061 281 TM 001 232

HOR Fisher, Maurice D.; Ward, Virgil S.

LE A Design for Evaluating Educational Programs for Culturally Disadvantaged Children.

DATE Apr 72
E 17p.; Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, April 1972

S PRICE MF-$0.65 HC-$3.29
CRIPTORS *Academic Performance; Cognitive Processes; *Culturally Disadvantaged; Disadvantaged Youth; Educational Objectives; *Educational Programs; Evaluation Criteria; Evaluation Needs; *Evaluation Techniques; Formative Evaluation; Instructional Programs; Models; *Program Evaluation; Rating Scales; Research Methodology; Taxonomy

TRACT

The evaluation design described in this paper cusses all significant aspects of program development and low-through for culturally disadvantaged children. Discussed are cational objectives, instructional methods, implementation cedures, and outcomes. (CK)

A DESIGN FOR EVALUATING EDUCATIONAL PROGRAMS FOR CULTURALLY DISADVANTAGED CHILDREN*

BY

Maurice D. Fisher

Center for Research and Development in Teaching,

Stanford University

and

Virgil S. Ward

Foundations of Education Department,

University of Virginia

*Paper presented at the 1972 American Educational Research Association meeting in Chicago -- Division H, Session 3.15, April 4th.

Compensatory education has been recognized as a significant method for improving the academic performance of culturally disadvantaged children because it concentrates upon removing deficits which appear to hinder their further educational progress. However, the outcomes of many compensatory education programs indicate that they do not produce the improvements in academic performance which are anticipated by their educational objectives. One particular reason as to why expected outcomes are not attained in many of these programs is that they are not based upon a theoretical rationale which would make them acceptable in the sense required by scientific inquiry.

It appears that educational programs which do not have a rationale derived from relevant educational and developmental theory, which explains why that given approach should be utilized with particular groups of culturally disadvantaged children, may eventuate in failure because the applied educational objectives and methods were inappropriate for the group of children under investigation. However, in such situations, it is difficult to determine the cause(s) of failure because no theoretical framework would have been involved to provide information concerning how the applied educational objectives and methods were expected to reduce specific types of deficits. On the other hand, a program which has a theoretical basis in developmental and educational theory can provide useful information concerning why it should be applied to certain types of culturally disadvantaged children, and this information can be utilized to explain either the success or failure of a program.

Programs which do not have an explicit and valid hypothetical framework can produce disillusionment among educators because they (educators) are not provided with clearcut guidelines concerning the populations to which these programs should be applied, and therefore, they will have difficulties in explaining why particular outcomes were obtained. These difficulties which are encountered in devising effective programs for culturally disadvantaged children demonstrate the desirability for establishing a procedure by which to evaluate them, since diversity of conception without disciplined criteria as guides permits activity which is likely to result in waste or possibly undesirable outcomes.

The evaluation design which is described in this paper takes account of all significant aspects of program development and execution, including educational objectives, instructional methods, implementation procedures, and outcomes, and the construction of this comprehensive design involves a set of criteria that can be applied to evaluating these four phases of educational programs. In the preparations leading to the evaluative system as finally conceptualized, the following activities were undertaken:

I. Educational Objectives: The selection of a scheme for arranging normal developmental processes into a chronological sequence that was based upon the research and theory of Piaget and Inhelder (1969), and which was then used to: (1) identify the deficiencies of culturally disadvantaged children, and (2) organize these deficiencies into a taxonomy that follows the order of normal cognitive, social, and affective development.

II. Instructional Methods: The identification of instructional methods which have been utilized by exemplary compensatory education programs, and the utilization of the Piaget and Inhelder scheme to construct a taxonomy of instructional methods that is arranged according to the areas of development they concentrate upon.

III. Implementation Procedures: The specification of implementation procedures which have been utilized by some of the most outstanding programs for culturally disadvantaged children, and the organization of this information in terms of the characteristics of superior implementation procedures.

IV. Outcomes: The identification of experimental designs with high amounts of internal and external validity, these designs being recommended by authorities in educational research, the isolation of standardized tests which have been utilized by exemplary education programs which are then arranged into a taxonomy that is parallel to the scheme of human development, and the specification of guidelines for testing culturally disadvantaged children that are derived from the recommendations of experts in this area of education.

Although a variety of educational programs have been developed for culturally disadvantaged children, it appears that many of these programs are similar in terms of their effectiveness in remedying particular types of deficits. The present evaluation design can be used as a frame of reference for deciding upon which one of several different programs should be selected for application to a particular educational situation, since this design is constructed to facilitate the selection of programs that demonstrate the most promising methods for removing deficiencies in these children. For example, this evaluation design will enable educators to choose between language training approaches which appear to produce similar outcomes but which utilize apparently different educational methods.

Illustrations of language training approaches which appear to be acceptable as to theoretical adequacy and as to final outcomes ould be those of Bereiter and Engelmann (1966) and Blank and Solomon (1969). In the former program, groups of about five disadvantaged children are given highly structured exercises in five areas of deficiency. These exercises consist of instruction in "verbatim repitition, yes-no questions, location tasks, statement production, and deduction problems." In contrast, the Blank and Solomon program has a 1-to-1 pupil/tutor ratio, and the child acquires verbal facility through unstructured conversations with the tutor. As a result of these conversations, he learns to use verbal concepts such as "circle, across, top and bottom" to describe relationships among objects and events. The educator who wishes to decide upon which one of these

programs to utilize must choose between two approaches that appear to be equally effective in terms of improving cognitive skills, since both of these programs produce increases in IQ's of approximately twelve points. Therefore, his decision as to which one to select should be based upon considerations of all aspects of an educational program that are related to the quality and validity of its objectives, instructional methods, implementation procedures and outcomes.

The arguments which exist concerning sensori-motor versus conceptual training (Kohlberg, 1968; Stendler-Lavatelli, 1968) further illustrate a situation in which a comprehensive evaluation design will enable educators to make a rational decision as to which one of these different approaches to adopt. Kohlberg argues that sensori-motor training, based upon the Montessori method (1964), is one of the most effective methods for improving cognitive abilities. Therefore, the culturally disadvantaged children in his program learned to match stimuli in terms of sensory qualities such as color or size, arrange stimuli along dimensions of light-dark and big-small, and differentiate between dimensions of loudness and pitch. In contrast, Stendler-Lavatelli contends that training in certain reasoning operations which have been studied by Piaget (1969) will improve cognitive abilities. As a result of this position, Stendler-Lavatelli developed a program in which culturally disadvantaged children learned the reversibility, associativity, combinativity and identity operations. The present evaluation design will also enable educators to choose between these two differnt types of programs, since it provides criteria for selecting remedial techniques that are most appropriate for promoting either sensori-motor or conceptual development.

SUBSTANTIVE ACTIVITIES INVOLVED IN CONSTRUCTING THIS DESIGN

The design for evaluating educational programs for culturally disadvantaged children was developed in the following manner in order to assure the adequacy of the theoretical basis of such programs:

Reference System for Evaluating Educational Objectives. The Piaget and Inhelder Taxonomy of Human Development (1969) was used as a basis for describing the major cognitive, social and affective processes of normal children, and this particular taxonomy was utilized because it is the most comprehensive one available. Subsequently, the literature on the deficiencies of culturally disadvantaged was studied, and this information was organized in a manner that was congruent with the psychological scheme of human development, i.e., deficiencies associated with cognitive, social and affective processes were classified according to whether they developed in the sensori-motor or concrete operations period, and in terms of their specific relationship to the developmental processes described by Piaget and Inhelder. Evaluative criteria for examining objectives were then derived from the taxonomies of human development and deficiencies by using these taxonomies to identify important developmental processes which should be concentrated upon, and these criteria evaluate educational objectives in terms of their underlying theoretical rationale, developmental significance, appropriateness for alleviating known deficiencies, comprehensiveness and the developmental sequence in which they are arranged.

Reference System for Evaluating Instructional Methods. The instructional methods utilized by exemplary educational programs evaluated by Hawkridge, et al. (1968a) were identified and arranged into a Taxonomy of Instructional Methods. This taxonomy was congruent with the scheme of human development described by Piaget and Inhelder (i.e., the instructional methods were classified according to the areas of development that they concentrated upon), and it included methods utilized by both preschool and primary school programs. The second group of evaluation criteria was based upon the Taxonomy of Instructional Methods, and the developed criteria specify that a current program's instructional methods should be similar to those described in the Taxonomy, comprehensive and arranged into a sequence which is parallel to the sequence of human development.

Reference System for Evaluating Implementation Procedures. The evaluation studies of Hawkridge, et al. (1968a, b) also provided information about the implementation procedures of exemplary programs, and this information was arranged into an inventory of the "characteristics of superior implementation procedures" that described activities such as teacher training methods and the intensity of instruction associated with these programs. The criteria for the evaluation of implementation procedures were set up to examine the six activities which were described as being the characteristics of superior implementation procedures. Thus, by utilizing these evaluation scales it can be determined whether a program was implemented in a manner congruent with some of the most outstanding programs in the nation; and the following activities were evaluated: the amount of planning in selecting personnel and organizing a program, teacher training methods, relevance of

instruction and materials to objectives, intensity of treatment, size of instructional groups and amount of parental involvement.

Reference System for Evaluating Outcomes. Initially, two experimental designs described by Campbell and Stanley (1963) were identified that were highest in both internal and external validity. The "Solomon" and "Post-test Only Control Group" designs were then described in terms of their requirements for the random selection of subjects and the choice of separate control groups composed of both disadvantaged and middle-class children, and further descriptions were given of the types of extraneous variables which are controlled by the use of these designs.

This reference system also consisted of information concerning the use of appraisal techniques with disadvantaged children. Thus, a Taxonomy of Standardized Tests was set up which classified representative assessment devices according to the areas of development (parallel to Piaget and Inhelder) which they measure. These tests were selected from the exemplary programs evaluated by Hawkridge, et al. (1968a, b), and the national evaluation of Head Start which includes approximately sixty tests (ETS, 1968). In addition, guidelines for using these tests with disadvantaged children were also presented, and this part of the reference system included information (from Deutsch, et al., 1964) about methods for determining and improving upon the reliability of tests, and recommendations for upgrading the validity of test interpretation.

The final set of criteria (for outcomes) was derived from the descriptions of superior experimental designs, the Taxonomy of Standardized Tests and the guidelines for testing disadvantaged children. Initially, the evaluation scales for examining a program's experimental design were constructed from the descriptions of superior experimental designs, and the scales for rating appraisal techniques were established by reference to the information about selecting tests that are developmentally appropriate, and which have high amounts of reliability and validity. Thus, the evaluative instrument for this last phase of the design examines the internal and external validity of a program's experimental design, the relevance of the utilized tests to measuring outcomes predicted by objectives, the reliability and validity of tests which have been administered and the interpretation of test results.

THE UTILITY OF THIS EVALUATION DESIGN

If responsible critics (Bereiter & Engelmann, 1966; Jensen, 1969) of the literature of the culturally disadvantaged can be used for assumptions, this field of education is characterized by poor conceptualization, inadequate substance in the treatment processes,

and imprecise or irrelevant appraisals of progress or outcome. In these connections, the present design would appear to be significant in that it provides an organized and disciplined scheme for:

a. Conceptualizing programs. Information is provided concerning which educational objectives are most suitable for inclusion in educational programs, and this information is based upon the study of deficiencies which are associated with a comprehensive developmental scheme.

b. Delimiting research to the most significant and theoretically integrable dimensions, variables, etc. Only those educational methods which are based upon sound theoretical positions are considered. Therefore, this evaluation procedure encourages the use of methods which are firmly grounded in educational and developmental theory.

c. Providing a structure for organizing information in this area of education. This inquiry provides a means of classifying different programs in terms of educational goals, procedures, and outcomes. Hence, it aids in determining how programs are related to each other.

d. Providing for more rigorous evaluation. Procedures are developed for evaluating all phases of educational programs, ranging from an examination of the entering behavior of disadvantaged children through the testing procedures which are utilized to measure outcomes.

PROBLEMS AND ISSUES

The most rigorous applications of this design require the use of rating scales which can reliably and validly measure how closely each of the four sets of criteria have been attained. Therefore, the problem of developing such scales will be treated in the following section. In addition, statements will be made concerning the educational qualifications of evaluators and the time at which the evaluation process should begin, since these factors will significantly influence the evaluation process. The discussion of these factors concentrates upon how they can affect evaluation in the public schools.

Evaluation Scales. This design is intended to comprise a necessary and sufficiently contained evaluation model and to be an advance over available resources in the field. Further developments of this design for use by public school systems require the construction of psychometrically reliable and valid scale standards in order to increase its sensitivity to the substantive differences between programs. Such scale standards can provide the necessary means for measuring how closely each criterion is fulfilled, if they include clear statements concerning the necessary characteristics which a program must have to receive a particular rating. However, this level of clarity can only be attained through field testing procedures that will yield information as to whether different evaluators give similar ratings when using the evaluation scales. Obviously, an instrument that is interpreted in different ways must be modified to produce a high level of agreement, since this change will improve its reliability. Therefore, procedures

for determining the clarity of the developed evaluation scales must involve taking measures of inter-rater reliabilities, and if these reliabilities are low, they can probably be increased by more precisely describing the characteristics that are required for assigning a particular rating, and by training evaluators until they clearly understand the statements included in each evaluation scale.

The developed scales must also have high amounts of external validity, since this measurement will indicate that they can correctly predict whether a program will be successful or fail, if it is applied to similar settings across the nation. However, this type of validity can only be assessed after observing the outcomes of such applications, and then measuring the relationship between the original evaluation given to a program and the later results. Clearly, such measurements of external validity will indicate which scales should be eliminated from the evaluation scheme, and which ones can be used with the knowledge that they enable the evaluator to make valid generalizations.

In order to utilize this evaluation design as it is presently constructed, a five point rating scale can be utilized to determine how closely each criterion has been fulfilled, with a score of one representing the lowest possible rating. This rating device will only provide the most elementary information concerning a program's success in attaining each of the four sets of criteria. But, it will also serve as a beginning point for the empirical testing process essential to the development of a valid device.

Qualifications of Evaluators. Apparently, a direct relationship should exist between the educational background of the evaluator and his effectiveness in utilizing these particular criteria, i.e., an individual who has thoroughly studied developmental theories and research, experimental methods, and procedures for evaluating educational programs will probably apply this design more validly than will persons without training in these areas. However, even those individuals who lack specific training can use this design to improve their evaluation activities, if they study each of the four reference systems carefully, and closely follow the procedures for determining whether each criterion has been fulfilled.

It appears that most evaluators of programs for disadvantaged children are trained in school administration rather than in areas such as child development and statistics because administrative personnel are usually responsible for overseeing and reporting upon the results of the extensive number of federally sponsored programs operating in the public schools. Therefore, it must be assumed that this evaluation scheme will be most frequently used by school administrators who do not have the "ideal" educational qualifications for thoroughly understanding and applying each of the four reference systems. In order for them to complete the most valid evaluation studies, it is essential that they acquire some fundamental information about developmental theory, research methods and educational evaluation. In addition, further information can be obtained by studying the applications of this

design by professional evaluators who have extensive backgrounds in these academic areas, since these individuals will have the most thorough qualifications for making valid applications.

Proactive Evaluation. It appears that the model presented here can be most effectively utilized to evaluate a program before it actually begins operating, i.e., the descriptions of objectives, instructional methods, implementation procedures and techniques for measuring outcomes which are presented in a project proposal can be cautiously examined to determine whether major weaknesses exist in any of these phases of a program's development. Thus, proactive evaluation may eliminate the difficulties which occur in modifying instructional methods, etc., based upon the results of formative evaluation studies because such changes are extremely difficult to make during the operational stage.

In summary, it is proposed that rigorous scale standards be developed for each evaluative criterion, evaluators acquire some essential knowledge of developmental theories and research methodology before applying this design, and applications be made proactively in order to locate potential weaknesses which may cause a program to fail in achieving commendable objectives. Presumably, these improvements and uses will produce evaluative data which will significantly increase the quality of education for disadvantaged children.

REFERENCES

Bereiter, ., & Engelmann, S. Teaching Disadvantaged Children in the Preschool. Englewood Cliffs: Prentice-Hall, 1966.

Blank, M., & Solomon, F. How shall the disadvantaged child be taught? Child Development, 1969, 40, 47-61.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally & Company, 1963. Pp. 171-246.

Deutsch, M., Fishman, J. A., Kogan, L., North, R., & Whiteman, M. Guidelines for testing minority group children. The Journal of Social Issues, 1964, 20, (No. 2), 129-145.

Educational Testing Service. Disadvantaged Children and their First School Experiences, ETS-OEO Longitudinal Study: Theoretical Considerations and Measurement Strategies, Appendices Related to Measures. ETS, 1968, PR-68-4.

Hawkridge, D. G., Chalupsky, A. B., & Roberts, A. O. A Study of Selected Exemplary Programs for the Education of Disadvantaged Children. Parts I-II, Final Report. Palo Alto, Calif.: American Institutes for Research, 1968a. (Distributed by National Cash Register Corp. for ERIC: Bethesda, Maryland. ED 023 776-023-777)

Hawkridge, D. G., Tallmadge, G. K., & Larsen, J. K. Foundations for Success in Educating Disadvantaged Children. Final Report. Palo Alto, Calif.: American Institutes for Research, 1968b. (Distributed by National Cash Register Corp. for ERIC: Bethesda, Maryland. ED 037 591)

Jensen, A. R. How much can we boost IQ and scholastic achievement? Harvard Educational Review, 1969, 39, 1-123.

Kohlberg, L. Montessori with the culturally disadvantaged: A cognitive-developmental interpretation and some research findings. In R. Hess & R. Bear (Eds.), Early Education: Current Theory, Research, and Practice. Chicago: Aldine, 1968. Pp. 105-118.

Montessori, M. The Montessori Method. New York: Schocken, 1964.

Piaget, J., & Inhelder, B. The Psychology of the Child. New York: Basic Books, 1969.

Stendler-Lavatelli,C. Environmental intervention in infancy and early childhood. In M. Deutsch, et al. (Eds.), Social Class, Race and Psychological Development. New York: Holt, Rinehart & Winston, 1968. Pp. 347-380.