

## ● Research in Review

COLUMN EDITOR: JAMES RATHS

### Evaluating E.S.E.A. Projects for the Disadvantaged

J. RICHARD HARSH

AS THE programs under the various titles of the Elementary and Secondary Education Act begin a second year, there is hardly an educator in the United States who would not exclaim, "This has been the busiest period of my professional life!" Most would agree this escalated pressure of activities has resulted from the pressing deadlines to write projects, obtain staff and materials, and organize new programs, while at the same time maintaining the regular educational program.

At present it is important that educators take time to make a "scorecard" of the process and products of the special project efforts to date. This appraisal should seek to review what has actually occurred in the first inning of the special projects. A review of the targets for accomplishment and the means to get there in relation to the results which have been attained should provide some illumination for developing better programs in the succeeding innings of the special projects. Hopefully, this would aid us in avoiding the characteristics of the lightning bug, ". . . who is brilliant, but hasn't any mind, for he blunders through existence with his headlights on behind."

#### Scoring Criteria: Steps in Evaluation

One approach to developing a scorecard or appraisal of the processes and products of the E.S.E.A. programs to date is to utilize the five steps of the evaluation process found in any explicit and systematic study. Each project may be considered in relation to how well its development and progress have met these five essential steps. Such an appraisal should provide important suggestions regarding the conceptualization of the project and the ways in which it may be improved.

*Step One—What was stated as the purpose of the project?* This step in the evaluation process is often completed with a minimum of effort and concern. Part of this is due to the fact that broad goals or objectives may be stated in sufficiently general terms so that they form an umbrella to encompass great diversity in process and product. Moreover, the statement of goals and objectives rarely differentiates among means, short-term educational experiences, ultimate achievements or outcomes.

Obviously it is desirable to refine and make objectives as specific and explicit as possible for the purposes of the study. However, statements of goals and objectives are often written to encompass a wide range of activities and pupil characteristics. Broad goals or objectives frequently become latch-strings to ambiguity and diffusion concerning the nature of the phenomenon that is to be studied. This condition creates the necessity for the second step in the evaluation process.

*Step Two—How can I recognize the outcomes if I see them?: Definition of objectives in behavioral terms.* This step in the evaluation process is a rigorous and demanding activity, for it requires that objectives or anticipated outcomes be defined and described in terms of behaviors that may be observed, recorded and analyzed. For example, the objective, "to develop and enrich the language of the disadvantaged," offers ambiguous direction for data gathering or assessments. The test of behavioral definition is to find consensus that the behavior's presence or absence represents attainment or non-attainment of the objective.

*Step Three—What is going to be done to achieve the objective?: The delineation of situations and experiences through which the attainment of the objectives of the program is to be accomplished.* This step in the evaluation process is frequently given more general than specific treatment. The situations in which the objectives are to be attained are often not examined in relation to their relevance for producing the desired behavior. Moreover, it is not uncommon to find that description of the

process, situation or environment of an educational program fails to identify specific methods, personnel or materials. This step in the evaluation process also demands rigor in order to give specificity to the "what," "when," "why" and "how" of the educational experience.

*Step Four—What information will be gathered as evidence of the attainment of the objectives?: The development or selection of appropriate assessments of the behavioral objectives.* The effectiveness of evaluation depends on the relevance and reliability of the data accumulated concerning the program. The selection of data-gathering devices should be directed by careful analysis of the anticipated behaviors that are defined in Step Two of the evaluation process. Too often general and "venerable" measures of achievement which have titles as tests of reading, language, math or interest, etc., are assumed automatically to provide a measure of any and all behaviors that relate to that domain of development.

It is essential that data gathering for programs dealing with diverse student populations and innovative programs should not be restricted to measures previously standardized on normal populations within traditional programs. The items found in standardized tests are usually selected because of their "general fit" to typical student populations in general areas such as reading or language. The studies by Frazier, Loban and others have indicated that language and reading have many dimensions with diverse relevance to various populations. Obviously, the generalized definitions of the materials which make up a content-domain may vary in dif-

ferent localities and may also be inappropriate for diverse populations. Assessments must be pertinent to the statement of objectives and the definitions of behavior to be attained.

Unfortunately, assessment has frequently been perceived to mean only standardized testing. In contrast, assessment includes many forms of data-gathering procedures. Observation is a form of assessment which includes the same properties of concern as any standardized test—reliability, standard error, validity, etc. Probably observation has greater flexibility for gathering data concerning specific behavioral objectives than general standardized assessments designed to give generalized appraisals of content-domain of the curriculum. There are some who would deery that observation lacks sufficient

objectivity and reliability to be considered as an adequate form of data gathering; however, it should be argued that a *systematic development of an observational process*, which explicitly pre-defines the time, circumstance, situation and behavior that will be observed may have as high reliability and as great validity for the questions that are to be judged or evaluated as some of the standardized instruments that lack the precision of "fit" to the particular objectives of the program.

Another very direct form of assessment is the analysis of students' products. Such an analysis done in a systematic manner, with the firsthand data of the productions from children, may offer evidence of the attainment of behavioral objectives. Some are loath to consider this because involved in prod-

## SCIENCE SERIES ACCENTS



### **INQUIRY, EXPERIMENTATION, INVESTIGATION**

Here's "Process Science" in a new elementary Science Series, Grades 1-6, that provides unlimited fresh opportunities and sets up an environment for pupils to become personally involved in science through experiments and open-end questions.

This process-centered, inquiry series consistently invites the pupil "to think like a scientist." It develops and uses the techniques of science based on analysis, observation, experimentation and careful thought.

We invite you to evaluate this fascinating new approach to science instruction. See the new LAIDLAW SCIENCE Series.



## **LAIDLAW BROTHERS**

A Division of Doubleday & Company, Inc.

Thatcher and Madison, River Forest, Illinois

uct analysis is always the need to develop criteria by which the products may be judged. However, with explicit behavioral objectives, criteria for judging the levels of adequacy or attainment in relation to the stated or anticipated behavioral outcomes may be made.

*Step Five—The fifth step in the evaluation process* deals with the summarizing and analysis of data (and the judgments or evaluation of the outcome). It has been observed that there is a natural tendency for many to approach the evaluation process by defining broad objectives and then jumping to the fifth step of making summary conclusions concerning the outcomes or products of the activity.

In the fifth step of the evaluation process, there is need for rigorous examination of the adequacy of the data which have been accumulated and summarized to represent attainment of the behavioral objectives described in Step Two of the evaluation process. The basic question is concerned with whether the data accumulated provide acceptable evidence upon which judgments or evaluations can be made. Furthermore, the data must also be examined in relation to adequacy for reflecting behavioral changes as a result of the situations designed for such attainment; perhaps the change in behavior may have resulted from concomitant factors unrelated to the particular project and design.

### **After a Scorecard, What Action May Be Taken?**

Each of us concerned with education should consider the projects we are engaged in from the standpoint of how

they may substantially contribute to further understanding of education and more effective learning of the students. Many members of the Association for Supervision and Curriculum Development are at this time directly or indirectly involved in some federal, state, local or professionally-sponsored project with research potential.

What might each of us do to increase the probabilities that the new projects and research studies will make a substantial contribution to the development of education? As we reflect upon our own attitudes and reactions to research and evaluation studies we have reviewed, the following guidelines or commitments for future activities are suggested:

1. To budget some time to review critically the adequacy with which the five steps of the evaluation process have been used.

2. To persist in defining the (a) content and process of the activity, (b) characteristics and circumstances of the learners, (c) location and situation for learning, (d) behavioral objectives or results anticipated until these are communicable to participants and observers with common meaning. For example, use the simple two-way test: If you and I observe the aspects or resultant behaviors of a project, will we be able to perceive that which is said to be occurring in the report?

3. To clearly differentiate planning and developmental experiences from investigations designed to determine the outcomes of a prescribed educational experience.

4. To insist that evidence to support or refute a research hypothesis or a

behavioral objective is composed of data with demonstrable reliability and validity.

5. To make explicit descriptions of the sampling procedure by which individuals and products were assessed in the data-gathering process.

6. To provide for the presentation of data in the form and units by which the data were gathered; e.g., incidents of observed behavior, quantity and quality of product or test performance should be presented in the units of measurement rather than summarized, converted, equated or interpreted by "perceptions of previous behavior," or "norms," or comparisons with other general but unassessed groups. Such pooling of data, conversions, or interpretations provoke questions or uncertainty as to the basis upon which the situation is to be judged.

7. To recognize and identify the multiple variables that are present in human behavior and devote special attention to the possibilities of planning cooperative or related studies which may control or examine the effects of various combinations of dependent variables.

8. To give special attention to determining the relevance of "venerable assessments" constructed and normed in another era for data-gathering in new programs of today. Common usage and tradition tend to promote "face validities" which may be erroneous, partial or distorted snapshots of what presently exists.

There are numerous examples in the current periodicals of education being analyzed and evaluated by nonprofes-

sional educators. Frequently the "case of one child" or the lacks of an adult population are cited as proof of the inadequacy of education. Insofar as the goals and objectives remain in broad, ambiguous and unbehaviorally defined terms, each person from any frame of reference may make evaluations on the basis of his unique and undefined point of view.

In contrast, the opportunities afforded by the special projects should result in more extensive information concerning the description, prediction and control of the course of education.

In the year 2000 when the inquiring student consults the reviews of educational research of the 1950s and 1960s, he will undoubtedly be impressed with the tremendous quantity of research projects during the 1960s. When we pause, as we must, to consider the status of these various new projects, we should give some thought to the question, "Can we anticipate that the reviewer of educational research done in the 1960s will find some substantial contribution to the understanding of human behavior and the learning process?"

#### **Suggested References**

*Review of Educational Research* 36 (3): June 1966; Chapters I, II, VI.

*Review of Educational Research* 35 (5): December 1965; Chapters I, II, III, IV, V.

*Language Programs for the Disadvantaged*. Champaign, Illinois: National Council of Teachers of English, 1965.

—J. RICHARD HARSH, *Educational Testing Service, Southern California Regional Office, 2200 Merton Avenue, Los Angeles.*

Copyright © 1967 by the Association for Supervision and Curriculum Development. All rights reserved.