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

ED 010 538 56

AN APPLICATION OF THE PRINCIPLES OF PROGRAMED INSTRUCTION TO A TELEVISED COURSE IN COLLEGE ECONOMICS.

BY- MCINTYRE, CHARLES J. AND OTHERS
ILLINOIS UNIV., URBANA

REPORT NUMBER BR-5-0841 PUB DATE APR 66
REPORT NUMBER NDEA-VIIA-1217
GRANT OEG-7-23-1020-248
EDRS PRICE MF-$0.09 HC-$1.44 36P.

DESCRIPTORS- *PROGRAMED INSTRUCTION, TELEVISED INSTRUCTION, *INSTRUCTIONAL IMPROVEMENT, COLLEGE INSTRUCTION, *ACADEMIC ACHIEVEMENT, *STUDENT ATTITUDES, COMPARATIVE ANALYSIS, PROGRAM IMPROVEMENT, *ECONOMICS, URBANA, ILLINOIS

TWO TECHNIQUES FOR REVISION OF INSTRUCTIONAL MATERIALS WERE STUDIED AND COMPARED--(1) THE USE OF SYSTEMATIC PROGRAMING PRINCIPLES AND (2) THE USE OF TRADITIONAL, MORE INTUITIVE, JUDGMENT PRINCIPLES. THESE TECHNIQUES WERE USED SEPARATELY TO REVISE A VIDEOTAPED COLLEGE LECTURE ON ECONOMICS. STUDENT SAMPLES WERE THEN USED TO EVALUATE THE TWO REVISIONS. PROGRAMED INSTRUCTION RESULTED IN BETTER STUDENT PERFORMANCE AND MORE FAVORABLE STUDENT ATTITUDES. (JH)

AN APPLICATION OF THE PRINCIPLES OF PROGRAMED INSTRUCTION TO A TELEVISED COURSE IN COLLEGE ECONOMICS

Principal Investigator: Charles J. McIntyre

Report Prepared by: Philip W. Tiemann
Donald W. Paden
Charles J. McIntyre

Title VII Project No. 5-0841
National Defense Education Act of 1958
Grant No. 7-23-1020-248

University of Illinois
Urbana, Illinois

April 1966

The Research Reported Herein Was
Supported by a Grant from the
U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education

ACKNOWLEDGMENTS

The research described in this report was made possible by the cooperation extended by the Department of Economics at the University of Illinois. Graduate teaching assistants from this department who contributed to the study include: Mrs. Barbara Jones, Mrs. Linda Metzler, John Mikesell, John Naylor, Milton Redman, Lawrence Rupley, Harvey Schwartz, Robert Thornton, and Leon Tucker.

We are indebted to the cooperation of Dr. Robert Tomlinson and Dr. Robert Lorenz, of the Instructional Materials Center, Office of Instructional Resources, and to Marty Kvame and Mignon Bishop of their staff, for timely preparation of the many visualizations required in the study.

Dr. Richard Spencer and his staff of the Division of Measurement and Research, Office of Instructional Resources, contributed very helpful assistance in the collection and evaluation of data. We would like to express our gratitude to James Hennes, Research Associate with the Office of Instructional Resources, for his assistance in processing and reporting the data analysis.

We would like to acknowledge the contribution of television director Stanton Saltzman, of the Division of Instructional Television, Office of Instructional Resources, and his production staff, who managed to meet the improbable schedule imposed by the investigators.

Our appreciation is also extended to Dr. Susan Meyer Markle, of the Office of Instructional Resources, Chicago Circle Campus, for reading portions of the final report and providing many helpful suggestions during the study.

We wish to thank Miss Susanne Sinnamon and Mrs. Gloria Hill for the preparation of manuscripts for the project and Mrs. Diane Paskus for her work on the numerous revisions of student outlines and television scripts.

TABLE OF CONTENTS

ABSTRACT

AN APPLICATION OF THE PRINCIPLES OF PROGRAMED INSTRUCTION TO A TELEVISED COURSE IN COLLEGE ECONOMICS

Title VII Project No. 5-0841; Grant No. 7-23-1020-248

This study has applied the basic procedures characteristic of programed instruction development to revision of a series of four video-taped lectures being used in an introductory college economics course. Several previous revisions of each lecture had been accomplished on the basis of an intuitive, more conventional procedure.

Two groups of students simultaneously viewed either the currently functional, conventional revisions or the programed revisions by means of two-channel, closed-circuit telecast as an integral part of the economics course. Relative student achievement was evaluated by means of a criterion test developed from specific instructional objectives derived from the existing conventional instruction in the revision process.

A three-lecture pilot study was conducted to develop procedures for applying the identified principles of programed instruction in the revision process. The programing procedures have been set forth and their application process described.

Analysis of criterion test performance in this pilot study by covariance on the basis of pre-test attainment indicated a 9% increment in mean achievement ($p < .015$; $t = 2.22$) favoring the students viewing the programed revision. The 9% increment continued to favor the experimental group ($p < .02$; $t = 2.16$) on a retention measure included in the course final examination.

A fourth lecture was incorporated and the series programed by applying the procedures developed in the pilot study. Similar covariance analysis indicated a 12% increment in mean achievement ($p < .0001$; $t = 4.30$) on the main study immediate post-test favoring students who viewed the programed revisions. Scores on a course evaluation questionnaire showed the programed revision students to be significantly more favorable in their attitudes than the conventional revision students after adjusting for pre-experimental differences. A mean achievement increment of 4% ($p < .025$; $t = 2.14$) continued to favor those students exposed to the programed revisions as evaluated by a retention measure embedded in the course final examination.

It has been suggested that the differential gain in student achievement may be a function of the change in teacher behavior between the revisions compared. Relevant pre- and post-study teaching-connected behavior of the major instructor has been summarized in order to provide a point of reference with respect to this change.

CHAPTER I

The Problem

During the past several years, televised instruction has evolved into a new phase. Initially, nearly all televised instruction was prepared and presented live, and thus was as evanescent as conventional instruction. Now, however, all television stations and a great many, if not the majority, of closed-circuit installations record and preserve instruction.

If properly managed, recorded instruction provides an opportunity for instructional improvement by making it feasible systematically to study, to analyze, and to revise the material. At the University of Illinois and, it is to be hoped, at most other institutions, revision of recorded instruction is both a right and responsibility of the academic departments concerned. In fact, major revision, frequently involving the entire course, is normally made after the first year. Some revision usually continues to be made thereafter.

It is common experience that instructors go about deciding upon revision largely on an intuitive basis. Rarely is revision undertaken for the purpose of bringing the instruction into harmony with some verbalizable, theoretical position about learning, and even more rarely is the result of the revision checked against student learning. Rather, the conscientious professor will watch his instruction, usually in the company of the students, gather impressions from his students, apply what he knows about teaching, and what he has been learning about television, and proceed to make such revision as his best judgment dictates.

This procedure is consistent with traditional methods, with the probably significant exception and addition that recorded televised instruction enables teachers to experience their own instruction. It may be that an experienced and sensitive teacher will accomplish as much by this procedure as by applying some more systematic and formal one.

However, a more systematic and formal approach to development of instruction has emerged from work on programed instruction and the pedagogical theory which undergirds it. Much of the success which has been reported for programed instruction in a variety of formats is very likely due not so much to the mode of presentation as it is to the method of development.

The basic and common developmental features applicable to various approaches to programed instruction seem to be:

1. The development of detailed specifications or objectives for each lesson in terms which permit one to test whether student performance is consistent with stated objectives.

2. The development of test items based upon each of these instructional objectives and a selection of some specific level of performance to serve as a criterion to judge the acceptability of the programed material.

3. Initial sequencing of the lesson content to an extent possible by logical analysis of the subject matter.

4. Conducting a series of empirical trials to eliminate ambiguity, interference, deviousness, etc.; checking each trial against the criterion test to determine whether the instructional objectives have been accomplished.

5. The provision for active response from students in such a way as to direct the attention of students to the most salient features of the lesson content.

6. The immediate reinforcement of correct responses, usually by providing knowledge of the correct results.

An additional characteristic, provision for individual pacing, which is probably not appropriate to the television medium has been shown by Carpenter and Greenhill[1] to be of doubtful importance when the student population is as homogeneous as a typical college population.

The purpose of this project then was to wed the versatile resources of television with the systematic developmental procedures of programed instruction, and to determine whether instruction prepared in this manner differed in effectiveness and in acceptability to students from televised instruction which had been developed and had undergone revision by the traditional, less systematic procedures.

Specifically, the following hypotheses were tested:

1. Programed televised instruction, having the characteristics described above, will result in student performance superior to that which results from recorded televised instruction which has been prepared and revised according to the best judgment of the competent, conscientious teacher.

2. The difference in achievement will persist for several weeks after exposure to the instruction.

3. Student attitudes toward programed televised instruction will be more favorable than toward a non-programed version of the same instruction.

---

[1] C. R. Carpenter and L. P. Greenhill, Comparative Research on Methods and Media for Presenting Programed Courses in Mathematics and English (University Park: Pennsylvania State University, 1963).

CHAPTER II

Related Research

Gropper and Lumsdaine[1] conducted three experiments using a different junior high school science lesson for each. In one they compared a "conventional" television lesson with a version of the same lesson which provided for active student response. In another they compared a conventional lesson with one which was sequenced so as to maximize the probability of correct active student response. In the third they compared a sequenced lesson with one which was pre-tested on a small group of students and revised so as further to reduce student difficulties in understanding material and thus increase the likelihood of correct student response.

They found that merely having the students respond had no beneficial effect on learning, presumably because students frequently made incorrect responses. However, the experimental students did significantly better on achievement tests when instruction was sequenced so as to increase the probability that they could respond correctly. Finally, when a lesson was pre-tested and revised, students seeing that version of the lesson did better than students who saw that version which had been merely sequenced but not pre-tested.

Our experiment, while very similar methodologically to the last Gropper - Lumsdaine experiment, differed from it in several important respects:

1. Gropper and Lumsdaine worked with junior high school students. Our subjects were freshmen and sophomore college students. The subject matter of the Gropper - Lumsdaine experiment was science. Ours was economics.

2. Gropper and Lumsdaine worked with a single lesson. We experimented with a unit of four lessons.

3. It appears from their report that the lesson for the Gropper - Lumsdaine experiment was prepared especially for the experiment and therefore it was probably a first effort which, from our experience, is usually inferior. Each of the four lessons of our control version had already undergone one or more revisions based on essentially conventional methods. Indeed, some student response and knowledge of results had already been provided for in the control version.

---

[1] G. L. Gropper and A. A. Lumsdaine, "An Experimental Evaluation of the Contribution of Sequencing, Pre-testing and Active Student Response to the Effectiveness of 'Programed' Television Instruction," Studies in Televised Instruction, Report No. 3 (Pittsburgh: American Institute for Research, 1961).

4. Response in the Gropper - Lumsdaine experiment usually consisted of students filling in a word or simple phrase. In our experiment, response sometimes took the form of generalizing more complex ideas from previous information, such as by drawing and completing simple graphs.

5. Gropper and Lumsdaine made no attempt to assess student attitude toward the different versions of the instruction.

The Carpenter and Greenhill research referred to previously compared student learning by means of programed, verbal, and numerical materials in English and mathematics with a variety of different programing formats, including teaching machines, film strips, and televised instruction. They found little difference in student achievement attributable to form of presentation. They concluded that with homogeneous groups, group pacing was feasible. They suggested that programed instruction, with the additional resources of television, could provide conditions as favorable for learning as those found in classes personally taught by above-average instructors. They did not seek to demonstrate, as does this study, that the process of course development based on programing principles yields instruction differing in quality from instruction conventionally produced. They do indicate the belief, however, that conventional patterns of television instruction must be broken if the medium is to be used with maximum effectiveness.

Schrag and Holland[1] revised a PSSC film so that each student's successful response assured a mastery of each concept or step leading to the concept in the film. Portions of the original film were deleted because they did not seem to deal with concepts in which motion was inherently important, and the material dealt with in that part of the film was programed in conventional paper-pencil form. A comparison of test scores of students viewing the original film with test scores of students viewing the programed version of it strongly suggested that the programed version was superior.

It would appear that the Schrag and Holland study was concerned primarily with providing systematic opportunities for student response in an already completed film. They did not revise the material of the film based on more clearly defined objectives or more careful sequencing.

---

[1] P. G. Schrag and James G. Holland, "Programing Motion Pictures: The Conversion of A PSSC Film into a Program," AV Communication Review, 13 (1965), 418-422.

CHAPTER III

Procedure

A. What Is Meant By "The Application of Principles of Programed Instruction"

Rapid expansion of interest in programed instruction in the decade since its inception (or revival, as the case may be) has led to a proliferation of programing "principles" as varied in number as in interpretation. The cornerstone of such principles has been a firm commitment to student accomplishment (i.e., to a demonstrable behavior change) as a result of the instruction so programed.

The basic developmental features of programed instruction previously noted in this report all exemplify this commitment. They served as guidelines for revision of the instructional sequence employed in the study.

1. Development of detailed specifications or objectives for each lesson in terms which permit one to test whether student performance is consistent with the stated objectives.

While it should seem obvious that anyone seeking to teach would clearly specify what he proposes to accomplish, such instructional objectives are rarely verbalized in an explicit manner. The task proved to be difficult, arduous, and highly informative. Content objectives implicit in the existing instruction were identified by analysis of previous tests, videotaped lectures, student outlines integrated with these lectures, and textual materials employed in the course. The major instructor functioned as a subject-matter specialist; the programer as a naïve student. Both proceeded with independent observations of these sources, and the listing of objectives employed in the pilot study was developed in many hours of joint discussion.

Subsequent analysis disclosed that some objectives were scarcely dealt with in the control, i.e., intuitively revised, lessons; other information in those lessons had little relationship to the explicit objectives identified. A constraint imposed by the comparative research design obligated the investigators to prepare a revised sequence, the instructional intent of which would be identical to that of the existing sequence. The resulting procedure of identifying only those objectives covered in the existing, conventionally-revised lectures was restrictive. Several opportunities for inclusion of additional supporting concepts occurred during the revision process, particularly as all lessons were dealt with as an integral unit when restructuring the sequence.

In those cases where the instructor did yield to desirable opportunities to include amplifying concepts, the instructional objectives were noted to be unique to the programed revision and were not included in the criterion post-tests employed to evaluate comparative performance.

It is acknowledged that instructional time devoted to these additional concepts could have been employed to expand upon or review the objectives identified in the existing lectures; however, it is noted that bias could only operate to the detriment of the research hypothesis, and the salutary potential of such revision was believed to be too desirable to dismiss.

2. Development of test items based upon each of these instructional objectives and the selection of some specific level of performance to serve as a criterion to judge the acceptability of the programed instruction.

A graduate assistant in economics worked from the listing of identified objectives to prepare criterion test items. These test items were reviewed jointly by the instructor, programer, and item writer to ensure their adequacy with respect to face validity and difficulty level. Joint item review also provided a check of the operational "interpretability" of the behavioral objectives, i.e., whether the instructor's intent was communicated to another subject-matter specialist. The graduate assistant who prepared criterion items served in no other capacity in the study.

3. Initial sequencing of the lesson content to an extent possible by logical analysis of the subject matter.

The content of economics is not hierarchical in nature; however, a certain logical sequence is found to exist and to be dependent upon the relative manner of representation or definition of major concepts. These relationships may be characterized as parallel in the sense that related sets of concepts must be mastered by the student before he can master more encompassing principles and concepts. An analysis for a logical preliminary sequence is possible within the definitional framework imposed by the interpretation of major concepts preferred by an individual subject-matter specialist.

4. Conducting a series of empirical trials to eliminate ambiguity, interference, deviousness, etc.; checking each trial against the criterion test developed above to determine whether the instructional objectives have been accomplished.

Once identified, this preliminary sequence was submitted to trial, usually with one or two individuals and--in the case of trials of a revised sequence--occasionally with groups of two to five individuals. Students employed in these trials had been exposed to only that instruction leading up to the lesson being revised. They participated in an additional one-hour session by immediately taking the criterion test items and discussing with the instructor and programer any difficulties encountered in understanding the lesson.

It should be noted that while a commitment to student attainment is undoubtedly characteristic of most conscientious instructors, such a commitment is difficult to translate into a behavioral frame of reference

merely in response to the exhortations of a cooperating programer. A noticeable shift by the instructor in this respect commences with the first empirical trials, particularly when criterion post-test "feedback" of individual student attainment, or lack of attainment, becomes painfully evident.

5. Provision for active response from students in such a way as to direct their attention to the most salient features of the lesson content.

While general student activity and involvement is necessary for any learning, the type of response frequently encountered in programed texts -- consisting of inserting words, completing sentences, and so on -- frequently seemed to be too trite, too contrived, or too boring. When programing the televised lectures and coordinated student outline, an attempt was made to provide for overt response in the outline which would cause the student to attend to discriminations of those similarities or differences identified as major points of importance. This identification had previously been made by means of either the analysis step in developing the initial content sequence or in subsequent analysis of the difficulties experienced by individual students in empirical trials.

Students were encouraged to refer to the outline before and after the televised lecture as well as during it. A concise introduction preceded each lecture. The body of the student outline was, in effect, an extensive set of structured notes which was completed by the student during the televised lecture.

Often the student was required to respond with a symbolic or "shorthand" response; for example, a plus or minus sign depicting the discrimination of a rising or falling graph trend. A deliberate attempt was made to minimize the manipulation of arithmetic during the lecture in preference to such shorthand responses, with opportunities for actual computation provided in the post-lecture review questions incorporated in the course outline.

Response opportunities had also been provided in the student outline integrated with the conventionally-revised televised lessons, but not on the basis of an explicit rationale supported by the trial data.

6. The immediate reinforcement of correct responses, usually by providing knowledge of correct results.

Reinforcement was provided in a gross manner in the form of confirming comments by the major instructor presenting the televised lecture. Confirmation was presented in as unobtrusive a manner as possible, such that it appeared as a natural, consequent statement in the lecture discourse. For example, after asking students to alter a graphic presentation in their student outline to indicate an increase in investment expenditure, the instructor might say, "If you have done this correctly, you have a line like the one you see on this graph here," and point to a visual being viewed by the television camera.

To the extent that the reinforcement technique employed did not provide for the student who had emitted an erroneous response, it was nonadaptive. Linear programing in general suffers from the theoretical deficiency of providing knowledge of correct results which while reinforcing for those who responded correctly is not reinforcing for those who responded incorrectly. In a linear program, individuals who respond incorrectly generally get no opportunity to make a correct, reinforced response.

B. Description of the Course

A one-semester terminal course in introductory economics served as the research vehicle for this experiment. The course, designed for students who are not majoring in commerce, was first presented by television at the University of Illinois in 1959. Since 1960, all television lectures have been revised at least once with most undergoing several revisions.

The course typically has an enrollment of about 800 students, mostly freshmen and sophomores, one-third to one-half of whom are taught in part with the use of television. The first and third 50-minute periods each week are instructed by television lectures with the second period devoted to a discussion and quiz section supervised by a graduate teaching assistant.

There is a student course outline of about 150 pages integrated with the television lessons. Each lecture is outlined in detail and includes appropriate visual presentation of supporting graphs, charts, and tables. Various questioning techniques are employed in the lectures with appropriate space for student response appearing in the outline in order to encourage student involvement.

When compared from time to time against student achievement in non-televised sections of the course, the usual finding has been one of no significant difference. Occasionally the televised sections have been superior, occasionally the conventional sections. Student attitudes toward the televised course have been found to be negative at the beginning of the course, but to improve before the course is completed.

C. Description of Pilot Study Procedures

A unit of instruction, previously dealt with in three televised lectures, was selected for programed revision in the pilot study. The unit, covering the theory of income determination, commenced the fifth week of the semester. Specific behavioral objectives for this instructional unit were identified jointly by the instructor and programer as described in step one of developmental features (Section A). Steps two and three (Section A) described the procedures followed to construct a criterion post-test, based upon the identified objectives, and to outline a preliminary instructional sequence.

A preliminary student outline was prepared for each trial presentation in the same format as that of the course outline used with the conventionally-revised lectures. Prior to the empirical trials, students who served as paid subjects were instructed in the content determined to be basic to the experimental sequence, i.e., normally covered in the sequence of lectures prior to the experimental treatment. The trial subjects actively responded in copies of the student outline as the instructor presented the entire trial sequence in a live simulation of his usual television studio manner. All pilot study trials were conducted in this face-to-face manner in a small conference room. The programer presented all visuals by overhead projection of transparencies.

Post-trial discussions with subjects were tape recorded during all trials of the first lesson revised by the instructor and programer. The tape recorder was dispensed with in subsequent trials as the investigators gained proficiency in recognizing the type of information most useful to them in further revision. It was found to be essential that the programer make a note of the amplifying information employed by the instructor to clarify student difficulties. The instructor, while genuinely committed to the improvement of his instruction, would occasionally -- and honestly -- fail to recognize that students required clarification with respect to particular points in the lecture. To engage in such clarifying explanations appears to be so natural a part of an instructor's behavior, developed over many years of classroom experience, that the essential character of such cogent and informative explanations tends to go unnoticed.

Individual trials with two to five students were continued until student performance on the criterion post-test items and general level of subjective proficiency expected by the instructor were both viewed as satisfactory. The second and third lessons were developed following the same procedures, with the exception that small groups of six to eight students were exposed to live simulation of the previous lessons in the sequence. Subjects for further individual trials were drawn from these groups exposed to all previous lessons in the sequence.

Revised student outlines were prepared for the three lessons in the same format and using the same production process (offset) employed for the material incorporated in the outline purchased by all students for use in the course. Rehearsals were conducted with television studio personnel and the three revised lectures were recorded on videotape. The visuals in the form of overhead transparencies produced for use in the empirical trials were projected on a screen in the studio and viewed in this manner by the television cameras.

It should be noted that the procedures for developing the programs for the pilot and main study did not involve a specialist in television production in the planning stages. At that time, such a person might have been able to suggest ways in which the medium could be used to enhance the effects of the programing procedures. In effect, the instruction was presented in much the same way as it had been prior to instituting programing procedures. A television specialist was involved in planning earlier

versions of the course, including the control version. Whether an imaginative and creative use of the medium designed specifically to enhance the systematic programing effort would be effective is a question of great interest, but it was not investigated in this study.

D. Conduct of the Pilot Study

Comparison of the control and experimental treatments was effected across four sections meeting simultaneously in conventional classrooms capable of seating up to 45 students. All classrooms were located in the same building and connected to the closed-circuit television transmission facility of the University. Each room was equipped with two television receivers located one to each side of the front of the class.

The major instructor, familiar with comparative reception quality of the four classrooms, rank-ordered the rooms in this respect. The treatment groups were constituted by pairing rooms ranked one and four as one group and rooms ranked two and three as the other group. These groups were randomly assigned to receive either the conventionally-revised or the programed sequences during the treatment period.

The pilot study sample consisted of all students in the 10:00 o'clock televised sections of Economics 108 during the Spring semester of 1965. All four class rosters were pooled and students assigned in a random manner to classes for the duration of the treatment period. There were 47 students in the control groups and 53 students in the experimental group. Nearly all the students were sophomores, with the sample averaging 51 semester-hours completed prior to the previous semester (Fall, 1964). Students were informed that they were to take part in a study to determine the effectiveness of three revised televised lectures. They were advised of their class reassignment for the two-week period by memorandum distributed during the class immediately prior to the treatment period. An attitude pre-test was also administered to all students on the occasion of this class meeting.

Revised course outlines for the three programed lectures were distributed to the experimental students as they arrived in class to attend the first treatment lecture. The two experimental sections were instructed to disregard the next three lessons in the student outlines they already possessed. Videotapes of both the conventionally-revised and programed lectures were played back simultaneously, with classroom receivers tuned to two different closed-circuit channels for appropriate lesson reception.

The immediate criterion post-test was administered simultaneously to all students at the first class meeting following the treatment period. Students had been informed that test results would be a part of their grade determination. An attitude post-test was administered to all students upon completion of the criterion examination. The delayed criterion post-test was included as part of the course final examination administered twelve weeks after the treatment period.

E. Description of the Main Study Procedures

An item analysis of immediate and delayed post-test results constituted the primary source of data for subjective decisions in programing the main study sequence. The post-test data indicated the possibility that Economics 108 content covered prior to the treatment period was not mastered to the level assumed in programing the pilot sequence. Naïve students employed in pilot study trials were instructed in concepts relevant to the sequence, e.g., familiarity with the factors of income; thus the apparent need for review was not disclosed in empirical trials.

A source of trial subjects more representative of the target population was needed. The major instructor had continued to teach one conventional (non-televised) lecture section of approximately 40 students each semester, a practice affording him a measure of direct contact with students while administering the televised course. In addition, this section (Section L) occasionally provided him with an opportunity to rehearse planned changes in the televised lectures.

The major instructor decided to postpone two of the first four weeks of Section L content, a portion relatively unrelated to the treatment sequence, until later in the course. Students exposed in the live class to the prior lecture in the sequence were solicited from Section L and paid to participate in the main study individual trials conducted in the same manner as reported for the pilot study. This, in effect, provided a pool of more representative subjects and permitted a two-week trial and processing time for each lecture to be re-programed in the main study.

Three procedural variations were incorporated in the main study to minimize student recognition of their experimental participation.

1. Random assignment was accomplished by program alterations in the computerized pre-registration system of the University which resulted in student assignment to one of two concurrent groups (control and experimental) of Economics 108.

2. The student outlines for the conventionally-revised lectures to be covered in the treatment period were excluded in the binding process. Course outlines purchased by the students, as a consequence, did not include these lessons. Students were advised that the missing lesson outlines were being revised and would be distributed at a later date without further cost. These outlines were delivered to the investigators and distributed to the control group at the time of distribution of the revised lesson outlines coordinated with the programed lectures.

3. The Course Evaluation Questionnaire (CEQ) pre-test was administered the first week of class, instead of during the

class prior to commencing the experimental treatment, as in the pilot study. The investigators noted that the decrement in attitude as measured by the CEQ in the pilot study (see Chapter IV) might have resulted from overexposure to treatment or measurement, i.e., randomizing by classroom reassignment, CEQ pre-test, distribution of new student outlines, CEQ post-test, and immediate post-test (all conducted in less than a three-week period). Several pilot students specifically mentioned that they had just taken the CEQ two weeks prior to receiving it at the end of the treatment period. Students in the main study were told that they would be asked to respond to the same CEQ later in the course.

Attention in the pilot study was directed to evolving techniques to develop detailed objectives and criterion test items, to sequence instructional content, and to record the programed lectures on videotape. The following steps were suggested as useful in programing instructional television.

1. Instructor specification of global objectives -- including several to be attained by "better" students.

2. Instructor/programer review of existing instructional materials to write detailed specification of behavioral objectives.

3. Preliminary identification by instructor of those behaviors which constitute minimum acceptable terminal performance.

4. Complete list of objectives given to criterion test item writer.

5. Review of criterion items produced. This step accomplished jointly by the item writer, instructor, and programer to ascertain (a) the representative adequacy of the items, and (b) the operational "interpretability" of the behavioral objectives.

6. Review of objectives by the instructor and programer to outline preliminary instructional sequence.

7. Instructor writes detailed outline of instructional sequence and assembles supporting visuals. Consultation with television production staff to avoid any unforeseen limitations on the reproduceability (or other production limitations) with regard to the proposed visuals.

8. Instructor and programer review the preliminary outline (at this stage, an effort to balance the technical accuracy of presentation with the need for representational clarity to the student).

9. Instructor prepares a student outline to be integrated with lecture.

10. Instructor and programer conduct initial trial with an individual student, administer criterion items, and tape record subsequent student interview -- which begins with discussion of criterion items and proceeds to discussion of the lecture and outline.

11. Instructor and programer discuss possible revisions. Instructor incorporates revisions decided upon and prepares new outline (and visuals, if necessary).

12. Steps 10 and 11 are repeated until, in the opinion of the instructor, the minimum criterion performance is reliably attained by representative trial students.

13. Instructional materials center prepares visuals as overhead transparencies.

14. Instructor finalizes script and student outline, which is reproduced in a form acceptable for group trial.

15. Instructor conducts a live trial with a full class, to include:

    a. instructor presentation of the sequence using overhead transparency visuals,

    b. instructor and programer evaluation of group behavior with respect to attention, response level, etc.,

    c. administration of criterion test items,

    d. large group discussion of test with comments solicited by instructor and programer, and

    e. instructor and programer discussion of this trial, as soon as possible, to reach agreement on further revision necessary.

16. Item analysis of group performance on criterion test to identify (a) previously undisclosed weaknesses in the instructional sequence, and (b) inadequate or possibly ambiguous criterion items.

17. Correction of deficiencies revealed in steps 15 and 16, with additional individual retrial when warranted (in the subjective judgment of the instructor and programer).

18. Completion of final printing of student outline, visual art work by the television and instructional materials center artists.

19. Rehearsal with television production crew and videotaping of the instructional sequence.

20. Replay of the videotape to disclose and correct any obvious production errors.

21. Large lecture use of videotaped lecture with observation of students by instructor and programer to check attention level, response level, etc. Discussion with individual students to gain a subjective impression of student reception.

22. Review of criterion post-test data for determination of persistent areas of weakness in the instructional sequence.

A further step not incorporated but recommended and planned as a guide to subsequent revision is a random sample of outlines, after the lecture, to determine the adequacy of student response.

These steps were followed in developing the programed lectures employed as the experimental treatment in the main study. In addition, sequence revision of the pilot study suggested that the content of an additional lecture should be incorporated in the sequence. A fourth lecture was added to include material dealing with the role of government in income determination.

The availability of Section L for group trial also permitted discussion of the criterion post-test items in a group setting (step 15d above). This possibility provided active discussion on which to base subjective judgments with respect to item ambiguity.

F. Conduct of the Main Study

The main study sample consisted of 242 students registering for 12 sections of Economics 108 offered by television during the Fall semester of 1965-66. Students were randomly assigned to control and experimental groups with the assistance of the computerized pre-registration system in use at the University.

Both sections met at 9:00 o'clock on Monday and Friday morning in comparable large lecture rooms equipped with television receivers mounted in overhead brackets about the room. Although these classrooms were located within different buildings, the buildings were relatively close to one another in the same section of the campus. Again by means of dual-channel closed-circuit transmission, the conventionally-revised (control) lessons were viewed by 126 students and 116 students viewed the programed (experimental) lessons.

It was not standard procedure to take attendance at the televised lectures. An apparently high absence rate was subjectively noted during the four weeks prior to the treatment period and during the two week treatment period. This prompted the investigators to consider attendance. Thus, upon conclusion of the treatment period, all students were asked

to indicate their attendance over the previous two weeks. It was found that nearly 100 students had either missed one or more of the four lectures during the treatment period, or had attended some of the lectures in the "wrong" classroom (i.e., not with their assigned groups) thus confounding their experimental or control status. Some 33% were absent from one or more lectures, while 5% viewed one or more lectures with the group to which they were not assigned.

The students who had been absent or had viewed one or more lessons in the wrong location were dropped from the sample. The sample size was reduced to 145, with 78 control students and 67 experimental students.

An immediate post-test composed of 54 multiple-choice, best-answer items was administered to control and experimental groups simultaneously in the first lecture meeting after the treatment period. Twenty-five of these items were repeated as a part of the final examination to measure retention. This examination was administered twelve weeks after conclusion of the treatment period.

G. Analysis of the Data

The data collected provided measurements on the following variables:

Dependent (criterion) Variables

1. Immediate post-experiment achievement test
2. Delayed post-experiment achievement test
3. Course Evaluation Questionnaire -- post-experiment
   a. General course attitude
   b. Attitude toward method of instruction
   c. Attitude toward course content
   d. Student interest and attention
   e. Attitude toward instructor
   f. Other specific items
   g. Total attitude score

Independent (experimental) Variables

4. Control vs. experimental group

Control Variables

5. School and College Ability Test (SCAT)
   a. Verbal
   b. Quantitative
6. Cooperative Reading Test
   a. Verbal
   b. Speed
   c. Comprehension
7. Cumulative grade point average
8. Cumulative number of semester hours completed
9. Attendance record during experimental period
10. Course Evaluation Questionnaire -- pre-experiment
11. Economics pre-experiment achievement test
    a. Objective (multiple choice) score
    b. Subjective (essay) score
12. Quiz section of student

The achievement tests were developed as described earlier. The test for the pilot study consisted of 32 multiple choice items. The test reliability (KR-14) in the control group was .75, and in the experimental group was .83. The test for the main study contained 54 multiple choice items, and had a test reliability of .70 in the control group and .82 in the experimental group. The first regular hourly examination served as the pre-experiment achievement test in the main study. It was developed by the major instructor and the teaching assistants and consisted of a subjective section (essay) and an objective section (multiple choice). The two sections correlated .40 with each other and individual pre-test scores correlated .46 with their post-test scores.

The Course Evaluation Questionnaire has been developed over some time by R. E. Spencer, Office of Instructional Resources, University of Illinois. The instrument consists of a series of statements eliciting opinions about various aspects of an instructional program, and provides subscores which have been normed over a population of several hundred

classes and several thousand students.[1] Factor analytic evaluation has resulted in five reproducible factors which, along with several unique items and a total score, provide the seven scores reported. Reliabilities of the subscores have ranged .50 to .90, generally being in the .70's or .80's. Reliability coefficients for the total scores have been consistently over .90.

The analysis was essentially a comparison of the control and experimental groups on the various criteria. While, in the main study, the students were randomly assigned to the experimental condition, additional control was provided in both the pilot study and main study by statistically controlling on certain variables. The statistical methods employed in the study consisted chiefly of correlations, analysis of variance, and analysis of covariance. Test evaluation methods and chi square were also used. Sample sizes varied from one analysis to the next, as indicated in Chapter IV, since complete data were not available for every student.

[1] Copies of the questionnaire and interpretation manual may be obtained from the Office of Instructional Resources, University of Illinois, Urbana, Illinois.

CHAPTER IV

Results

The principal question under investigation was whether or not a difference in achievement and attitude would occur favoring students viewing lectures revised by application of principles of programed instruction over students viewing traditionally revised lectures.

A. Pilot Study

The first analysis of the pilot study sample included the covariables of School and College Ability Test (SCAT) scores, Cooperative Reading scores, grade point average (GPA), and semester hours completed. The criterion was the achievement test administered three days after the last experimental lecture and again twelve weeks later at the end of the semester. The achievement test correlated from .02 to .18 with the various SCAT and Cooperative Reading subscores, all nonsignificant. Grade point average correlated .38 ($p < .001$) with the immediate post-test scores, but only .16 (n.s.) with the delayed post-test scores. The immediate post-test and delayed post-test scores correlated .57 ($p < .001$) with each other.

Table 1 indicates the results of the first analysis of covariance, covarying on SCAT, Cooperative Reading, GPA, and semester hours. The immediate post-test scores showed a significant difference favoring the experimental group (one tail $p < .025$). The delayed post-test score difference was slightly less significant (one tail $p < .05$).

Table 1

Pilot Study Analysis of Covariance Results--First Pass

Criterion	Control Group Mean N = 38	Experimental Group Mean N = 41	t Ratio (Adjusted)	Sig. Level (1 tail)
Immediate Post-Test . . .	21.55	23.56	2.04	.025
Delayed Post-Test	21.29	23.29	1.84	.05

When comparing the differences between the experimental and control groups on the covariables, it was found that only the Cooperative Reading verbal test score showed significant differences (the control group scored higher). Since only one of the SCAT-Coop covariables showed differences, and since 19 students were missing SCAT scores and therefore had to be dropped from the analysis, a second analysis was made. This time only the student's grade point average and semester hours completed were covaried. By not using the SCAT results in the equation, the older students who were more likely to be missing SCAT scores were included in the sample. For example, the average semester hours completed increased from 46.7 to 50.8 when the SCAT scores were dropped. Table 2 indicates the results of the second analysis.

Table 2

Pilot Study Analysis of Covariance Results--Second Pass

Criterion	Control Group Mean N = 43	Experimental Group Mean N = 53	t Ratio (Adjusted)	Sig. Level (1 tail)
Immediate Post-Test. . . .	21.65	23.47	2.22	.015
Delayed Post-Test.	21.44	23.42	2.16	.02

Scores for students viewing the programed lessons were nine per cent greater than the scores of students viewing the conventionally-revised lessons. There was very little loss in retention over the 12 weeks before the final exam. Neither the change in score for either group nor the relative difference in change in score between control and experimental group was significant. That is, there was no advantage or disadvantage in measured rate of retention related to viewing programed lectures.

The Course Evaluation Questionnaire was administered just before and just after the pilot study experimental period. Both before and after the experimental period the control group was less favorable in evaluation that the experimental group. None of the seven pre-experiment differences between control and experimental groups was significant, and only one of the seven post-experiment comparisons (instructor evaluation) showed a significant difference (using two-tailed significant tests). Both the control and experimental groups became less favorable toward the course over the period of the experiment, with the control group decreasing significantly on all seven scores, and the experimental group decreasing significantly on only one of the seven subscales (general course attitude).

The percentile ranking (all university norms) for the total course evaluation score in the pilot study control group was 49 before the experimental period and 20 after the experimental period. The percentile ranking for the experimental group on total score was 60 before the experimental period and 38 after the experimental period.

B. The Main Study

Table 3 indicates the results of the analysis in the main study. The immediate and delayed post-test scores were covaried on the pre-test scores.[1] The delayed post-test consisted of 25 of the 54 items from the immediate post-test. Each post-experimental Course Evaluation Questionnaire score was covaried on the same pre-experimental subscale.

Table 3

Main Study Analysis of Covariance Results

Criterion	Control Group Mean N = 66	Experimental Group Mean N = 54	t Ratio (Adjusted)	Sig. Level
Immediate Post-Test. . . .	36.48	41.02	4.30	.0001[a]
Delayed Post-Test.	19.67	20.73	2.14	.025[a]
CEQ -- Total	2.29	2.13	2.36	.02[b]
CEQ -- Subscores				
General Course Attitude	2.04	1.93	1.32	.20[b]
Method of Instruction .	2.54	2.22	2.13	.04[b]
Course Content.	2.32	2.22	1.94	.06[b]
Interest and Attention	2.44	2.16	2.62	.01[b]
Instructor.	2.18	2.10	1.32	.20[b]
Specific Items.	2.24	2.13	2.11	.04[b]

Note: CEQ scores were scored with 1.00 as the most favorable score and 4.00 as the least favorable score.

[a]One-tailed test

[b]Two-tailed test.

[1]Inclusion of additional covariance factors in the pilot study resulted in little change in level of significance. As the pre-test scores of the control group exceeded those of the experimental group and the post-test data resulted in an extreme significance of difference in the opposite direction, use of further covariance factors in the main study did not appear to be necessary.

The mean increase in immediate post-test scores for the experimental group over the control group was twelve per cent, significant at the .0001 level. Five of the seven Course Evaluation Questionnaire scores showed the experimental group significantly more favorable than the control group after adjusting for their pre-experiment differences.

Again, as in the pilot study, both before and after the experimental period, the control group was less favorable in evaluation than the experimental group. The control group became significantly less favorable toward the course over the period of the experiment on four of the seven attitude scores. The experimental group did not show any significant changes. As Table 3 indicates, five of the seven attitude scores show favorable evaluations for the experimental group relative to the control group.

Figure 1 shows the relative change in evaluations for the groups in both the pilot and main study. The pilot study groups expressed larger shifts in attitude, consistently downward. The experimental group in the main study expressed relatively little attitude change, with two of the seven scores (method of instruction and specific items) showing slight positive shifts. It may be noted that, in general, the students rated Economics 108 at or below the average rating for courses in the norming population.

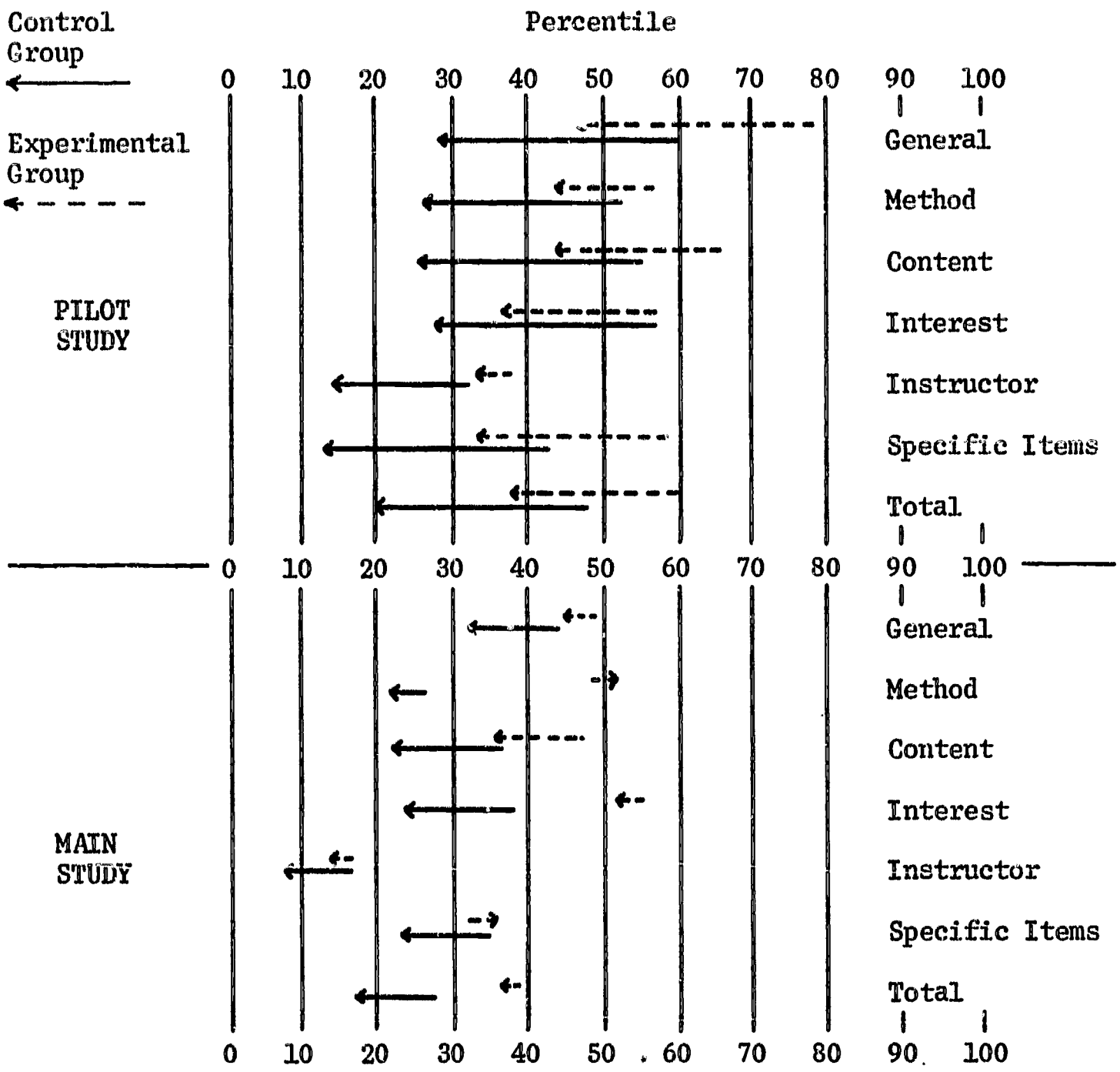
C. Discussion

The comparison of primary concern in this study has been investigated by means of simultaneous transmission of two series of videotapes. Each of these tapes may be thought of as recorded episodes which reflect the teacher behavior of specifying objectives, either explicitly or implicitly, and of reaching decisions as to content appropriate to these objectives -- whether specifically guided by actual trial with students or generally guided by years of classroom experience with students, or both. The tapes disclose how the instructor sequenced various verbal and graphic examples to portray the concepts and principles, and how he provided opportunities for students to be actively involved in the learning process. All these and many other decisions reflect the total behavior of the major instructor acting upon assumptions he currently held with reference to how students learn.

The data set forth in this chapter are derived from the evaluation of effectiveness attained by one instructor acting upon earlier assumptions with regard to instructional development (conventional revisions) compared with a situation in which, functioning as his own control, he acted on the basis of new assumptions (programed revisions). It is suggested that the differential student achievement gain, reflected in the data, may be attributed to the change in behavior of the instructor as he acted on the basis of new assumptions which, as a result of engaging in the process of empirical revision, he perceived to be more appropriate in developing instructional sequences. For purposes of comparison, relevant pre- and post-study behavior of the major instructor is summarized -- necessarily in somewhat subjective manner.

FIGURE 1

Shifts in Attitude Measures From the
Course Evaluation Questionnaire Expressed
in All-University Percentile Norms



NOTE: Each arrow represents the shift from pre-experiment attitude position to post-experiment attitude position.

Pre-study teaching-connected behavior. The major instructor had developed and continuously revised a series of 26 televised lectures and an associated 150-page student outline over a period of six years on essentially an intuitive basis. The outline incorporated intuitively-based provision for active student response in the lecture-coordinated portions. A sizable bank of test items had been generated with reference to data from statistical item analysis. The major instructor had developed and used a television-oriented student attitude measure and had conducted informal studies to investigate the relative performance of students viewing televised lectures. Several journal articles describing the procedures employed in developing the televised lectures had been published.[1]

Post-study teaching-connected behavior. The major instructor has accepted and will implement the use of empirical trials as a basis for revision of his remaining lectures.[2]

Having specified explicit objectives in terms of student performance, the major instructor recognized that some students might already be able to demonstrate such performance. A one-hour pre-test covering four content areas has been developed and administered to all second semester, 1966, televised sections. Scores in the four areas will be correlated with final course grades. Lacking this normative data, students who achieved above the 90th percentile were counseled as to their areas of strength and weakness and advised to take a three-hour proficiency credit in the course. Items evaluating basic graphic interpretation and arithmetic skills were included in the pre-test. If normative data indicate a relationship between low final grades and low test scores on these items, a series of remedial aids will be prepared for students.

The major instructor has demonstrated an increasing concern for the effectiveness of his graduate teaching assistants. Recognizing the

---

[1]Charles J. McIntyre and Donald W. Paden, "Economic Considerations in Televised Instruction," Journal of Higher Education, 23 (1962), 319-323.

Donald W. Paden, "The Teaching of Economics via Television at the College Level," supplement to Televised College Courses, Ed. by John W. Meaney. New York, Fund for the Advancement of Education, 1962.

Donald W. Paden, "The Teaching of Economics by Television." Quarterly Review of Economics and Business, 2(4) (1962), 29-39.

Donald W. Paden, "How Teachers can use Television to Improve Instruction -- Higher Education." in Improvement of Teaching by Television, Ed. by Barton L. Griffith and Donald W. MacLennan, Columbia, University of Missouri Press, 1964, 77-81.

[2]Four additional lectures have been programed as of the date of this report, with representative students serving as individual subjects.

benefit to an instructor who attempts to apply the procedures developed in this study, he assigned for revision each of three lectures to pairs of assistants. They prepared detailed objectives as a first revision step. Test items were proposed to evaluate student performance with respect to these objectives. The assistants then developed lectures and coordinated outlines were prepared and printed. The lectures were recorded on videotape by the assistants and later telecast as an integral part of the course. No observation data have been recorded to evaluate the results of these experiences. However, the major instructor has proposed that such training in programing procedures be conducted for all Economics teaching assistants.

Implications of teaching-connected behavior. The pre-study behavior of the major instructor is believed to have indicated an above-average commitment both in terms of time and energy to the development and implementation of high quality instruction. The data indicate that functional, high quality instruction derived from the efforts of this instructor has been revised such that a significant gain in student achievement has been obtained.

While the procedures developed and applied in the study are believed to possess general applicability, the results necessarily depend upon the competence reflected in the teaching behavior of the instructor who developed and videotaped both revisions. The relevant pre- and post-study behavior of the major instructor have been described in order to provide a point of reference for those who may wish to generalize the results of the study to their instructional situation.

CHAPTER V

Summary and Conclusions

Instructional improvement is usually a largely intuitive process. A conscientious professor will gather impressions from his students, draw upon ideas found to be useful in whatever experience he has had, and proceed to revise his instruction as his best judgment dictates. This study has sought to investigate a more systematic and formal approach to the problem of instructional development.

The capabilities of videotape to record and preserve instruction were employed to compare the relative efficiency of two sequences of instruction telecast to students in an introductory college economics course. The basic features of programed instruction, which afford a systematic approach to development of instructional sequences, were combined with the reproduction capabilities of the television medium.

The following principles of programed instruction served as a basis for development of the programed revision employed in the study.

1. Development of detailed specifications or objectives for each lesson in terms which permit one to test whether student performance is consistent with the stated objectives.

2. Development of test items based upon each of these instructional objectives and the selection of some specific level of performance to serve as a criterion to judge the acceptability of the programed instruction.

3. Initial sequencing of the lesson content to the extent possible by logical analysis of the subject matter.

4. Conducting a series of empirical trials to eliminate ambiguity, interference, deviousness, etc.; checking each trial against the criterion test developed above to determine whether the instructional objectives have been accomplished.

5. Provision for active response from students in such a way as to direct their attention to the most salient features of the lesson content.

6. Immediate reinforcement of correct responses, usually by providing knowledge of correct results.

The study was conducted in two phases. A pilot study was undertaken to develop and try out procedures by which the principles of programed instruction could be applied to televised instruction. A representative unit of three existing lessons was selected and programed in this phase of the study.

The programed revisions were developed in a series of individual and small group (2-5 students) trials employing representative students as subjects. These subjects actively responded in trial copies of the coordinated student outline as the major instructor presented the entire trial sequence in a live simulation of his televised instruction. Trial visuals were used in all these face-to-face simulations, conducted in a small conference room. Subjects responded to applicable criterion test items after the presentation and then discussed the presentation at length with the major instructor and programer.

The existing instruction consisted of two videotaped, 50-minute lectures presented each week by closed-circuit television to a large lecture section. A small group (20-30 students) mid-week quiz session led by a graduate teaching assistant provided an opportunity for discussion of the lecture content. The body of a 150-page student outline was, in effect, a structured set of notes coordinated with the televised lectures. It also included an introduction to each lecture and a set of post-lecture review questions.

All of the videotaped lectures and their associated student outlines had been revised several times by an essentially intuitive, "conventional" procedure in the six years that the major instructor had offered the course by television.

Two groups of students simultaneously viewed either the currently functional, conventional revisions or the programed revisions by means of two-channel, closed-circuit telecast as an integral part of the course. Relative student achievement was evaluated by means of a criterion test developed from specific instructional objectives derived from the existing conventional instruction in the revision process.

The pilot study sample consisted of all students in the televised sessions meeting at the same hour during the Spring semester of 1965, with 47 students viewing the conventional revisions and 53 students viewing the programed revisions. Analysis of criterion test performance in this pilot study, by means of covariance on the basis of pre-test attainment, indicated a 9% increment in mean achievement ($p < .02$; $t = 2.22$) favoring the students viewing the programed revision. The 9% increment continued to obtain ($p < .02$; $t = 2.16$) and to favor the experimental group on a retention measure included after 12 weeks in the course examination.

Attention in the pilot study was directed to evolving techniques to develop detailed objectives and criterion test items, to sequence instructional content, and to record the programed lectures on videotape. The description of the main study procedure in the body of this report enumerates 22 procedural steps suggested as useful in programing instructional television. These procedures were followed in revising the four lectures employed in the main study, an additional lecture having been suggested as appropriate in the sequence as a result of considerable restructuring of content in the pilot study.

The main study sample consisted of 242 students in all televised sections of the course during the Fall semester of 1965 - 1966. Students were randomly assigned to control and experimental groups. Various corrections to these groups to adjust for attendance and missing data resulted in a 66-student sample viewing the conventionally-revised lectures and a 54-student sample viewing the programed revision. Similar covariance analysis indicated a 12% increment in mean achievement ($p < .0001$; $t = 4.30$) favoring students who viewed the programed revisions. Twenty-five of the 54 items on this criterion post-test were readministered as a measure of retention 12 weeks later as a part of the course final examination. A mean achievement increment of 4% ($p < .025$; $t = 2.14$) continued to favor those students who viewed the programed revisions.

A Course Evaluation Questionnaire, consisting of a series of 50 opinion statements to which the student indicated a level of agreement or disagreement, was used to evaluate relative student attitude. Subscores on the CEQ have been normed over a population of several hundred classes and several thousand students, with reliability coefficients for total score consistently over .90. Five of the seven CEQ subscores showed the students viewing the programed revisions significantly more favorable than the students viewing the conventional revisions after adjusting for their pre-experiment differences.

Three hypotheses were tested in the study:

1. Programed televised instruction will result in student performance superior to that which results from recorded televised instruction which has been prepared and revised according to the best judgment of a competent, conscientious teacher.

2. The difference in achievement will persist for several weeks after exposure to the instruction.

3. Student attitudes toward programed televised instruction will be relatively more favorable than toward a non-programed version of the same instruction.

The three research hypotheses have been confirmed by the data.

It has been suggested that the differential gain in student achievement may be a function of the change in teacher behavior between the revisions compared. Relevant pre- and post-study teaching-connected behavior of the major instructor has been summarized in order to provide a point of reference with respect to this change.