DOCUMENT RESUME

ED 092 565 TM 003 667

AUTHOR Alderman, Donald L.
TITLE Planning for the Evaluation of the PLATO and TICCIT Computer-Based Instructional Systems: The Comparison of Performance for Community College Students.
PUB DATE [Apr 74]
NOTE 11p.; 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 *Academic Achievement; *Community Colleges; Comparative Analysis; *Computer Assisted Instruction; Conventional Instruction; Educational Improvement; Educational Innovation; Educational Needs; *Evaluation Methods; Instructional Technology; *Program Effectiveness; Program Evaluation
IDENTIFIERS PLATO; Programed Logic for Automatic Operations; TICCIT; Time Shared Interactive Computer

ABSTRACT

Plans for the evaluation of the Time Shared Interactive Computer-Controlled Information Television System (TICCIT) and Programed Logic for Automatic Operations (PLATO) are discussed. These projects represent attempts to implement innovative curricular systems and to demonstrate the effectiveness of computer assisted instruction. Plans for the evaluation of PLATO and TICCIT emphasize the measurement of student achievement to reflect program accomplishments. The evaluation of the systems will be in the form of a comparative field study concerned with the educational impact of TICCIT and PLATO. Results from instruction with the computer curricular materials will be compared with the results from conventional instruction. (Author/BB)

Planning for the Evaluation of the PLATO and TICCIT Computer-Based Instructional Systems: The comparison of performance for community college students.

Donald L. Alderman

Educational Testing Service

Paper presented at the Annual Meeting of the
American Educational Research Association
Chicago, Illinois

April 19, 1974

Planning for the Evaluation of the PLATO and TICCIT Computer-Based Instructional Systems: The comparison of performance for community college students.

Donald L. Alderman

Educational Testing Service

Introduction

The TICCIT and PLATO projects represent ambitious attempts to implement innovative curricular systems and to demonstrate the effectiveness of computer-assisted instruction. The TICCIT demonstration entails extensive computer use to support introductory mathematics and English courses at two community colleges (Bunderson, 1973; MITRE Corp., 1974). The PLATO program involves the application of computer-based education as an integral component of community college courses in several subject areas (for a general description, see Alpert & Bitzer, 1970). These projects provide us with an unusual opportunity to document the impact of computer-assisted instruction on a wide range of college services.

As projects targeted toward actual trials in educational institutions, TICCIT and PLATO necessitate an evaluation which gives particular attention to their context. That context influences the projects' progress and carries significant implications for the design of an appropriate educational study. Despite rapid growth over the last decade and certain pockets of computer experience, community colleges remain a neglected sector of higher education with respect to computer use for instruction (Miller, 1972). Yet the need to serve a diversified student body and adapt to changing demands makes the community college a prime candidate for the use of instructional technology (Bushnell, 1973). The evaluation for TICCIT and PLATO must consider the scope of potential effects within the community

college (Kerr, 1972), as well as the particular position of those colleges participating in the demonstrations.

While plans for the educational evaluation call for a broad investigation to determine the respective advantages and disadvantages of TICCIT and PLATO (Alderman & Mahler, 1973), student performance deserves special emphasis for its importance to the colleges and for its relation to project goals. This category of information includes both achievement and conduct (e.g., attrition, time to completion) as measures of program results. The apparent simplicity of an evaluative focus on student performance to convey results is deceptive. For our obligation is to report on the respective positive and negative aspects of two specific computer-assisted instructional systems, rather than label either demonstration a success or failure. In recognition of the colleges' responsibilities, the evaluation must also avoid interfering with college operations and services during the course of the demonstrations. So the measurement of student performance and the control of factors which might affect that performance pose critical problems to the evaluator.

Measurement

Perhaps the first severe problem that arises in assessing program results through student achievement is the choice of an appropriate instrument. Of course, this choice depends on the program under study and the evaluation's design. In order to determine TICCIT and PLATO's impact at participating colleges, the evaluation relies on results from regular classes without computer use as a relative standard for the comparison of program accomplishments. Therefore the instrument chosen must be appropriate for the curriculums supported by the respective computer systems as well as the coverage in other classes without

computer use. This approach to the evaluation as a comparative field study is consistent with a view of the projects as demonstrations of alternative instructional methods (see Glass, 1972).

At least three major options exist in selecting an instrument that reflects student achievement. Due to their validity and widespread administration across situations, standardized tests offer a convenient solution for obtaining measures of results. But such a test is often insensitive to the particular program under study in that its match with the curricular coverage omits much detail (Shoemaker, 1972) and its aim is only to assess a student's knowledge rather than a program's results. Criterion-referenced tests present an alternative that offers direct correspondence with the curriculum. However, this mode of testing implies a firm grasp of what constitutes mastery, a concept that still is subject to numerous interpretations (Stake, 1970), and bypasses measuring the extent to which material is learned (Ebel, 1973). A third tool with which to capture the dimensions of student achievement is special tests developed for the specific program's evaluation. Given the requisite experience in test construction available for the TICCIT and PLATO evaluations, this option does appear to insure proper instrumentation for the programs since it would allow interaction with project developers and college faculty to determine appropriate test composition.

Of course tests tailored to the instructional programs cannot alone guarantee intended comparisons and avoid misrepresentation of results. It is important to recognize TICCIT and PLATO as separate projects with distinctive features and different premises for computer application to instruction. Their respective demonstrations entail field trials at different sites under different conditions

rather than a competition or comparison between two computer-assisted instructional systems. Therefore, the development of special achievement measures depends on the particular system's courseware (computer-supported curricular materials) and the instruction given in regular classes of the same courses which employ computer materials.

A further delineation of test items according to objectives held by the instructional treatments enables us to compare results in terms of both common and unique attributes (Popham, 1969; Shoemaker, 1972). Since the participating community colleges play a critical role in the design of courseware and retain the authority to decide on matters of curriculum integration, there is substantial overlap between objectives from computer materials and those from regular courses in both projects. With this basis for the analysis of test results, it becomes possible to compare students' achievement under dissimilar instructional programs in terms of subscores for (a) common objectives, (b) objectives unique to the computer curricular materials, (c) objectives unique to regular classes, and (d) transfer tasks beyond the scope of the instructional programs. In cases of great disparity between TICCIT or PLATO objectives and those of the regular instruction, it is still the evaluator's obligation to document project results with measures developed to reflect the particular objectives.

The process for developing program measures dependent on objectives is a sensitive component of instrumentation. It necessitates a thorough familiarity with the curriculums and extensive advice from developers and faculty. The evaluations for TICCIT and PLATO rely on project documentation, availability of curricular materials, and contacts with courseware developers for obtaining information about their objectives; conferences conducted at the participating

colleges provide us with information on the target courses' objectives as expressed by faculty. Such exchanges between the evaluator and participants in the demonstrations lead to preliminary pools of test items to be submitted to further review by developers and faculty, and to be given trial administrations before actual use. Toward the conclusion of the demonstration period the evaluation also plans to solicit faculty and developer ratings, on item importance and emphasis in instruction, for the tests used to assess program results.

Beyond measures of student achievement tailored to the particular curriculums, the tests serving as the basis for program comparisons must provide evidence for reporting the project's specific advantages and disadvantages. Since instructional programs often lead to different patterns of achievement (Walker & Schafferzick, 1974), it is appropriate to consider two schemes for item classification. The first concerns abilities, such as knowledge, comprehension, and application (see Bloom, 1956; Gagne, 1970), and gives attention to learner outcomes in terms of cognitive levels. A second structure for item categories relates to content and establishes subject matter topics within courses (Williams, 1973). Together ability and content classifications for test items afford greater detail on program results as well as categories with appeal to general interests outside the specific projects.

The evaluations for TICCIT and PLATO, then, include the use of special tests which reflect program objectives and also provide information on student achievement with regard to certain abilities and specific topical content. This instrumentation design represents our approach to the problem of measurement in a comparative field study.

Control

The distinction between an experiment and a demonstration suggests the evaluative constraints involved in the TICCIT and PLATO projects. The evaluation cannot impose strict experimental control which disrupts college operations and precludes the refinement of courseware; but it can strive to document the impact of computer-assisted instruction at an educational institution. In such a situation, a variety of factors might affect student performance besides the instructional programs under study. It is therefore important for the evaluation to attempt a balance in anticipated threats to validity, and to track a broad range of variables.

Since the evaluation plans require a relative standard for the comparison of achievement results, there are two student groups fundamental to the demonstrations. Students enrolled in course sections which make extensive use of TICCIT or PLATO curricular materials constitute the primary group under study. Students from other sections of the same courses form a second comparison group, in effect a control group, that receives regular instruction without computer dependence.

Since students must retain their perogative to choose courses and change sections, complete random assignment is not suitable for the demonstrations. Instead, the evaluator expects to arrange a quasi-random assignment of students registered for target courses to the two comparison groups. This implies that students would register for classes without knowledge of the particular instructional conditions or be given section assignments for a particular course on a random basis. While no method can totally compensate for a lack of randominzation, alternative procedures for identifying comparable student groups, such as matched samples, will be available (see Anastasio, 1972).

In addition to careful plans for establishing comparison groups, close attention to test administration can also enhance control over conditions irrelevant to the instructional programs and allow data collection which covers possible effects on learner achievement. Plans for test administration differ for the TICCIT and PLATO projects as a result of their dissimilar goals and implementation, but their respective evaluations will employ a combination of four test situations. These include:

1. Pretests. The pretests assess students' entrance skills relevant to the designated courses for the demonstration, and provide a baseline measure of student achievement related to course content. Pretest administration will take place during the registration period or first week of classes for an academic term. Students' scores on these instruments will give us a critical covariate in analysis as well as an indication of initial performance through which gains from instruction can be estimated.
2. Topical tests. Achievement measures targeted toward a particular component of a course enable the evaluation to focus on certain topics. Such topical tests place emphasis on those aspects of a course with extensive use of PLATO's adjunctive materials, and provide detailed information on specific units of TICCIT's "mainline" instruction. Furthermore, with variations in instructional sequence across course sections and in coverage across programs, these tests afford a flexibility in administration that allows for data collection on acquisition and on unique curricular components.
3. Posttests. Measures of overall program effects on student achievement will be obtained from posttests. These instruments cover an entire academic term of course work and, therefore, assess the

cumulative results of study under an instructional program. Scores on posttests as well as other tests designed to document instructional outcomes will follow the item classification schemes for objectives, abilities, and content explained above. Of course posttest administration will occur after students complete instruction in target courses.

4. Tests for retention. In order to gather information on retention and students' preparation for subsequent courses, the evaluation plans a fourth test situation. Tests for retention will be given to those students from the comparison groups who later enroll for successive courses in a sequence dependent on a target course.

This discussion of the evaluation plans clearly focuses on the measurement of student achievement to reflect program accomplishments. As a comparative field study concerned with the educational impact of TICCIT and PLATO, the evaluation will address results from instruction with the respective computer curricular materials in contrast with results from regular instruction. This contrast must extend beyond achievement measures. The innovative scope of the TICCIT and PLATO projects also holds the potential to affect aspects of student performance such as attrition, rate of learning, and study habits. Perhaps this potential to improve college services through effects on student conduct is as important as our traditional concern with achievement.

References

Alderman, D. L., & Mahler, W. A. The evaluation of PLATO and TICCIT: Educational analysis of the community college components. Educational Testing Service, Project Report 73-49, 1973.

Alpert, D., & Bitzer, D. L. Advances in computer-based education. Science, 1970, 167, 1582-1590.

Anastasio, E. J. Evaluation of the PLATO and TICCIT computer-based instructional systems -- A preliminary plan. Educational Testing Service, Project Report 72-19, 1972.

Bloom, B. S. Taxonomy of educational objectives, Handbook I: Cognitive domain. New York: McKay, 1956.

Bunderson, C. V. The TICCIT project: design strategy for educational innovation. ICUE Technical Report No. 4, Brigham Young University, 1973.

Bushnell, D. S. Organizing for change: new priorities for community colleges. New York: McGraw-Hill, 1973.

Ebel, R. L. Evaluation and educational objectives. Journal of Educational Measurement, 1973, 10(4), 273-279.

Gagne, R. M. The conditions of learning (2nd ed.). New York: Holt, Rinehart, and Winston, 1970.

Glass, G. V. Two generations of evaluation models. In P. A. Taylor and D. M. Cowley (Eds.), Readings in curriculum evaluation. Dubuque, Iowa: Wm C. Brown Company, 1972, pp. 103-109.

Kerr, C. The fourth revolution: Instructional technology in higher education. New York: McGraw-Hill, 1972.

Miller, L. W. How can two-year colleges use the computer? College Management, October, 1972, 30-32.

The MITRE Corporation. An overview of the TICCIT program. Washington: MITRE Corp., M74-1, 1974.

Popham, W. J. Program fair evaluation -- Summative assessment of instructional sequences with dissimilar objectives. National Society for Programmed Instruction, 1969, 8, 6-9.

Shoemaker, D. M. Evaluating the effectiveness of competing instructional programs. Educational Researcher, 1972, 1(5), 5-9, 12.

Stake, R. E. Objectives, priorities, and other judgment data. Review of Educational Research, 1970, 40, 181-212.

Walker, D. F., & Schaffarzick, J. Comparing curricula. Review of Educational Research, 1974, 44(1), 83-111.

Williams, J. D. Course-component evaluation. Part I: rationale and outline. Journal of Structural Learning, 1973, 4, 25-31.