R E P O R T R E S U M E S

ED 019 896 FL 000 430

EVALUATION AND TESTING OF FLES PROGRAMS.

BY- CARTON, AARON S.
NEW YORK UNIV., N.Y., SCH. OF EDUCATION

PUB DATE 18 OCT 61

EDRS PRICE MF-$0.25 HC-$0.60 13P.

DESCRIPTORS- *FLES PROGRAMS, *PROGRAM EVALUATION, *CURRICULUM EVALUATION, *RESEARCH METHODOLOGY, *EVALUATION METHODS, ARTICULATION (PROGRAM), LANGUAGE RESEARCH, RESEARCH CRITERIA, TESTING METHODS, TESTING PROBLEMS, AUDIOLINGUAL SKILLS,

AN INTRODUCTION WHICH STATES THE PURPOSES OF AN EVALUATION OF FLES PROGRAMS AND STRESSES THE DIFFERENCE BETWEEN VALUE JUDGMENT AND EVALUATION IS FOLLOWED BY A DISCUSSION OF THE RELATION BETWEEN EVALUATION AND CURRICULUM. TEST AND MEASUREMENT PROCEDURES FOR EVALUATING FLES PROGRAMS ARE THEN SUGGESTED. A CONCLUDING SECTION INDICATES THE DESIGN OF A SPECIFIC EVALUATION PROJECT CONCERNED WITH TWO OBJECTIVES--(1) TO DETERMINE WHETHER FLES SHOULD BE GIVEN ON A SELECTIVE OR A UNIVERSAL BASIS IN THE SCHOOL SYSTEM, AND (2) TO ASCERTAIN THE EFFECT OF THE FLES PROGRAM ON JUNIOR HIGH SCHOOL FOREIGN LANGUAGE LEARNING. THIS SPEECH WAS DELIVERED AT THE ANNUAL MEETING OF THE NEW YORK STATE TEACHING ASSOCIATION, SOUTH NASSAU ZONE (116TH, OCTOBER 18, 1961). (AR)

EVALUATION AND TESTING OF FLES PROGRAMS

TEXT OF A PAPER GIVEN AT

116th ANNUAL MEETING OF THE NEW YORK STATE TEACHING ASSOCIATION

SOUTH NASSAU ZONE, OCTOBER 18, 1961

Aaron S. Carton

Experimental Teaching Center

School of Education

New York University

Ladies and gentlemen:

The function of an evaluation project attached to a program that is essentially new to education is not readily understood. I am pleased to have the opportunity today to try to explain this function to myself. I am pleased, indeed, to have you listen while I speak to myself. And I look forward eagerly to your comments. They shall constitute the high point of my visit here.

In my part of this conversation, I shall first spend some time describing what I see to be the purpose of evaluation. Under this heading we shall discuss what is evaluated. This topic will lead us to a discussion of the relation between curriculum and evaluation. The discussion of "course content" takes us to a description of the "tests and measurements" appropriate for the particular evaluation we are conducting. The design of our evluation project is also to be described here. The logical conclusion to our talk would be a discussion of results at one or several programs of foreign languages in the elementary schools (FLES). We shall not, however, conclude our discussion today. We are not prepared to discuss results beyond giving a sketchy description of what we guess is taking place. In the course of the discussion it will become clear that it is necessary for at least some years to elapse before results can be published.

The Function of Evaluation

Dr. Childers has given us a description of a burgeoning educational movement. Shall we evaluate the validity of this movement? The answer is emphatically: No! The movement represents a stage of growth in our educational philosophy. As Americans, are we becoming aware of the place of foreign language study in the school curriculum as we once had to learn

to understand the place of social studies or science. Essentially we feel today that being able to speak to someone in a foreign language is about to become as frequent a necessity as counting the correct change in the grocery store. Thus the school systems have the responsibility of teaching foreign languages as thoroughly and efficiently as possible. The FLES movement is a reflection of that feeling.

It becomes our task, however, to try to learn exactly what it is that the children learn in their FLES classes. It is the responsibility of an evaluation project to describe what is learned and to find ways of comparing what is learned in one program of instruction with what is learned in another program. The question: Is foreign language study good? is a question of value. It is not a question for evaluation.

Because a description for making educational decisions is useful only if it is a perfectly objective one, the research worker may approach a problem from a point of view opposite to the one he subscribes to. With a "mock-negativistic" frame of mind, it is more likely that an evaluator will think up more relevant questions that ought to be checked than he might think up if he allowed his personal enthusiasms to guide him. It is the job of a research worker in charge of an evaluation project to be skeptical. Without skepticism, the findings of any study can turn out to be perfectly useless when subjected to further scrutiny or when these findings are tested against the real hard outcomes of an educational decision which was based on the study. In the remainder of my talk, I may express some skepticism. I hope, however, that this skepticism will be taken for no more than what it is -- the assumption of the position I feel obligated to assume, if my work as an evaluator is to be of service to you.

What do the children learn in a FLES program? We know that that will depend, first and foremost, on what they have been taught. Next, what children learn might depend on factors such as how old the children are; on how smart and skilled they are in learning; on how much they are willing to apply themselves; and so on. The final important set of considerations which determine what the children learn might be the choice of the method of instruction; the skill of the teacher in executing this method; the appropriateness of the method to the particular children being taught and to the curriculum or objectives that are being taught; and so on.

On a most simple level, in order for the evaluator to determine what and how much the children have learned, he might merely prepare tests to determine whether the children have mastered the curriculum. But such a testing project would fail to tell us what the children <u>can</u> learn. Nor would we discover which children learn best. Nor would we learn anything about how they ought best be taught. In order to get information on these considerations we should have to have data on many different curricula, on many different types of instruction, and on children of various ages and "brightnesses." Thus, if the work of the evaluator is to have some general value, it should be concerned with as many of these considerations he can possibly get information on in any given evaluation project.

Dr. Childers brought to our attention a study of FLES instruction which complained that the movement represented too many approaches to be adequately evaluated. The criticisms on planning and lack of adequately organized course progressions are well taken. But I would hardly agree that a diversity of approaches -- if they are intelligently planned by people who know something about language, linguistics and children -- is

a matter to complain about. I agree wholeheartedly with the position Dr. Childers takes. I hold that until we know for certain that we are applying the best possible curricula at the best possible time to precisely the group of pupils we want to teach and until we know for certain that we are using the best possible methods, we shall benefit from a diversity of approaches. Diversity represents experimentation. Since it is, even from the viewpoint of simple logic, impossible to know that we have the best possible solution to any problem, we shall always benefit from diversity and experiment. And the most successful and useful evaluation projects will be those which describe in most accurate detail the exact nature of the curriculum, the methods of teaching and the character of the pupils.

Relation Between the Curriculum and Evaluation

An educational program is successful if the pupils who are subjected to it learn to be able to do what the program sets out to make them able to do. In the case of FLES programs, the most widely accepted objectives are the development in the children of an ability to understand the spoken foreign language and to speak it. A test is said to be a valid measure of what it is intended to measure, if the scores of individuals on a test are actually predicators of the extent to which the individuals are able to perform the skill they are tested for in an actual life situation. Thus, in FLES programs we should seek to develop tests which measure the ability to understand and speak foreign languages.

The evaluator is charged with preparing tests to measure what the program is supposed to teach. But, if the description of what the program is

supposed to teach is limited to such general terms as "the ability to understand the spoken language and to speak it" the evaluation process bogs down. The reality of the situation remains that the pupils in a FLES program will not -- at any given time during the program or immediately after it -- be able to understand or say much more than they have been taught. Thus in addition to needing to know a curriculum in order to evaluate it in terms of the outcomes of <u>that</u> curriculum we see that another concern of the evaluator with the curriculum occurs in the technical process of preparing tests.

Now, if FLES programs are to have a diversity of curricula, it follows that evaluation projects attached to them must have a diversity of different tests. If, however, we are to compare curricula and programs with each other, we must discover some common dimensions along which performance can be measured. I would submit that all foreign language courses consist of two major categories of knowledge. These categories run parallel to the nature of language. The first category is the "structure" of the language. The second is the lexicon or vocabulary. The term structure, as used here, is almost synonomous with the term grammar. Were I to teach the structure of a language -- and, remember, in FLES this can be done by means of pattern practice without any reference whatsoever to the fact that we are doing it -- I would be concerned

(a) with how to build sentences and express grammatical relationships by means of sentences in the language, that is, the "syntax of the language;"

(b) with the way a language builds its words and how grammatical relationships are expressed in words, that is, the "morphology of the language;" and

(c) with the phonetics of the language and with the system of sound contrasts, that is, the "phonemics of the language."

Please note that I have inverted the usual sequence of descriptions of these elements of language. The implication is merely that there is nothing significant about the usual sequence of treating with phonetics first. In audio-lingual foreign language instruction, syntax, morphology, and phonemics are taught concomitantly. The examples used for pattern practice contain elements of each and there is no way, nor is there any reason, for assigning precedence to one or to another element.

With the understanding of the nature of language, as consisting of structure and lexicon, I am in a position to construct tests for whatever curriculum or whatever program of instruction I am asked to evaluate. The way I might proceed is as follows: First I must ask the teaching personnel of the program for a body of data. These data are obtained with the following series of questions:

(a) What vocabulary items were taught in your program?

(b) What kinds of sentences do you expect your pupils to be able to understand and to produce? Examples of more detailed questions belonging to this category are:

Can they comprehend and ask questions?

Can they understand and issue commands?

Can they understand a statement indicating where something is?

Can they indicate the location of something?

Can they make a statement indicating that someone or something has acted on another someone or something?

(This last example is merely a subtle paraphrase of: Can they make up sentences which contain transitive verbs?)

(c) What kinds of alterations should the pupils be able to make on words? Examples of detailed questions in this category are: Can they distinguish between singulars and plurals? Can they distinguish among and indicate the persons and tenses of a verb correctly? Can they recognize alterations and can they produce the appropriate alterations on a noun (in the case of an inflected language?)

(d) Have the pupils mastered the phonetics of the language? The detailed questions under this heading are: Can the pupils recognize and produce the crucial phonemic distinctions in a language? Can they produce all or most of the sounds of the language in a manner similar to a native speaker?

The preparation of tests to discover whether or not a pupil can do the kinds of things the program intended he should be able to do is the next step in the process of evaluation. This is a task which requires a good deal of ingenuity and many teachers and test-makers have devised a large array of interesting item-types. It is not my intention to go into the technical aspects of test production today, although we will be concerned with some tests later.

The point I wish to make here is this: If I can prepare tests which indicate the extent to which the pupils have a mastery of the underlying structure of a language, I think I am in a position to state whether they have acquired the basic skills of understanding and speaking the language. The particular vocabulary used, is the vehicle by which the structure is taught and by which it is tested. The degree of mastery of a language for a single individual can be stated as a fraction of the total number of structural components of the language to the total any individual can handle. Programs can be compared with each other by noting which structural com-

ponents each program emphasizes and by studying the rate with which each program approaches a complete mastery of the structural features of that language.

Design of an Evaluation Project

The evaluation project we have been concerned with has two objectives. First it is concerned with whether FLES should be given on a selective or on a universal basis in the school system we are concerned with. The second is to ascertain the affect of the FLES program on the junior high school foreign language learning.

The dilemma of who should receive FLES instruction is a common one. It will be recalled that the school systems which reported that they already had foreign languages in the grades at the White House Conference in 1953 reported programs which were aimed mainly at the gifted. With American educational philosophy being what it is, the notion that all children had a right to learn a foreign language became generally attached to the FLES movement as it began to achieve national proportions. But school administrators were faced with the problem of whether to give foreign language instruction, which to many seemed to be an adjunct and a frill, to children who had not achieved adequate mastery of certain other more crucial areas of the school curriculum, such as arithmetic and reading. Shall FLES be given only to those children whose achievement in other school areas is deemed satisfactory enough to add another area of study? But what if these are not the children who are most gifted in learning with the audio-lingual approach? Suppose some of the children who are duller in other subjects have a knack for learning a foreign language and can be helped to learn in other areas, if they are allowed to develop this knack? On the other hand, what shall we do with the otherwise able student who has no talent for foreign language learning?

In many of its aspects, the problem is one of values, of philosophy, and of ethical dilemmas. But in order to make a sensible resolution of the conflict, the administrator needs to know something about the foreign language achievement of children with various configurations of ability. In order to help unravel the nature of the relationships which may exist here, we are studying (in a FLES program which offers language instruction on a universal basis to fifth and sixth graders) the relationship between achievement in various school areas and foreign language learning, the relation between foreign language aptitude and degree of foreign language mastery, and the relation between foreign language aptitude and achievement in other school areas. In our study, we are using achievement scores which are available in the various other school subjects for the first of measures named above. We are using scores on the Elementary Form of the Carroll-Sapon Modern Language Aptitude Test for aptitude scores. (This aptitude test is a version, adapted for grade school, of the Modern Language Aptitude Test for adults. The elementary school form is in its experimental phases and a special arrangement with the authors of the test accounts for our use of it. This test is expected to be ready for general use in the reasonably near future.) The third set of measures, that is, the measures of achievement in foreign languages, will be discussed shortly.

We are interested in the effect of FLES programs on the junior high school level language learning (language learning level) first, because it should give us some data on the problems of articulation, and second, because it is here, that the actual value of a FLES program can begin to be seen. If the junior high school pupil does learn more foreign language as a result of having been exposed to a FLES program, we have a justification for FLES. And it is because this remains to be seen, that the result of this evaluation study

will not be meaningful until at least two years have elapsed. The question of shifting from audio-lingual to writing and grammar instruction becomes a problem of the junior high school teacher in a school system where there is FLES instruction. The problem is actually well known to many high school foreign language teachers who start their language instruction with the audio-lingual approach and who have prepared special materials for the transition to the writing and grammar approach.

Our attempt to discover what the effect of FLES is on junior high school foreign language learning includes tests of speaking and aural comprehension for 7th graders last year -- these are students who have had no prior FLES instruction -- and for 7th graders this year -- these are students who have had one year of FLES training last year. We are also studying their grades on the conventional translation and grammar tests.

I hold that the burden of testing is heavy on our children. I feel that it would be well to avoid testing in FLES programs where the children can learn without anxiety. Further, we tend to spend so much time testing our pupils, we are coming to a point at which we owe them more instruction. Thus, it is important to try to make the foreign language speaking and comprehension tests pleasurable and instructive. I do not know how well we have succeeded in doing this in our study. (It seems that the children did enjoy some of our tests.) The comprehension test used, consisted of arrays of pictures with foreign language statements for each array. The task of the child is to match up the foreign language statement with the correct picture. The speaking tests require tape recording. We have had to use sampling techniques in our study, because analysing tape recordings is a costly and time-consuming process. We have made tape recordings of approximately one per cent of the population in the FLES program we are

studying. These youngsters each responded to a standard question tape. The techniques for analysis require tedious efforts on the part of fluent and competent foreign language scorers.

A Short Concluding Word

FLES is today in the national limelight. It promises to make foreign language education successful in areas in which it has failed before. It promises to make us able to speak French when we get to France. But already a colleague in educational research challenges me to: "Say something in FLES." If the FLES movement fails to live up to the expectations some have for it, it is possible that foreign language education in general will suffer. The older foreign language teachers among us know how foreign language education in America has swayed in the breezes of political considerations and they, together with the younger ones among us, hope that those days are over.

We have a responsibility today to help make the FLES movement achieve its goals. The fact that the children seem to learn better because they are young will help the foreign language teacher achieve more success in the areas of understanding and speaking than he has before. But continued success depends on the development of better and better methods. It requires the imagination of the teacher and the psychologist to improve methods and curricula. And the skeptical attitude of the research worker -- be he a psychologist, or a teacher -- is required to make sure that our innovations are meaningful. The research worker, however, does not often conduct the teaching. He might lose some of his valuable skepticism if he did. Besides he frequently has little time for anything but writing designs and conducting data analyses, if he is skilled at these things. In order to

devise techniques for analyzing the problems teachers have in their attempts to improve their methods, teachers must communicate with interested research workers and tell them what those problems are. At this point, it is appropriate to relinquish the floor and listen to your part of the conversation.