DOCUMENT RESUME

ED 091 448 TM 003 666

AUTHOR Breland, Hunter M.
TITLE Planning for the Evaluation of the PLATO Computer-Based Instructional System: Attitudinal and Performance Evaluation in the Elementary Schools.
PUB DATE [Apr 74]
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research association (59th, Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Achievement; *Computer Assisted Instruction; Mathematics Instruction; *Program Attitudes; *Program Evaluation; Reading Instruction
IDENTIFIERS PLATO; *Programmed Logic for Automatic Operations

ABSTRACT

The planning for the evaluation of achievement and attitudes in Programmed Logic for Automatic Operations (PLATO) elementary school reading and math programs is discussed. To measure the achievement of the students in the program, standardized tests will be used along with special tests custom-tailored to the PLATO reading and math programs. Pupils will also be observed on a one-to-one basis. To assess the students' attitudes toward the program, group tests will be administered as a background to a more direct approach of asking individual children a standard set of questions about their feelings toward the program. Finally, in the evaluation of the PLATO program there is no control group. However, there is a need to observe the differences that occur in a naturalistic setting. Consequently, four types of comparisons among different groups are planned, and it is believed that arguments concerning the effects of PLATO can be constructed. (Author/MLP)

Session #26.12

PLANNING FOR THE EVALUATION OF THE PLATO COMPUTER-BASED INSTRUCTIONAL SYSTEM: ATTITUDINAL AND PERFORMANCE EVALUATION IN THE ELEMENTARY SCHOOLS

Hunter M. Breland

Educational Testing Service

Paper presented at the Annual Meeting of the

American Educational Research Association

Chicago, Illinois

April 19, 1974

ATTITUDINAL AND PERFORMANCE EVALUATION IN THE ELEMENTARY SCHOOLS

Hunter M. Breland

Educational Testing Service

Although the title of this paper--a title created last August--begins with "Attitudinal," I will reverse the order and begin with a discussion of the performance evaluation for this session. I want to outline some of the difficulties which have been encountered by the ETS evaluation team in planning for the evaluation of achievement and attitudes in the PLATO elementary school reading and math programs. I will organize the discussion in terms of three major headings: achievement measurement, attitude measurement, and the problem of control.

Achievement Measurement

Because traditional methods of measuring achievement--i.e., standardized tests--are currently undergoing some uncertainties, those engaged in the evaluation of achievement face the necessity of considering other approaches. One of these alternative approaches is what is called content-referenced, objective-referenced, criterion-referenced, crijective-referenced--or whatever--testing. Some believe that we are in a transition phase from a preoccupation with prediction to more of a concern for diagnosis (Mercer, 1974). To the evaluator, the situation appears less like a state of transition than like a state of limbo. For one finds that acceptance and application of the newer methods is not sufficient--one must <u>also</u> retain the old. And the difficulties do not end there. Critics of all types of pencil-and-paper instruments insist on

still another additional procedure in which the evaluator interacts on a one-to-one basis with each of a number of pupils individually. This procedure is sometimes called that of the work sample. And yet if the evaluator agrees that the work sample is a good procedure and should perhaps replace either the standard[illegible]esting or the content-referenced testing--or both--the cry is jus[illegible]. The evaluator thus faces the risk of overextending his resources in dealing with the issue of achievement alone, when a number of other issues deserve attention as well. In fact, it has become fashionable to disparage the importance of achievement in educational program evaluation. The National Science Foundation, however, has indicated that achievement is a central concern in the PLATO evaluation.

Even though traditional standardized tests will be used in the evaluation, their limitations are recognized. With respect to the PLATO program, in particular, a major disadvantage of standardized tests relates to their content coverage. The PLATO elementary school mathematics program--especially--covers material not ordinarily a part of mathematics instruction in the 4 to 6 grade range. Accordingly, standardized tests applicable to this grade range are not altogether appropriate for the evaluation of the PLATO mathematics program. In order to be widely applicable, standardized tests must be fairly general in nature to insure their appropriateness to many different types of instructional programs. Yet as Hartnett (1971), and others, have noted, making a test broadly applicable may also make it insensitive to important program differences. With reference to standardized tests, Shoemaker (1972) has noted that "such a test is not likely to contain both the breadth and depth of coverage necessary to make a detailed assessment of any instructional program." Such insensitivity may lead to conclusions of no difference among

programs that do have distinctly different characteristics.

Nevertheless, standardized tests have certain merits that are often overlooked. These tests have usually been carefully constructed, they are accompanied by rigorous validity data, the level of difficulty is known, and they provide a basis (norms) of comparison with other programs and other times. Because of the widespread use of standardized tests, they can also be used for more specific comparisons beyond that of comparing norms (e.g., comparing specific item responses). Standardized tests are also useful for another purpose: whatever additional values a new program of instruction may have, there is a need to provide public assurance that elementary school students are achieving adequately in the areas of reading and mathematics as reflected by scores on standardized tests.

To compensate for some of the disadvantages of standardized tests, ETS is developing special tests custom-tailored to the PLATO reading and math programs. These tests are being developed in collaboration with PLATO lesson designers. Where instructional objectives have been specified in advance, the special tests will be geared to such objectives--especially in areas where the standardized tests have no coverage. In cases where objectives are apparent but have not been stated clearly by the PLATO lesson designers, objectives are inferred from observable lessons already developed, PLATO program documentation, and from discussions with lesson designers.

Specific instruments to be used are:

Metropolitan Achievement Test - Primer. This widely-used standardized achievement test, developed by Harcourt Brace Jovanovich, Inc., is to be used as one means of assessing reading skills in the K to grade 1 range. Since Part 3 of the MA Primer relates to quantitative skills, only Part 1 (Listening for Sounds) and Part 2 (Reading) are to be used.

Comprehensive Test of Basic Skills. This CTB/McGraw-Hill instrument has been chosen as the standardized test in mathematics because of its emphasis on computational skills, its use as a regular part of the Urbana, Illinois, school district evaluation program, and its widespread national usage. Only that portion of the CTBS relating to mathematics, however, is to be used. Level 2 of the CTBS measures mathematics achievement for the 4 to 6 grade range--the same range targeted in the PLATO elementary school mathematics program.

Special Reading Test. This instrument is being developed by ETS with the assistance of the PLATO elementary school reading staff.

Special Mathematics Test. This instrument is being developed by ETS with the assistance of the PLATO elementary school math staff.

The above described instruments will, of course, not answer all the concerns about achievement. They will not cover all content that might be desired, they are not completely adequate for the measurement of change, they have all the limitations of the pencil-and-paper format, and there are a number of problems in transfering from PLATO display format to a format equitable for non-PLATO comparison groups. To alleviate some of these difficulties, pupils will be observed on a one-to-one basis through the use of work samples to enrich and corroborate evidence obtained from the standardized and special instrumentation.

Attitude Measurement

While a number of instruments exist for probing attitudes toward school subjects, all have the problems of response sets, replicability, reliability, and others sources of inaccuracy. And the younger the children are the more severe the measurement problem. For these reasons, no published instruments have been found that are adequate for the task of assessing pupil attitudes in the PLATO elementary school evaluation. Other ETS projects now in progress, however, have experienced some success with paper-and-pencil instruments for young children and these--or modified versions of them--are to be used in the PLATO evaluation. For the youngest children, those in the PLATO reading program, no attempt will be made to assess attitude toward reading specifically, but only to assess undifferentiated attitude toward school. For the older 4 to 6 grade children in the mathematics program, the instruments will differentiate among attitudes toward mathematics, attitudes toward reading, and attitudes toward PLATO and computer-based instruction in general.

These group-administered instruments will be considered as a background to a more direct approach: that of asking individual children a standard set of questions such as how they feel about what they are learning, what problems they are encountering with PLATO, which things they prefer to learn from the terminal and which from their teacher, and whether they would rather make a mistake on the terminal or with the teacher. These standard "chats" will be conducted by observers after having obserbed the child during his turn at the terminal. Extreme care is being made to design the interaction so as to minimize influencing the child's attitude.

The Problem of Control

Although no attempt is being made in the evaluation to conduct an experimental procedure with control groups, random assignment, etc., there is a need to observe differences as occur in a naturalistic setting. To provide for arguments more convincing than mere testimonials will require that comparisons be made among groups and that the results of those comparisons be attributable to PLATO. Otherwise, the demonstrations will fail to demonstrate since any claim for the utility of PLATO as an educational device could be easily attributable to causes from other parts of the system of events under consideration. Failure to recognize fundamental differences among groups of pupils being compared, differences in schools, differences in teachers, and differences in home activities would lead to so many challenges of the outcomes that the demonstrations would have little chance of success.

One type of comparison planned is that of differences between a baseline and later performances. That is, if the demonstrations begin on schedule in September, 1974, and continue to May, 1976, then observations made in May, 1974, may be compared with those made in later years at approximately the same time. The pupils being compared, of course, are not the same ones--but similar students of the same age, grade, intelligence, and home background. One of the difficulties encountered in the [illegible]ent evaluation, however, is that it is not known for certain in what classes and schools the demonstration is to be conducted. Even though a pool of volunteer teachers has been assembled by the PLATO developers in the Champaign/Urbana area, there is no guarantee that enough terminals will be available for all of these teachers. Nor is it

certain that all of these teachers will continue their interest in the program. Additionally, there exists the possibility that some of the teachers may prove to be unsuitable for the demonstrations. Thus, the evaluator is faced with the problem of collecting baseline information in schools and classes which may not be involved in the evaluation. Since the collection of data necessarily requires some effort on the part of schools and teachers, the possibility exists of offending such schools and teachers to a degree that they will not participate in future data collections.

A second type of comparison that is planned is what Campbell & Stanley (1968) called the "non-equivalent control group comparison." We would perhaps not call the comparison groups control groups even though it is certain that they will be non-equivalent. Despite these difficulties, it is believed that useful arguments can be derived from such comparisons--as long as they are used in concert with other, distinctly different comparisons. Accordingly, attempts are being made to locate teachers and classes, not using PLATO, in which instrumentation and other observations may be made so as to compare events in these classes with those occuring in the PLATO demonstration classes. Since the developer has made no effort to obtain these comparison classes, the burden has fallen upon the evaluator. And since the implementation of the PLATO terminals is being planned so that all pupils in a school will have access to PLATO (even though they are not in official demonstration classes), the comparison classes will have to be obtained outside of schools participating in the demonstrations. Because of the pervasive impact of PLATO throughout the entire Champaign/Urbana area, it may be difficult to locate appropriate comparison classes within this area.

A third type of comparison that is possible is what might be called a static-group comparison. That is, at some future time--say near the end of the PLATO demonstrations--classes will exist in which there are students of varying degree of experience with PLATO. There is a large turnover in the schools and some of the pupils will have only recently arrived in the area. Accordingly, these new students will have experienced very little of PLATO--but other students in the same class will have had PLATO instruction for possibly two years. And within the group of experienced pupils differing degrees of exposure will be ascertainable through automatically recorded data on time spent at terminals. One would hypothesize that pupils receiving the most exposure to PLATO will have learned the most about specific material presented on PLATO. Also of interest will be comparisons of the new and old students with respect to material non-specific to PLATO but which is a standard part of school instruction.

A fourth type of comparison will be that with other computer-based programs currently in progress. Each day over 50,000 elementary students sit down at a computer terminal and receive 10-20 minutes of instruction in reading, mathematics, and other areas. To the degree that comparative data can be obtained from these other programs useful analyses can be made.

Through these multiple sets of comparisons it is believed that convincing arguments concerning the effects of PLATO can be constructed. To guard against the possibility of evaluating a "non-event" as described by Charters & Jones (1973), however, careful descriptions of the process of implementation will be made as described in the next presentation.

References

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally & Company, 1968.

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

Hartnett, R. T. Accountability in higher education: A consideration of some of the problems of assessing college impacts. New York: College Entrance Examination Board, 1971.

Shoemaker, D. M. Evaluating the effectiveness of competing instructional programs. Educational Researcher, May 1972, 5-8.