involved in a curriculum provide the evaluator with relevant critical reviews.

Analysis of student work

This involves examination of student workbooks and practical work. It provides helpful information about student response to materials and learning experiences.

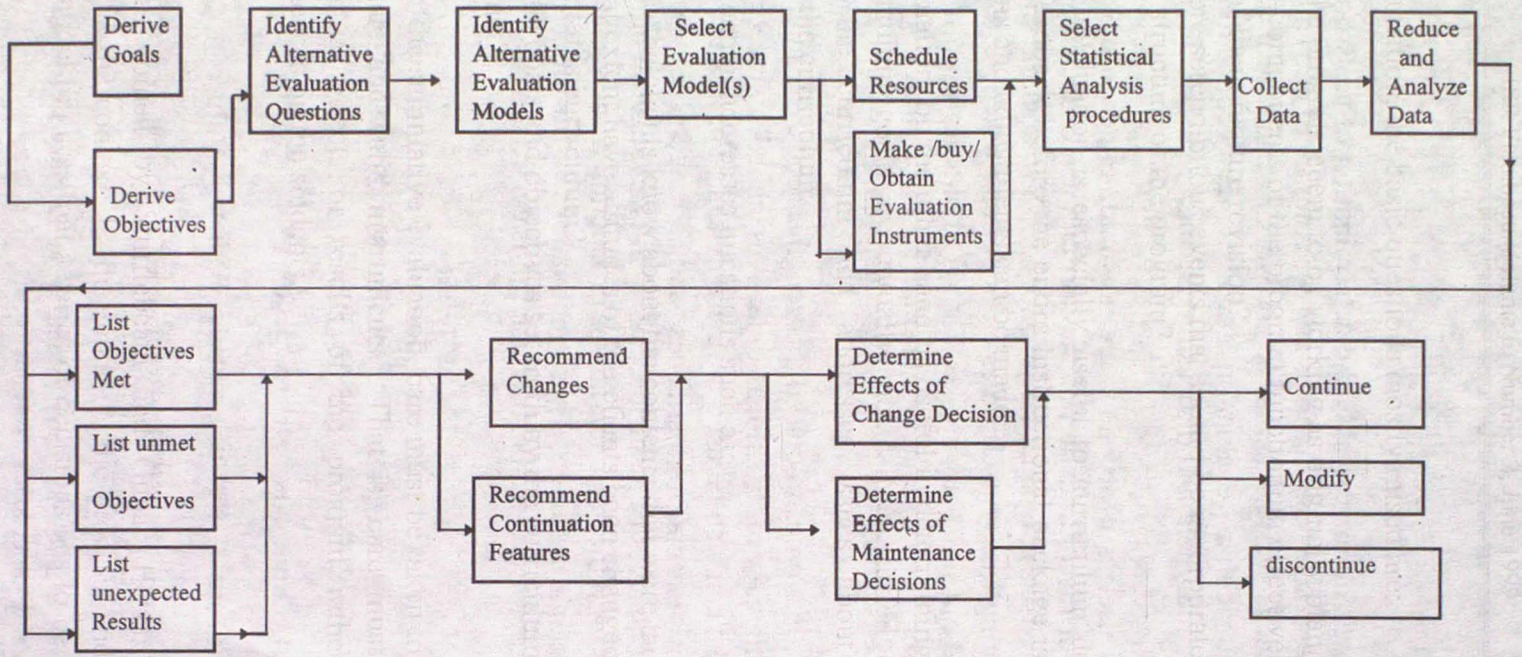
Discussions with teachers and students

Private records

These records include recording of absences, quality of assignments, number of books borrowed from the library, disciplinary action, peer-group participation.

Finally, it seems worthwhile to present here a simplified evaluation procedure. One might use the following flow chart as a guide in planning and conducting evaluation. But, this is a simple guide and can be adapted according to needs.

A simplified Evaluation Procedure



Source: Kaufman and Thomas (1980:222-223)

Conclusion

Curriculum evaluation is simply the determination of the worth of curriculum practices. It examines the strengths and weaknesses of curriculum practices. It pinpoints what is working, what is not working, and what should work. In doing so, curriculum evaluation involves data collection, analysis and interpretation, judgment and discussion making.

Evaluation is useless if decisions are not made either to the improvement or continuation or discontinuation of the curriculum. If decisions are not drawn, it is considered as if evaluation is skipped. Yet documentation of decision made is meaningless by itself unless it is put into practice. Evaluation should not be carried out for its own sake. That is, it should be problem oriented and problem solver. It has to have impact. The incorporation of results into the actual practice is a part and parcel of evaluation.

Results of evaluation can be used to strengthen ends. Results can be the basis for introducing new instructional objectives aimed at meeting demonstrated needs. Results can also be used to revise means. They can serve as a guide in planning new learning experiences and arrangements in alleviating the identified deficiencies of a curriculum. That is, evaluation pinpoints needs and guides a person in the selection and/or production of new material, procedures and organizational patterns. These innovations in turn must be tried out and their results appraised. In short, evaluation is cyclic.

References

Anderson, S B and Ball, S. (1978). **The Profession and Practice of Program Evaluation**. California: Jossey-Bass, San Francisco.

Brady, L. (1987). **Curriculum Development**. New York: Prentice Hall.

Kaufman, R. and Susan, T. (1980). **Evaluation Without Fear, New View Points**. New york.

Taylor, P. A. and Doris M. C. (1972). **Readings in Curriculum Evaluation**. Ontario: WM.C.Brown Company Publishers.

Wishon, Phillip M, et al (1978). **Curriculum for the Primary Years: An Integrated Approach**. Ohio: Jones, Worthington.