A COMPARATIVE ANALYSIS OF THE IMPACT OF VARIOUS METHODS OF INSTRUCTION ON ACHIEVEMENT AND UNDERSTANDING IN MATHEMATICS FOR ELEMENTARY TEACHERS

By

PHILIP EDWARD GIBBONS

Bachelor of Arts
King's College
Wilkes Barre, Pennsylvania
1955

Master of Arts
Boston College
Chestnut Hill, Massachusetts
1961

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University in
partial fulfillment of the requirements
for the degree of
DOCTOR OF EDUCATION
July, 1967

A COMPARATIVE ANALYSIS OF THE IMPACT OF
VARIOUS METHODS OF INSTRUCTION ON
ACHIEVEMENT AND UNDERSTANDING
IN MATHEMATICS FOR
ELEMENTARY
TEACHERS

Thesis Approved:

Thesis Adviser

Dean of the Graduate School

ACKNOWLEDGMENTS

The writer wishes to express his deep appreciation to each member of the dissertation committee: Drs. W. Ware Marsden, Vernon E. Troxel, Milton E. Berg, Gerald K. Goff, Wayne K. Hoy. Special thanks is extended to Dr. W. Ware Marsden, chairman, for his continued, sincere interest and effort, and to Dr. Vernon Troxel for his guidance in developing the statistical analysis.

Many persons are to be thanked for their efforts related to the experiment. These include: Dr. Charles Pickens, Mr. Kelvin Casebeer, Dr. Raymond McKellips, and Dr. Earl Reynolds.

Special gratitude is extended to my wife, Judy, whose confidence and understanding helped to make this study a reality.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

THE PROBLEM

During the past decade many changes and improvements have taken place in the field of education. One of the more significant changes has come from the new emphasis being placed on the mathematical preparation of elementary teachers. Some of the reasons for this new emphasis have been (i) the realization of the importance of mathematics education in general, (ii) the new programs being offered in secondary mathematics, and (iii) the new mathematics programs being offered at the elementary level. Today's elementary teachers must teach more mathematics, and do so in a more meaningful way, than have the elementary teachers of the past. Further, more elementary students go to high school and college than ever before; therefore, elementary teachers must be concerned with each student's understanding of mathematics as well as his computational skills. Elementary teachers, present and future, will not be able to meet present demands unless they are prepared in a more meaningful manner.

Recently, experiments have been carried out to provide basic materials for the continued investigation of problems in the improvement of elementary mathematics programs. Certain of these investigations indicate rather widespread evidence of the elementary school teachers' incomplete mastery of mathematics. As a result of the many studies, stress is now being placed on the need to find ways to remedy the situation rather than the gathering of additional data to reemphasize that elementary teachers are deficient in their mathematics preparation.

The careful preparation of prospective elementary teachers in mathematics subject matter is a prerequisite to an improved program in mathematics at the elementary school level. Therefore, the question of elementary teachers being fully prepared to teach today's elementary school mathematics is one that has been raised by many mathematicians and mathematics educators. There has been much written to support the fact that today's elementary teachers need to improve their basic knowledge and fundamental understanding of mathematics (2, p. 296), (10, p. 4), (28, p. 51).

In a study conducted by Glennon, Weaver, and Phillips (18) comparing mathematics facilities and understanding of elementary teachers in the United States and Canada it was reported that there was a need for reappraising elementary teacher education programs. The study involved data concerning the competency of elementary teachers in mathematics from Alberta in Canada, Illinois, and Massachusetts in the United States. This data indicated that the mathematical competence of prospective elementary teachers in Alberta was relatively higher than the mathematical competence of prospective elementary teachers in Illinois and Massachusetts.

The study concluded by suggesting that those individuals who are concerned about the relative mathematical attainments of students in the United States and elsewhere might find the solution to this problem lies in improving the preparation programs in mathematics for prospective elementary teachers.

Further, Melson (28) reported the results of a study based on a test of thirty-three items for grades one through six in modern elementary mathematics. This test was given to forty-one elementary teachers in September, 1963. The teachers tested had been graduated in June, 1963. The results showed the median score to be twelve correct responses out of thirty-three, 36 per cent; two of the forty-one teachers scored about 75 per cent, twenty-seven below 50 per cent, and twelve below 25 per cent. The most disturbing factor in this report was that all teachers involved indicated that they had successfully completed a course in modern mathematics. Melson pointed out that this indicated either inadequate preparation of the course or faulty mastery of it.

Two major sources of information for evaluating teacher competence are (i) an actual test of their knowledge (this was done in the above study), and (ii) a study of their self-judgment. Groff (21) chose the latter. He investigated the pre-service elementary teachers' self-judgment as to how well they felt they were prepared to teach elementary school mathematics. The teachers felt they were very well prepared in modern elementary mathematics. The results further pointed out that these future elementary teachers felt they were better prepared to teach arithmetic than anything else except reading. This finding, in conjunction with the previously reported results, indicates there is need to do further research that combines attitudes, self-evaluation, and content understanding.

Garstens (17) also stated that it was not necessary to point out to any group concerned with elementary education that an elementary teacher should have a background that is broader and deeper than the level at which he is teaching. Furthermore, educators must accept the obligation to develop appropriate mathematics courses for elementary education majors, courses that will be suitable, stimulating, and significant.

After reviewing the relevant literature, it became apparent that there was need to develop a method of instruction that will better prepare future elementary teachers in the fundamental concepts of elementary mathematics. The development of such a method offers a distinct challenge to those interested in mathematics education. One possible solution would be to supplement the present mathematics courses in such a way as to improve the future teachers' knowledge and understanding of basic concepts. One method of supplementation that has been suggested is the use of programed materials.

There has been much research on the use of programed material. Ripple (35) reported the results of a study at

Cornell University that compared learning through programed material with learning by what has been called "comparable" textbook material or "conventional" instruction. The groups tested in this study were selected from sophomores enrolled in the beginning psychology course at Ithaca College, Ithaca, New York, for the fall semester of 1964. Since these groups were carefully selected, no difference was expected or found between the groups on the pretest. Furthermore, no difference was found between the two formats on gains.

Research which indicates that programed material alone is not significantly better than the traditional lecture method has resulted in a trend toward the production of materials that can supplement and aid rather than replace classroom teachers in their daily instruction. Brown and Mayor (2) reported in their study that much research is needed on methods of instruction, improvement of teaching aids, and patterns of learning. Educators need to know a great deal more about developing special courses for teachers. Carr (31), Goodlad (20), and Popham (21) emphasized a lack of carefully executed experiments that unequivocally demonstrate the superiority of automated instruction, either programed text or machine, over the usual classroom procedures.

Many have speculated concerning the usefulness of programed material as a supplementary aid in instruction. The following statement by Stolurou is probably the strongest of these predictions:

> These devices (automated instruction) are here to stay. Future research will concern itself with important characteristics of the developments, a theory of teaching will emerge. The devices of the future will be either books (programed or scrambled) or computer based machines, small devices will drop out. The results of experiments in programed instruction suggest an impressive contribution to education, and, if the right programs can be developed and combined with economical and effective means of presentation, the applications of programed instruction will be widespread. (39, p. 85)

The use of programed materials is strongly urged by some educators. The following statements have been selected as they point strongly and directly to the aims of this paper.

> Experimental studies should be undertaken in order to determine what content material and what types of presentation provide teachers with the knowledge and understanding that is most valuable to them as teachers of elementary arithmetic. They need competence much more than they need factual recall. (37, p. 398)

> Specifically, we must consider programed instruction in proper perspective among other educational techniques and attempt to discover what combination of methods will lead to most efficient learning under specified conditions. Future research must be directed toward the discovery of optimal combinations of educational techniques for specific student and task characteristics. (8, p. 373)

> Unfortunately, indication of experimentation either in curriculum offerings or in methods used in teacher education could not be found in the research offerings from 1958-1963. Opportunities exist for experimentation with two new media in teacher education: (a) Programed instruction, and (b) closed circuit television. Much experimentation and research, especially with techniques other than questionnaire surveys, are needed in relation to all aspects of teacher education programs. (34, p. 377)

This last statement, accompanied by the fact that very little research was reported from 1963-1966 on improvements in the presentation of elementary mathematics to prospective elementary teachers, indicated a need for experimentation in this area.

A principal purpose of this study was to investigate potential ways to improve prospective elementary teachers' knowledge and understanding of elementary mathematics. A second purpose was to investigate whether or not the mastery of this mathematics was affected by the way it was taught at the undergraduate level.

The research problem was designed to determine whether or not undergraduate classes that were exposed to a combination of programed learning, lecture, and discussion could achieve greater understanding in elementary mathematics than undergraduate classes that received only the lecture form of instruction. A control group, which received no instruction in elementary mathematics, was used in order to evaluate the influence of maturation. Finally, if it had been determined that one of these elements in the teaching-learning process helped some students achieve significantly more than the others, then a partial solution might be available for use by those interested in increasing the supply of mathematically competent elementary teachers.

In summary, the literature seemed to emphasize the need to improve elementary teachers' mathematical background. Further, it appeared to emphasize that research should be directed toward discovering an effective combination of educational techniques that would lead the student to maximum understanding. The discovering of an effective combination of educational techniques would be a step toward developing a theory of instruction. This theory of instruction would be based on certain theories of learning that were commonly and currently accepted. Further, any theory of instruction would be prescriptive in that it would set forth rules concerning the most effective way of presenting knowledge. It would also be normative in that it would set up criteria and state the conditions for meeting them (7, p. 41). Finally, a theory of instruction would need to be concerned with how what one wishes to teach can best be learned; it also would need to be concerned with improving learning rather than describing it.

Since no comprehensive theory of learning was available, one had to make use of certain micro theories that did exist. Some of these theoretical frameworks were (i) Mitzel's paradigm, (ii) Smith's paradigm, (iii) Ryan's paradigm and (iv) Stone-Leavitt's paradigm (13, p. 121). These models have the following common characteristics: (i) a perceptual and cognitive process on the part of the teacher, (ii) action elements on the part of the teacher, (iii) perceptual and cognitive processes on the part of the learner, and (iv) action elements on the part of the learner. These common characteristics have contributed greatly to the theoretical design of the experiment referred to in this paper.

It was assumed that necessary conditions for effective learning should include (4, p. 308), (6, p. 40):

(i) An instructional situation that specifies the experiences that most effectively implant in the learner a predisposition toward learning, whether it be specific or general; one that produces motivation, directs perception, elicits responses, and provides supplementation (15, p. 276). Further, an effective instructional situation should provide pre-instruction procedures that many times enhance learning in a given situation (14, p. 260). It should provide situations that (i) require active student response rather than passive listening (14, p. 638), (ii) provide for a wide range of stimulating materials and situations as these usually increase the amount of learning (45, p. 300), and (iii) provide immediate and continuous supplementation (25, p. 541).

(ii) The specific ways in which a body of knowledge should be structured so that it can be readily grasped by the learner. The structure of a body of knowledge should be such that it has power for simplifying information, for creating or generating new propositions, and for increasing the manipulability of the knowledge. Structure is related to the status and ability of the learner and should enable the learner to grasp facts, principles, and inter-relationships. Therefore, the size of the steps in learning should be varied. If they are too small, general principles are not understood. If they are too large, specific facts and principles are overlooked or underestimated (14, p. 626). Also, since learning is developmental and is a process in which earlier learning greatly influences later learning, the structure of a body of knowledge should be spiral in nature and highly developmental (6, p. 504).

(iii) The most effective sequence of topics and/or methods of presentation. This proper sequence of topics or methods of instruction is essential to both the logical and psychological development of a body of knowledge. An effective sequence should not introduce new material until prior material is thoroughly consolidated (1, p. 506). New materials and new methods should have a derivative relationship with prior materials and methods for maximum learning (1, p. 507). Maintaining and improving desired responses increases learning (25, p. 542). Also, a mixture of prompted and unprompted trials is more effective than using complete prompting throughout (27, p. 345). Finally, practicing responses in varied conditions facilitates their establishment (28, p. 57).

(iv) A system of evaluation that specifies the nature and pacing of rewards and punishments. No teaching-learning situation is complete without proper evaluation.

A knowledge of results should come at a point when the learner is comparing the results of his tryout with some criterion of what he seeks to achieve (4, p. 315). Evaluation should be given periodically and frequently for effective learning (27, p. 355). Immediate feedback of results aids length of retention and transfer of learning to new situations (24, p. 208). Finally, the proper balancing of extrinsic and intrinsic evaluation aids the learning progress (8, p. 41).

In summary, effective instruction must provide many stages for learning. The sequence must provide an introduction and a motivation. There must be small steps which culminate as a "principle" which is enriched by the large step sequence. The total instructional program must attempt to evoke, maintain, supplement, and improve desired responses. Finally, the learner must be considered as an individual within a group of individuals.

Using these points and considerations, the following rationale is presented concerning the various experimental groups within the related research.

The first experimental group, the Lecture Program Discussion group (hereafter denoted the L. P. D. group), received the following method of instruction. Each new concept, or set of concepts, was first introduced through a lecture that was supplemented by a homework assignment that consisted of reading a certain number of frames from related programed materials. The concepts were then discussed in detail, by both students and instructor, at the next class meeting. This cycle was repeated throughout the entire course.

The total method of instruction applied to this group best fitted the above theoretical design for the following reasons. (i) It provided the best sequence of methods of presentation by introducing a concept, or set of concepts, through a well structured lecture. This introductory lecture provided any necessary pre-instructional procedures and created in the learner the proper predisposition toward the given concepts. (ii) The concepts were then immediately maintained and supplemented in a logical and sequential manner by use of the programed materials. These programed materials supplemented the general structure of the lecture by providing the small-step derivative type of structure that is necessary for understanding any body of knowledge. (iii) The concepts were then further supplemented and clarified by the succeeding discussion. (iv) The use of three distinct stages of instruction provided a wide range of materials and situations for the learner. (v) The programed materials and discussion provided situations where the learner could actively respond to the given concepts. (vi) The discussion provided a situation in which the instructor could evaluate the class' general understanding, and make possible postponing the introduction of new concepts when general understanding was not satisfactory. (vii) The programed materials provided an opportunity for the learner to continuously evaluate his understanding of the given lectures. This immediate feedback of results enhanced the length of retention and made the succeeding discussion more meaningful. (viii) The programed materials allowed for much individualization with respect to pacing.

The second experimental group, the Program Lecture Discussion group (hereafter denoted the P. L. D. group), received the following method of instruction. Each new concept, or set of concepts, was first introduced through programed materials. The learner read these materials prior to attending a given lecture. These concepts were then supplemented and enlarged upon by a related lecture. Finally, the programed materials and lecture were then discussed at the next class meeting. This cycle was then repeated throughout the entire course.

This method of instruction has many of the characteristics of the above method. However, it did not appear to be as complete as the L. P. D. method for the following reasons (i) It is difficult for programed materials, consisting of small steps, to give as complete a structural introduction to a set of concepts as can a lecture. Further, it is difficult for such materials to provide either the necessary pre-instructional procedures or proper predisposition toward the learning situation. (ii) It is difficult for the supplementation by the lecture to be as sequential and logical as that provided by the programed materials.

The third experimental group, the Lecture Textbook group (hereafter denoted the L. T. group), received the following method of instruction. Each new concept, or set of concepts, was introduced through a lecture. The assignment for the succeeding class was to solve a set of exercises from a related textbook. This method represented the traditional approach that has been and continues to be used at most colleges. There was no discussion unless a student requested the answer to, or an explanation of, a given exercise.

This method of instruction did not appear to be as complete as the L. P. D. or the P. L. D. method for the following reasons. (i) Supplementation, either positive or negative, is rarely immediate and is most often non-existent. (ii) Variations in the learning situation are minimal. (iii) Active student responses are virtually absent. (iv) Immediate and continuous evaluation on the part of both learner and instructor is rarely considered. (v) Individualization is very difficult to achieve.

The following hypotheses were deduced from the theory and rationale presented:

1. Those students involved in the L. P. D. method will show a significantly greater level of achievement and understanding in mathematics than those students involved in the P. L. D. method.

2. Those students involved in the L. P. D. method will show a significantly greater level of achievement and understanding in mathematics than those students involved in the L. T. method.

3. Those students involved in the P. L. D. method will show a significantly greater level of achievement and understanding than those students involved in the L. T. method.

CHAPTER II

THE EXPERIMENT

Introduction

The experiment was conducted at Oklahoma State University, Stillwater, Oklahoma and Southwestern State College, Weatherford, Oklahoma. The first semester of the 1966-67 academic year was selected for carrying on the experiment, the purpose of which was to evaluate the impact of various methods of instruction on achievement and understanding in mathematics for elementary teachers.

The classes chosen for the experimental part of the study were six sections of mathematics for elementary teachers. Four sections at Oklahoma State University and two sections at Southwestern State College were selected. No attempt was made to control enrollment in any of these sections. However, the samples may be assumed to be much like those sections of students that regularly enrolled in this particular course. The control group consisted of elementary education majors enrolled in Education 213 at Oklahoma State University during the fall semester of the 1966-67 academic year.

The instructors involved in the experiment were interested in the mathematical preparation of elementary teachers and were experienced classroom teachers.

The pretest, The Structure of the Number System (Form A) was administered to each group during the first week of the semester, in September. The posttest, The Structure of the Number System (Form B), was administered to each group during the last week of the semester, in January. All statistical analysis related to the experiment was completed by using the adjusted posttest results.

Subject Matter

The subject matter involved in the experiment is commonly referred to as modern mathematics for elementary teachers. Topics covered included set theory, the whole numbers, systems of numeration, fractions, the integers, the number line and its uses, and the rational numbers.

In the unit on set theory the following concepts were developed: set, set membership, set notation (including set-builder notation), set measurement (empty set, finite set, and infinite set), set relationships (equality, equivalence, nonequivalence, greater than, less than, disjointedness, subset, proper subset), universal set, complement set, set operations (union, intersection, complementation, cross-product, and partition), and set-operation properties (closure, commutativity, associativity, identity, and distributivity).

In the unit on whole numbers the following concepts were developed: number, number names, counting, counting numerals, place-value, expanded notation, addition, subtraction, multiplication, division, order, and ordinal numbers. The properties for the four operations (addition, subtraction, multiplication, and division) were also developed. These included closure, commutativity, associativity, identity, cancellation, and distributivity. Understanding of each property was reinforced by applying it in the solution of problems and mathematical proofs. All the above concepts were developed by relating them to an appropriate concept from set theory. For example, the foundations of addition were developed using the union of disjoint sets. Finally, the algorithms for each operation were developed in great detail.

In the unit on systems of numeration the important concepts from base ten were reviewed. During this review base ten was presented as a mathematical system consisting of ten basic symbols, a place-value principle, two primary operations (addition and multiplication), and two secondary operations (subtraction and division). The concept of grouping was developed and then used to illustrate that a given number idea may have many different symbolizations. The operations (addition, subtraction, multiplication, and division) were presented through the use of expanded notation and regrouping. This method added much to the meaning of each operation, and served to reinforce the understanding of the grouping procedure. Following each of these detailed presentations, the given algorithm was introduced and explained. For example, in base five (23 + 14) was presented in the following manner: 23 + 14 = (20 + 3) + (10 + 4) = (20 + 10) + (3 + 4) = 30 + (10 + 2) = (30 + 10) + 2 = 40 + 2= 42. Finally, the properties for each operation were discussed, and it was pointed out that these properties are independent of any given system of numeration.

Fractions were introduced by carefully defining a fraction through the use of set partitions. Following this, the concepts of unit fraction, ordered pairs, and equivalent fractions were developed by diagram and definition. The operations of addition, subtraction, multiplication, and division were illustrated by diagrams and then defined by mathematical equations. The properties for these operations (closure, commutativity, associativity, identity, multiplication inverse, and distributivity) were proved as theorems, which were based on previous definitions and whole number properties. For example; given that a, b, c, and d were whole numbers with b and d not equal to zero, commutativity for the addition of fractions was developed in the following manner: $a/b + c/d = (ad + bc)/bd = (da + cb)/db = (cb + da)/db = c/d + a/b$. Order was introduced ($a/b < c/d$ if and only if $ad < bc$) in such a manner as to enable the student to determine simple inequality and direction. Although not stated directly, this chapter introduced the student to the basic concepts involved in mathematical proofs.

The integers were developed by using ordered pairs of whole numbers. The concepts of equivalence, addition, and multiplication were defined and developed through the use of these ordered pairs. Also, the properties of addition and multiplication (closure, commutativity, associativity, identity, inverse, and distributivity) were proved as theorems based on ordered pairs. Subtraction and division were developed from the additive and multiplicative points of view. Next, the ordered pairs were defined in such a way, (a, b) is equivalent to (a - b), as to enable the student to interpret them as signed numbers. Finally, the various properties for signed numbers were proved by using these ordered pairs. For example, the proof that a negative integer multiplied by a negative integer is a positive integer was developed in the following manner: (o, x) and (o, y) are considered as negative x and negative y, and $(o, x) \cdot (o, y) = (o \cdot o + x \cdot y, o \cdot x + o \cdot y) = (xy, o)$ which is considered as positive xy.

The number line was introduced at this time as an aid in understanding ideas presented in the first five units. It was used to illustrate number facts, not to prove them. The number line was presented as an arbitrary line (usually horizontal) with an arbitrary point as the origin and an arbitrary unit of length for determining the position of each integer. Each of the four operations (addition, subtraction, multiplication, and division) was explained using whole numbers, integers, and fractions. Also, the properties for each of these operations were illustrated using both integers and fractions.

The unit on rational numbers was introduced by defining a rational number as an ordered pair of integers with the second element being positive. This definition was then used in defining an equivalence relation, addition, subtraction, multiplication, and division. The properties for these operations (closure, commutativity, associativity, identity, inverse, and distributivity) were developed as theorems based on the above definitions and the related properties from the integers. Definitions for order and density were given and many related theorems were proved. For example, it was shown that if $a/b < c/d$ then $(a/b + c/d)/2$ was between a/b and c/d by showing $a/b < (a/b + c/d)/2$ and $(a/b + c/d)/2 < c/d$. The final topic in this unit was decimals. Included under this topic were the following concepts: numerator, denominator, basic units, place-value, expanded notation, exponents and the rules for operating with exponents, converting rational numbers to terminating or repeating decimals, and converting terminating or repeating decimals to rational numbers.

Methods of Instruction

Three methods of instruction were employed in the experiment. They were (i) the Lecture Program Discussion method, (ii) the Program Lecture Discussion method, and (iii) the Lecture Textbook method.

The L. P. D. method was a three step method of instruction. Each new concept, or set of concepts, was first introduced through a lecture. The number of concepts developed in a given period varied in relation to the complexity of the given concepts. The lecture was then supplemented by a homework assignment that consisted of reading a certain number of frames from related programed materials. The concepts were then thoroughly discussed at the next class meeting. This cycle was repeated throughout the entire course.

The lecture presented essentially the same content as was to be assigned in the programed materials. Each lecture began with a brief overview of the concepts to be presented. Then, the individual facts, principles, and examples were structured in such a way as to put them in proper perspective with regard to the total unit. The lecture was then summarized by reviewing the concepts just presented. Finally, the instructor concluded by making suggestions that would aid the student in his reading of the programed materials.

The programed materials were structured to add the small-step logic and sequence that was necessary for developing more complete understanding of concepts presented in the lecture. The number of frames needed to develop a given concept depended upon the complexity of the concept. There were approximately forty to forty-five frames assigned for each class meeting.

The discussion period provided time for each student to ask questions, make comments, and attempt generalizations whenever possible. It also provided an opportunity for the instructor to make comments, ask probing questions, and pass subjective judgment on general class understanding.

Once the cycle (lecture, programed materials, and discussion) was set in motion it appeared that fifteen to twenty minutes was sufficient for each discussion period. Therefore, each class meeting consisted of fifteen to twenty minutes of discussion and thirty to thirty-five minutes of lecture. This is illustrated by the following diagram:

LECTURE ⟶ FRAMES ⟶ DISCUSSION, LECTURE ⟶ FRAMES ⟶

The actual subject matter was contained in a programed text consisting of seven chapters. Each chapter was completed in approximately two weeks. There were one hour examinations at the end of chapters two, four and six. The last examination was two hours, and it was cumulative. There were no unannounced quizzes. The distribution of class periods for each of the first three examination intervals was (i) ten periods for discussion and lecture, (ii) one period for review, (iii) one period for the examination, and (iv) one period for explaining the examination. The last examination interval consisted of six discussion-lecture periods, two review periods (one for chapter seven and one cumulative), and one final examination period.

The P. L. D. (Program Lecture Discussion) method was also a three step method of instruction. Each new concept, or set of concepts, was first introduced through programed materials that were read prior to attending a given lecture.

Again the number of concepts developed varied in relation to the complexity of the given concepts. These programed materials were then supplemented by a related lecture. The concepts were then thoroughly discussed at the next class meeting. This cycle was repeated throughout the entire course.

The programed materials, having been read before the lecture, not only provided for the student the small-step logic and sequence, but they also provided a thorough preview of the succeeding lecture.

Each lecture was prepared in advance and presented essentially the same content as was contained in the programed materials. However, the students were allowed to present questions and reactions prior to the actual lecture. This was done in order to enable the instructor to adjust his lecture in such a way as to satisfy existing questions and reactions. If no questions or reactions were presented, the instructor presented a few of his own in order to motivate the students toward the succeeding lecture. For example, he (the instructor) might motivate the students toward the properties of addition in fractions by reviewing the properties of addition in the whole numbers. Each lecture was presented in the following pattern: (i) a brief overview of the topics contained in the programed materials, (ii) a structured presentation in which the individual facts, principles, and examples were put in proper perspective with regard to the total unit, and (iii) a summary that attempted to completely interrelate the lecture and the programed materials.

The discussion period again provided time for the students to ask further questions, make comments, and attempt generalizations whenever possible. It also provided time for the instructor to make comments, ask probing questions, and pass subjective judgment on general class understanding.

Once the cycle was set in motion it was found that thirty to thirty-five minutes was sufficient for each lecture. Therefore, each class meeting consisted of thirty to thirty-five minutes of lecture, and fifteen to twenty minutes of discussion devoted to interrelating the programed materials and the lecture. This is illustrated by the following diagram.

FRAMES → LECTURE, DISCUSSION → FRAMES → LECTURE, DISCUSSION

The subject matter and programed text for this method was the same as that of the L. P. D. method. Each chapter was completed in approximately two weeks. There were one hour examinations at the end of chapters two, four, and six. The last examination was two hours, and it was cumulative. There were no unannounced quizzes. The distribution of class periods for each of the first three examination intervals was (i) ten periods for lecture and discussion, (ii) one period for review, (iii) one period for the examination, and (iv) one period for explaining the examination.

The last examination interval consisted of six lecture-discussion periods, two review periods (one for chapter seven and one cumulative), and one final examination period.

The programed material employed in these two methods, Basic Mathematics, A Programed Introduction (19) is unique. It is neither linear programing nor branch programing. It is a hybrid form of programing that combines both the linear and branch forms. This combination was accomplished in the following manner: (i) a series of Skinner-type frames that are single response, completion statements (these statements usually require less thinking on the part of the reader than the Crowder-type frames), (ii) a Crowder-type frame which is a multiple-choice statement (this frame usually requires some thinking or generalizing on the part of the reader), and (iii) a repetition of parts (i) and (ii). The number of Skinner frames between Crowder frames ranged from five to fifteen.

The L. T. (Lecture Textbook) method was the conventional method of instruction found in many colleges. The lecture presented (i) a general overview of the concepts and how they were related to the past material, (ii) the main body which consisted of a sequence of facts, principles, and examples that were in proper perspective with regard to the total unit, and (iii) a structured summary that reviewed previous material, related it to the presented material, and related both to future material.

The assignment following each lecture consisted of solving problems, from the accompanying textbook, that were related to the lecture.

These problems were not collected and were not discussed unless a student requested a solution or explanation. This cycle was repeated throughout the course. It is illustrated by the following diagram.

LECTURE → RELATED PROBLEMS → LECTURE → RELATED PROBLEMS →

The actual subject matter was contained in a standard textbook consisting of seven chapters. These chapters were essentially the same as those of the programed material that was employed in the L. P. D. method and the P. L. D. method. Each chapter was completed in approximately two weeks. There were one hour examinations after chapters two, four and six. The last examination was two hours, and it was cumulative. There were no unannounced quizzes. The distribution of class periods for each of the first three examination intervals was (i) ten periods for lecture, (ii) one period for review, (iii) one period for the examination, and (iv) one period for explaining the examination. The last examination interval consisted of six lecture periods, two periods for review (one for chapter seven and one cumulative), and one final examination period.

The textbook for the L. T. method was Today's Mathematics (23). This is a standard textbook with the added feature of practical classroom applications. When each concept is developed, the authors immediately illustrate where and how it can be applied at the elementary school level.

The control group received no instruction in mathematics for elementary teachers.

Evaluation Instruments

The instruments that were used to measure the levels of achievement that resulted from the various methods of instruction were: (i) American College Test in Mathematics (A.C.T.), (ii) The Structure of the Number System (Form A), and (iii) The Structure of the Number System (Form B).

The A.C.T. Mathematics Test was developed by the American College Testing Program. It is a mathematical aptitude test that is considered to be a good predictor of future achievement in college mathematics (3, p. 9). The test consisted of thirty-six multiple choice questions that sampled aptitudes related to pre-college mathematics. The results of this test were used as one of the two covariates in the statistical analysis of the posttest results.

The Structure of the Number System (Form A) was produced by Educational Testing Service, Cooperative Mathematics Tests Division. This test is an achievement test that measures understanding of the real number system up to the rational numbers. The test consisted of forty multiple choice questions that sampled the following topics: arithmetic judgment, operational properties (closure, commutative, associative, and distributive), inverses and identities, properties of the integers, place value, (factors, divisors, and multiples), prime numbers, number lines, zero denominator, number systems (bases other than ten), modular arithmetic, and Roman numerals. This test was used as a pretest in the experiment, and the results were used as one of the two covariates in the statistical analysis of the posttest results.

The Structure of the Number System (Form B) was also produced by Educational Testing Service, Cooperative Mathematics Tests Division. It is also an achievement test that measures understanding of the real number system up to the rational numbers. The test consisted of forty multiple choice questions and was used as the posttest in the experiment. Form B is considered an alternate form of Form A and, thus, covered the same topics as Form A.

The two Number Systems tests were designed by the Educational Testing Service staff and some forty-six high school and college mathematics teachers. The tests were pretested throughout the country in May, 1960. After analyzing the results, they were revised in May, 1961 and re-pretested in May, 1962. The results from the second pretesting indicated the tests were appropriate for the intended population.

These two tests were selected because they were the only commercially produced tests directly related to the objectives of the experiment. They stress understanding of facts, principles, and relationships, and do not emphasize computational skills. Furthermore, the tests are measures of developed abilities, and thus their content validity is very important. Educational Testing Service feels (47, p. 62) they have insured this by entrusting test construction to persons well-qualified to judge the relationship of test content to teaching objectives. The reliabilities reported by E.T.S. are measures of internal consistency, computed by using the Kuder-Richardson Formula 20. The reliability of Form A was .86 with a standard error of measurement of 2.73. The reliability of Form B was .84 with a standard error of measurement of 2.75. The correlation of Form A with the SCAT-Quantitative Test was .78, and that of Form B was .74. Educational Testing Service pointed out (47, p. 64) that this was lower than expected, but this was due to the fact that Forms A and B measure understanding while the SCAT-Quantitative emphasizes computational skills. Form A had an item-total score discrimination correlation of .50 and that of Form B was .48. These results indicate that the tests are effective in discriminating between high and low ability students (47, p. 64). Finally, the equivalence of these two alternate forms was very good. The converted raw scores differed by no more than two at all levels of performance. These results are tabulated in the Educational Testing Service mathematics booklet (47, p. 67).

Sample

The sample for this study consisted of one hundred forty-one undergraduate students. Ninety-six of these were enrolled at Oklahoma State University, and forty-five were enrolled at Southwestern State College. The experimental groups were distributed in three sections in the following manner: (i) forty-seven students (L. P. D. method, Oklahoma State University), (ii) forty-five students (L. T. method Southwestern State College), and (iii) thirty-three students (P. L. D. method, Oklahoma State University). The control group consisted of sixteen students from Oklahoma State University. Any student who did not complete the course or any student for whom it was impossible to obtain related data was not included in the sample analysis. In the P. L. D. group eleven dropped the course, and four were discarded due to lack of related data. In the P. L. D. group nine dropped the course, and four were discarded for lack of related data. In the L. T. group fifteen dropped the course, and seven were discarded for lack of related data. In the control group three dropped the course, and three were discarded for lack of related data. All students involved in the study were elementary education majors.

The L. P. D. group had a mean score of 19.32 on the A.C.T. mathematics test. This test had a possible score of thirty-six. This group also had a mean score of 19.32 on the pretest. This test had a possible score of forty.

The P. L. D. group had a mean score of 18.30 on the A.C.T. mathematics test and a mean score of 18.30 on the pretest.

The L. T. group had a mean score of 16.40 on the A.C.T. mathematics test and a mean score of 16.22 on the pretest.

The control group had a mean score of 17.19 on the A.C.T. mathematics test and a mean score of 18.38 on the pretest.

An analysis of covariance (44, p. 18) indicated there were significant differences among the mean scores of these four groups on both the A.C.T. mathematics test and the pretest.

Other data such as I. Q. scores and personality test evaluations were not available for the groups.

Analysis

Each group was administered the pretest, The Structure of the Number System (Form A), during the first week of the semester in September, 1966. The posttest, The Structure of the Number System, (Form B), was administered during the last week of the semester in January, 1967. The data that were used to test the hypotheses were the A.C.T. mathematics test scores, the pretest scores, and the posttest scores.

Since there were significant differences between the mean scores of the groups on both the A.C.T. mathematics test and the pretest, analysis of covariance was employed in comparing the groups on the posttest results. Most authors (16), (26), (11), in explaining the application of the analysis of covariance, let the covariant score represent a pretest score. In this analysis, the pretest score was used as one covariable, but the A.C.T. mathematics score was also used as a covariable for the dependent variable, the posttest score. Garrett (16, p. 295) explains the use of analysis of covariance when he states:

> Analysis of covariance represents an extension of the anlaysis of variance to allow for the correlation between initial and final scores. Covariance anlaysis is especially useful for experiments in the behavioral sciences where for various reasons it is impossible or quite difficult to equate control and experimental groups at the start: a situation which often obtains in actual experiments. Through covariance analysis one is able to affect adjustments in final or terminal scores which will allow for differences in some initial variable.

Further analysis for comparing adjusted individual means used Tukey's procedure for comparing individual means (13, p. 330) which consisted of (i) testing for a significant gap, (ii) testing for a "straggler," and (iii) testing for excessive variability. This procedure allows an experimenter to draw as many conclusions as are reasonable about differences that are present among means. The basic plan of the procedure is to classify the means into groups that are alike among themselves but differ from each other.

CHAPTER III

ANALYSIS OF THE DATA

Introduction

This chapter contains the findings of the statistical tests used to determine the significance of the results of this investigation. The .05 level of probability was used to judge the significance of all statistical tests. The rejection of any hypothesis was directed; therefore, one-tailed tests of significance were employed. The four major headings, which represent the four major statistical analyses, will be (i) multiple analysis of covariance--four groups, (ii) test for a significant gap, (iii) test for a "straggler", and (iv) test for excessive variability. A summary of the results will follow the presentation of the statistical analyses.

Multiple Analysis of Covariance--Four Groups

This statistical technique is a combination of analysis of variance and multiple regression techniques. The method enabled the writer to statistically equate the means of the groups with respect to the covariates before drawing conclusions about treatment effect. Further, this statistic allowed the writer to control for differences in A.C.T.

mathematics test scores and differences in pretest scores while comparing differences exhibited on the posttest scores.

The data for the three experimental groups and the one control group were prepared for an IBM 1620 computer system. The Northern Oklahoma College Computer Center provided a fortran program for analyzing this data. The multiple analysis of covariance program of Winer (44, p. 618) was utilized. This program calculated the F ratio for the adjusted treatment means, the Beta coefficients and their standard errors, and the adjusted treatment means with their accompanying standard errors. The findings concerning these four groups (L. P. D, P. L. D, L. T., and control) are presented in Table I.

Table I

ANALYSIS OF COVARIANCE--FOUR GROUPS

Source of Variation	Adjusted Sum of Squares	df	Mean Sum of Squares	F
Treatments	255.62	3	85.21	
Error	1776.07	135	13.16	6.48
Total*	2031.69	138		

*Covariates were the A.C.T. mathematics test scores and the pretest (Structure of the Number System (E.T.S.), Form A) scores.

From Table I, the calculated F value is shown as 6.48 correct to two decimal places. The critical F value, for the given degrees of freedom, was 3.23. This result disclosed the fact that significant differences existed among the four groups on the adjusted posttest results. However, since the control group was included in the analysis, significant differences were expected. Continuing, this result indicated that one or more of the following conditions may be observed (13, p. 101): (i) there is a wide gap between adjacent means when they are arranged in order of magnitude; (ii) one of the means is a "straggler"; and (iii) the means taken as a group show excessive variability.

The method selected for further analyzing the adjusted posttest results was Tukey's procedure for comparing individual means (13, p. 330). This method classifies the means into groups that are alike among themselves but differ from each other. There were three basic subdivisions in this procedure. They were (i) testing for a significant gap, (ii) testing for a "straggler", and (iii) testing for excessive variability.

Test for a Significant Gap

The first step in this test was to arrange the adjusted posttest means for the four groups in order of magnitude as shown in Table II.

TABLE II

ADJUSTED MEANS ARRANGED IN ORDER OF MAGNITUDE

	Experimental Conditions			
	Control	L. T.	P. L. D.	L. P. D.
Adjusted $\overline{Y}$	20.59	24.26	25.40	25.76

The statistic used in this test was given by the formula

$$\text{Significant gap} = (t_{.05})\ (\sqrt{2})\ (S_{\bar{x}})\ ,$$

where $S_{\bar{x}}$ was the standard error of the mean, and $t_{.05}$ was the tabulated value of t at the 5 per cent level for the degrees of freedom associated with the mean square of the error from Table I.

For the data of Table I, t at the 5 per cent level for 135 degrees of freedom was 1.98. $S_{\bar{x}}$ was 1.24 and was calculated by the following formula,

$$S_{\bar{x}} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_1^2}{n_2} + \frac{S_3^2}{n_3} + \frac{S_4^2}{n_4}}\ ,$$

where S_i^2 was the adjusted mean square of the error for each group and n_i was the number of subjects in each group. Substituting in the first formula with these values, it was found that

$$\text{Significant gap} = (1.98)\ (1.14)\ (1.24) = 3.46$$

Inspecting the differences between the adjacent pairs of means from Table II, it was found that the gap between 20.59 and 24.26 was a significant gap. Thus, the significant gap test had divided the four adjusted means into two subgroups: the control group by itself, and the three experimental groups (L. P. D., P. L. D., and L. T.) as another subgroup.

Test for a "Straggler"

Tukey's second test was employed in order to determine if there was a "straggler" among the three means found in the second subgroup of the above test. The "straggler" test involved finding: (i) the adjusted grand mean $\bar{Y}$ for the L. P. D., P. L. D., and the L. T. groups, (ii) the most straggling mean $\bar{Y}_1$, and (iii) the differences between these two divided by $S_{\bar{x}}$ which was the standard error of the mean for these three groups.

The statistic to be computed was the z statistic. It was computed by the formula (10, p. 332)

$$z = \frac{\dfrac{\bar{Y}_1 - \bar{Y}}{S_{\bar{x}}} - \dfrac{1}{2}}{3\left(\dfrac{1}{4} + \dfrac{1}{df}\right)} .$$

$\bar{Y}_1$ was 24.26 (the L. T. adjusted mean), and $\bar{Y}$ was 25.14. The value of df was the number of degrees of freedom associated with the mean square of the error within the groups. Substituting in the formula with these values, it was found that

$$z = \frac{-1.07 - .50}{3\ (.26)} = -2.04$$

Since the value of a significant z at the 5 per cent level is -1.96 and $-2.04 < -1.96$, then the L. T. group is separated from the L. P. D. group and the P. L. D. group.

Test for Excessive Variability

This third test was used to determine whether there was excessive variability in the remaining subgroup consisting of the L. P. D. and the P. L. D. groups. The statistic used in this subgroup was the *F* statistic. The sum of the squares of the deviations of the adjusted individual means $\overline{Y}_i$ from the adjusted grand mean $\overline{Y}$ was found. Dividing this by one less than the number of means involved yielded an estimate of the variance of the means within the group. The following formula,

$$F = \frac{\dfrac{\sum\left(\overline{Y}_i - \overline{Y}\right)^2}{K - 1}}{S^2_{\overline{x}}},$$

was calculated where k is the number of means in the new subgroup, and $S^2_{\overline{x}}$ is the square of the standard error. The individual means were 25.40 and 25.76, $\overline{Y}$ was 25.58, and $S^2_{\overline{x}}$ was .30. The degrees of freedom for evaluating *F* of the formula were (k-1) for the numerator and for the denominator the degrees of freedom were those associated with the mean square of the error. Substituting in the formula with these values, it was found that

$$F = \frac{\dfrac{.06}{1}}{.30} = .20$$

Since the critical *F* value for the given degrees of freedom was 1.99, it was found that no excessive variability existed between the means of the L. P. D. group and the P. L. D. group.

Summary of the Results

The purpose of this section is to summarize the results of the statistical analyses carried out in conjunction with the three hypotheses of the experiment. The final conclusions, limitations, and recommendations are presented in chapter four.

The analysis comparing the three experimental groups and the control group, when considering the total number of 141 subjects, disclosed the fact that significant differences existed among the four groups on the adjusted posttest results. The calculated F value was 6.48 while the critical F value, for the given degrees of freedom, was 3.23. This result was expected as the control group, which received no instruction, was included in the analysis. Finally, this result necessitated further anlaysis in order to determine just where the significant differences existed.

The analysis selected was Tukey's procedure for comparing individual means. This procedure allowed the writer to classify the adjusted posttest means into groups that were alike among themselves but differing from each other. The procedure consisted of three statistical analyses: (i) the test for a significant gap, (ii) the test for a "straggler", and (iii) the test for excessive variability.

The test for a significant gap consisted of finding a gap that could be used in determining whether gaps between the adjacent adjusted posttest means (Table II) were large enough to be considered significant. The significant gap was found to be 3.46. Upon inspecting the differences between adjacent pairs of means in Table II, it was found that there was a significant gap between the mean for the control group and the mean for the L. T. group. Therefore, this result enabled the writer to divide the four means into two subgroups: one containing just the control mean and one containing the L. P. D. mean, the P. L. D. mean, and the L. T. mean.

The test for a "straggler" was used to investigate the subgroup formed by the test for a significant gap. This test was employed for investigating the subgroup consisting of the L. P. D. mean, the P. L. D. mean, and the L. T. mean. The purpose of the test was to determine whether one of these three means was an excessive "straggler" when compared to the total subgroup. The statistic used in this test was the z statistic, and it was found to be a -2.04 when applied to the L. T. mean (the most straggling mean of this subgroup). The critical value for z at the 5 per cent level was -1.96; therefore, the L. T. mean was accepted as an excessive "straggler" when compared with the L. P. D. mean and the P. L. D. mean.

The test for excessive variability was employed in order to determine whether there was excessive variability in the remaining subgroup consisting of the L. P. D. mean and the P. L. D. mean. The calculated F value was .20 while the critical F value, for the given degrees of freedom was

1.99. This result indicated there was no excessive variability between the L. P. D. mean and the P. L. D. mean.

In summary, these three tests indicated the following: (i) The control group mean (20.59) was significantly smaller than the L. T. mean (24.26). (ii) The L. T. mean was significantly smaller than the P. L. D. mean (25.40) and the L. P. D. mean (25.76). (iii) There was no evidence of significant variability between the L. P. D. mean and the P. L. D. mean.

These findings allowed the writer to (i) accept hypotheses two and three of chapter one, and (ii) reject hypothesis one.

The following points, though not included in the statistical analysis or related to the hypotheses, may be of interest to the reader. (i) There were six male and one hundred nineteen female students enrolled in the experimental groups. The male students had a mean score of 18.16 on the A.C.T. mathematics test, 16.50 on the pretest, and 20.00 on the posttest. The female students had a mean score of 18.01 on the A.C.T. mathematics test, 17.95 on the pretest, and 25.17 on the posttest. These results indicated that the female sample scored significantly higher on the posttest than did the male sample. However, one must be careful in generalizing these results as the male sample was extremely small. (ii) The three experimental groups contained twenty-eight freshman, seventy-two sophomores, and twenty-five juniors. The freshmen had a mean score of 17.61 on the

A.C.T. mathematics test, 17.96 on the pretest, and 25.68 on the posttest. The sophomores had a mean score of 18.12 on the A.C.T. mathematics test, 18.22 on the pretest, and 25.36 on the posttest. The juniors had a mean score of 17.00 on the A.C.T. mathematics test, 17.21 on the pretest, and 22.70 on the posttest. Although there were no significant differences among the adjusted posttest means, the three classes were ranked in the following descending order with respect to these adjusted means: freshmen, sophomores, juniors.

CHAPTER IV

SUMMARY, LIMITATIONS, AND CONCLUSIONS

Summary

The primary purpose of this study was to investitage experimentally the comparative effectiveness of three experimental methods of instruction at the undergraduate level of education. The subject matter that was presented through these methods was mathematics for prospective elementary teachers.

Two of the three experimental methods (the L. P. D. method and the P. L. D. method) were three-step methods that employed lectures, programed materials, and discussions. These particular methods were selected as they seemed to best satisfy the many assumptions considered necessary for effective learning. These assumptions were selected from a review of the writings of various psychologists who are considered to be authorities in the field of learning theory. The third experimental method of instruction (the L. T. method) employed only one step in the instructional procedure, the lecture.

One hundred forty-one undergraduate students enrolled in elementary education were used as subjects in this experiment. Ninety-six of these were students at Oklahoma State University and forty-five were students at Southwestern State College.

The basic design of the experiment was pretest-treatment-posttest. The pretest was administered to all subjects during the first week of the semester in September, 1966. The treatments (the methods of instruction) were applied three times per week for the entire semester. The posttest was administered to all subjects that completed the course during the last week of the semester in January, 1967. A control group (which received no instruction in mathematics for elementary teachers) was part of the experiment in order to evaluate the effectiveness of all methods of instruction.

The independent variables were the three methods of instruction: the L. P. D. method, the P. L. D. method, and the L. T. method. The dependent variables were the adjusted scores of these groups on the posttest.

Evaluation of the instruction was accomplished through the use of commercially made tests. The pretest and the posttest (The Structure of the Number System, Forms A and B) were produced by the Educational Testing Service, Cooperative Mathematics Tests Division. These tests were alternate forms and were used to measure the achievement of the subjects after one semester of mathematics for elementary teachers. The A.C.T. mathematics tests were produced by the American College Testing Program. These tests are aptitude tests, and the results were used as one of the two covariates in the statistical analysis. The pretest results were used as the other covariate.

There were two major statistical analyses in the experiment, the analysis of covariance and Tukey's procedure for comparing individual means. The analysis of covariance was used in analyzing all four groups in order to determine if there were significant differences between the groups. This statistic was selected as it allowed the writer to draw conclusions about treatment effect after variables that affected the observation were adjusted statistically. Tukey's procedure for comparing individual means was selected as it allowed the writer to separate the four groups into subgroups that were alike among themselves but differing from each other. The procedure consisted of three separate tests (i) the test for a significant gap, (ii) the test for a "straggler", and (iii) the test for excessive variability.

Limitations

Before considering the conclusions, it seems important to point out some conditions that may cast limitations on the findings. The reader should be aware of these limitations so that any tendency to overinterpret or over generalize may be reduced.

First, the reader should be cautioned that the sample was not necessarily a representative sample of elementary education majors as they were not randomly assigned. Further, the sample subgroups were different with respect to the A.C.T. mathematics tests scores and the pretest scores. However, these differences were statistically controlled by employing the analysis of covariance.

Second, the writer recognizes the limitations introduced by having the samples at different locations. Different campuses and different educational environments may have affected the results.

Another consideration in interpreting the results of this investigation is the Hawthorne effect (46). The experimental groups realized they were part of a study, and this may have affected the results.

Finally, each subject was administered a pretest and a posttest. Therefore, the effects of taking the pretest may have affected the posttest results. However, this effect was controlled to some extent as each group had four examinations during the semester.

Conclusions

The evidence resulting from the analysis of the data appears to support several conclusions.

First, there were significant differences among the adjusted posttest means of the four groups. This conclusion was accepted as a result of the analysis covariance--four groups.

Second, all three methods of instruction allowed their respective groups to raise the level of their mathematical achievement and understanding. This was indicated by the significant gap test which pointed out that a significant gap existed between the adjusted posttest mean for the control group and the adjusted mean for the L. T. group. However, the reader is cautioned not to extend this result to individual subjects as each method of instruction contained subjects that had either zero gainscores or negative gainscores.

Third, hypothesis two of chapter one was accepted. That is, the students involved in the L. P. D. method showed a significantly greater level of achievement and understanding in mathematics than those students involved in the L. T. method. This conclusion was accepted as a result of the "straggler" test. The results of this test indicated that the adjusted posttest mean of the L. T. group was significantly less than the adjusted posttest mean of the L. P. D. group.

Fourth, hypothesis three of chapter one was accepted. That is, the students involved in the P. L. D. method showed a significantly greater level of achievement and understanding in mathematics than those students involved in the L. T. method. This conclusion was also accepted as a result of the "straggler" test. Again, the results indicated that the adjusted posttest mean of the L. T. group was significantly less than the adjusted posttest mean of the P. L. D. group.

Fifth, the students involved in the L. P. D. group did show, as predicted, a greater level of achievement and understanding in mathematics than did the students in the P. L. D. group. However, this level of achievement was not significantly greater. Therefore, hypothesis one of chapter one was not accepted.

No unified science of learning exists, and, therefore, any application of learning theory to a particular method of instruction is difficult (11, p. 25). Furthermore, the fact that these two methods were highly similar and differed only with respect to the order of motivation and supplementation made predicting their relative effectiveness even more difficult for several reasons: (i) Motivational variables are perhaps the most elusive concepts with which psychologists have worked (11, p. 38). (ii) The effects of supplementation on learning are complicated in that they serve not only to confirm preceding behavior but also maintain the motivational level (11, p. 34). (iii) Desire to succeed has been a highly dependable source of motivation for learning (11, p. 37). (iv) Once learning has proceeded to a certain level of proficiency, so that the desired behavior is dominant, it may be that the nature of any motivation makes little difference; any source of motivation may sustain the performance (11, p. 38).

These statements seem to imply the following: (i) In each method the lecture and the programed material may have been performing dual roles (motivation and supplementation). (ii) In each method the desire to succeed may have reduced the effects of the lecture or the programed material as a motivator. (iii) After some point in the course, the nature of the motivation may have made little difference.

Therefore, with the above two paragraphs and the fact that each method included a follow up discussion period, it was not possible to adequately measure the effectiveness, if any, brought about by interchanging the lecture and the programed material. The measuring instruments measured only final changes in behavior, they did not measure any conditions that existed during the experiment.

In sum, the writer found no evidence for rejecting the theoretical design, however, under the conditions of this experiment, he could not accept the statement that the L. P. D. method was the best fit to the theoretical design.

Sixth, conclusions three and four indicated a more general conclusion concerning methods of instruction. These results seemed to indicate that a method of instruction which consisted of many phases was significantly more effective than a method of instruction that consisted of only one phase.

Recommendations

The fact that the L. P. D. group and the P. L. D. group achieved a significantly higher level of achievement and understanding than the L. T. group encourages the writer to recommend that future research might investigate these methods of instruction with much larger groups. Groups of one hundred or more would be appropriate for carrying on this further investigation.

A second recommendation is that an experiment, similar in nature, be conducted in which (i) it is possible to evaluate each step in the instructional sequence, and (ii) the instruments for evaluation are designed to evaluate the actual subject matter under consideration. Such an experiment should enable future research to make more specific conclusions regarding the effectiveness of a particular instructional sequence.

A further recommendation would be to apply the theoretical design, or one highly similar to it, to other areas of mathematics. If it has success in these areas, then experimentation might be carried on in other subject matter areas.

Finally, it is recommended that other methods of instruction be developed and investigated. For example, one that combines lectures, programed materials, discussion periods, and problem sessions or laboratory periods might be investigated.

Such research is recommended as it might enable future research to make conclusions concerning the feasibility of adding an additional step to the instructional sequence.

SELECTED BIBLIOGRAPHY

(1) D. P. Ausubel. et al., "Meaningful Learning and Retention: Interpersonal Cognitive Variables." Review of Educational Research, XXXI (December, 1961), 500-510.

(2) Brown, J. A. and J. K. Mayor, "The Academic and Professional Preparation of Teachers of Mathematics." Review of Educational Research. XXVII (October, 1957), 296-301.

(3) Buros, O. K., ed., The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1965.

(4) Bruner, J. S. "Some Theories on Instruction Illustrated with Reference to Mathematics." National Society for the Study of Education: Sixty Third Yearbook, 1964, 306-336.

(5) Bruner, J. S. "The New Educational Technology." American Behavioral Scientist, VI (November, 1962), 5-7.

(6) Bruner, J. S. The Process of Education. Cambridge, Massachusetts: Harvard University Press, 1960.

(7) Bruner, J. S. Toward A Theory of Instruction. Cambridge, Massachusetts: Harvard University Press, 1966.

(8) Coulson, J. E. "Programed Instruction: A Perspective." Journal of Teacher Education, XIV, (December, 1963), 372-378.

(9) DeCecco, J. P., ed., Human Learning in the School. New York: Holt, Rinehart and Winston, 1963.

(10) DeVault, V. M. Improving Mathematics Programs. Columbus, Ohio: Charles E. Merrill Books, Inc., 1959.

(11) Dixon, W. J. and F. J. Massey. Introduction to Statistical Analysis. New York: McGraw Hill Book Company, 1957.

(12) W. Dutton. et al., "Background Mathematics for Elementary Teachers: National Council of Teachers of Mathematics: Twenty Fifth Yearbook, 1960.

(13) Edwards, A. L. Statistical Methods for the Behavioral Sciences. New York: Rinehart and Company, Inc., 1959.

(14) Gage, N. L. "Instruments and Media of Instruction" Handbook of Research on Teaching. Chicago: Rand McNally Co., 1961, 605-655.

(15) Gage, N. L. "Theories of Teaching." National Society for the Study of Education: Sixty Third Yearbook, 1964, 268-285.

(16) Garrett, H. E. Statistics in Psychology and Education. 5th ed. New York: David McKay Company, Inc., 1958.

(17) Garsten, H. L. "Mathematics and Elementary Education Majors." The Arithmetic Teacher, XII (December, 1964), 540-542.

(18) Glennon, V. J., P. L. Weaver, and J. W. Phillips, "Mathematical Competence of Prospective Elementary Teachers in Canada and the United States." The Artithmetic Teacher, VIII (April, 1961), 147-150.

(19) Goff, G. K. and M. E. Berg. Basic Mathematics, A programed Introduction. New York: Appleton, Century, Croft, 1967.

(20) Goodlad, J. O. "An Analysis of Professional Laboratory Experiences in the Education of Teachers." Journal of Teacher Education, 1965, 263-270.

(21) Groff, P. J. "Self-Estimates of Ability to Teach Arithmetic." The Arithmetic Teacher, X (December, 1963), 479-480.

(22) Guthrie, E. R. The Psychology of Learning. Rev. Ed. New York: Harper, 1952.

(23) Heddens, S. J. Today's Mathematics. Chicago: Science Research Associates, 1965.

(24) Hill, W. E. Learning: A Survey of Psychological Interpretations. San Francisco: Chandler Publishing Co. 1963.

(25) Krumboltz, J. D. "Meaning, Learning, and Retention: Practice and Reinforcement Variables." Review of Educational Research, XXXI (December, 1961), 535-546.

(26) Linquist, E. F. Statistical Analysis in Education. Boston: Houghton Mifflin Company, 1940.

(27) Lumsdaine, A. A. "Educational Technology, Programed Learning, and Instructional Science." National Society for the Study of Education: Sixty Third Yearbook, 1964 371-401.

(28) Melson, R. "How Well Are Colleges Preparing Teachers for Modern Mathematics." The Arithmetic Teacher, XII (January, 1965), 51-55.

(29) Melton, A. W. "The Science of Learning and the Technology of Educational Methods" Harvard Educational Review, XXIX (Spring, 1959), 96-106.

(30) Meurhenry, W. C. et al., Trends in Programed Instruction. Department of Audio-Visual Instruction, National Society for Programed Instruction, 1964.

(31) Moore, W. J. and W. I. Smith. Programed Learning: Theory and Research. Princeton, New Jersey: D. Van Nostrand Company, Inc., 1962.

(32) Pressy, S. L. "Teaching Machine (and Learning Theory) Crisis." Journal of Applied Psychology, XXXXVII (February, 1963), 1-6.

(33) Pressy, S. L. "Autoinstruction: Perspectives, Problems, Potential." National Society for the Study of Education; Sixty Third Yearbook, 1964, 354-370.

(34) Reynard, H. E. "Preservice and In-Service Education of Teachers." Review of Educational Research, XXXIII (October, 1963) 369-380.

(35) Ripple, R. E. "Programed Instruction: A New Approach to Teaching and Learning." Journal of Educational Psychology, 1965, 56, 133-139.

(36) Ross, W. Teaching Machines: Industry Survey and Buyers Guide. New York: The Center of Programed Instruction, Inc., 1962.

(37) Sparks, J. N. "Arithmetic Understandings Needed by Elementary School Teachers." The Arithmetic Teacher, VIII, (January, 1961), 395-402.

(38) Steele, R. G. D. and J. H. Torrie. Principles and Procedures of Statistics. New York: McGraw Hill Book Company, 1960.

(39) Stolorou, L. "Implications of Current Research and Future Trends." Journal of Educational Research, 1962 55, 519-527.

(40) Thorndike, E. L. The Fundamentals of Learning. New York: Teachers College, Columbia University, 1932.

(41) Travers, R. M. W. Essentials of Learning. New York: MacMillan Company, 1967.

(42) Walker, H. M. and J. Lev. Statistical Inference. New York: Henry Holt and Company, 1953.

(43) Watson, J. B. "Psychology as the Behaviorist Views It." Psychological Review, 1913, 20, 158-177.

(44) Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw Hill Book Company, 1962.

(45) Woodring, P. "Reform Movements from the Point of View of Psychological Theory." National Society for the Study of Education; Sixty Third Yearbook, 1964, 286-305.

(46) Zaleznik, A. And D. Moment. The Dynamics of Interpersonal Behavior. New York: John Wiley & Sons, 1964

(47) Educational Testing Service. Cooperative Mathematics Tests Handbook. Princeton, New Jersey: Educational Testing Service, 1964.

APPENDIX A

INDIVIDUAL SCORES OF SUBJECTS PARTICIPATING IN THE STUDY

CONTROL			L. T.			P. L. D.			L. P. D.		
ACTM	Prt	Pst	ACTM	Prt	Pst	ACTM	Prt	Pst	ACTM	Prt	Pst
23	23	26	11	12	19	14	13	18	20	19	26
16	27	23	13	10	18	23	26	35	23	22	30
27	23	24	21	22	28	18	20	25	18	15	26
22	20	22	24	28	28	17	16	24	30	28	34
10	13	13	19	21	31	14	13	19	27	34	35
16	17	17	22	18	30	14	22	28	25	25	30
15	13	17	10	12	16	17	15	21	17	15	20
7	19	26	12	12	17	20	28	32	17	18	25
10	13	16	19	17	25	21	32	33	18	12	22
20	19	26	25	29	33	17	15	25	16	12	25
14	10	13	16	12	17	16	21	26	21	20	29
21	19	20	13	10	14	10	14	19	25	22	31
25	19	25	25	32	37	22	22	30	29	26	31
22	26	30	15	25	27	24	22	29	21	30	31
5	9	12	15	15	21	18	12	23	25	32	34
22	24	25	17	16	24	14	14	17	22	26	36
			14	17	25	24	21	32	13	8	13
			21	20	22	21	20	25	19	13	24
			17	26	25	14	15	25	18	14	28
			19	13	21	3	14	21	12	14	20
			15	9	11	20	19	28	24	24	25
			15	9	17	23	21	28	24	18	28
			19	16	24	15	15	18	18	18	27
			14	15	28	25	15	30	19	13	24
			9	9	10	17	13	18	14	18	28
			23	20	31	21	23	29	7	16	14
			14	11	20	15	10	15	14	23	24
			15	15	21	19	18	25	27	21	21
			19	20	27	20	13	28	17	20	26
			16	10	14	21	16	33	34	26	38
			16	21	22	23	32	37	18	24	25
			17	10	19	17	16	23	18	19	31
			14	9	19	27	18	28	3	12	24
			24	17	17				14	11	21
			17	16	26				26	24	30
			12	14	10				11	13	31
			10	9	16				17	17	26
			15	19	20				22	24	32
			8	11	17				17	23	24
			12	12	25				18	15	21
			15	17	31				16	16	25
			18	22	30				18	19	32
			25	12	22				14	14	30
			11	24	28				18	16	22
			17	16	24				28	22	29
									18	17	24
									18	20	32

APPENDIX B

STATISTICAL EQUATIONS USED IN THE ANALYSIS OF COVARIANCE--FOUR GROUPS

STATISTICAL EQUATIONS

1. For adjusted sum of squares of error:	Value
$E_{xx} = \sum_{j} (E_{xxj}) = \sum_{j} \sum_{i} (X_{ij} - \bar{X}_{j})^2$	3866.42
$E_{zz} = \sum_{j} (E_{zzj}) = \sum_{j} \sum_{i} (Z_{ij} - \bar{Z}_{j})^2$	4516.71
$E_{yy} = \sum_{j} (E_{yyj}) = \sum_{j} \sum_{i} (Y_{ij} - \bar{Y}_{j})^2$	4641.39
$E_{xy} = \sum_{j} (E_{xyj}) = \sum_{j} \sum_{i} (X_{ij} - \bar{X}_{j})(Y_{ij} - \bar{Y}_{j})$	2530.64
$E_{zy} = \sum_{j} (E_{zyj}) = \sum_{j} \sum_{i} (Z_{ij} - \bar{Z}_{j})(Y_{ij} - \bar{Y}_{j})$	3523.44
$E_{xz} = \sum_{j} (E_{xzy}) = \sum_{j} \sum_{i} (X_{ij} - \bar{X}_{j})(Z_{ij} - \bar{Z}_{j})$	2564.06
$d = E_{xx}E_{zz} - (E_{xz})^2$	10,899,111.
$b_{xy} = \dfrac{E_{zz}E_{xy} - E_{xz}E_{zy}}{d}$	.220

$$b_{yz} = \frac{E_{xx}E_{zy} - E_{xz}E_{xy}}{d}$$ .655

$$E'_{yy} = E_{yy} - b_{yx}E_{xy} - b_{yz}E_{zy}$$ 1776.07

2. For adjusted sum of squares of treatment:

$$T_{xx} = \sum_j n_j (\overline{X}_j - \overline{X})^2$$ 209.38

$$T_{zz} = \sum_j n_j (\overline{Z}_j - \overline{Z})^2$$ 229.26

$$T_{yy} = \sum_j n_j (\overline{Y}_j - \overline{Y})^2$$ 667.65

$$T_{xy} = \sum_j n_j (\overline{X}_j - \overline{X})(\overline{Y}_j - \overline{Y})$$ 356.36

$$T_{zy} = \sum_j n_j (\overline{Z}_j - \overline{Z})(\overline{Y}_j - \overline{Y})$$ 317.60

$$T_{xz} = \sum_j n_j (\overline{X}_j - \overline{X})(\overline{Z}_j - \overline{Z})$$ 207.76

$$T'_{yy} = S'_{yy} - E'_{yy}$$ 255.62

3. For total sum of squares:

$S_{xx} = T_{xx} + E_{xx}$	4075.80
$S_{zz} = T_{zz} + E_{zz}$	4745.97
$S_{yy} = T_{yy} + E_{yy}$	5309.04
$S_{xy} = T_{xy} + E_{xy}$	2887.08
$S_{zy} = T_{zy} + E_{zy}$	3841.08
$S_{xz} = T_{xz} + E_{xz}$	2771.82
$d'' = S_{xx}S_{zz} - (S_{xz})^2$	11,660,679.
$b''_{yx} = \dfrac{S_{zz}S_{xy} - S_{xz}S_{zy}}{d''}$	.215
$b''_{yz} = \dfrac{S_{xx}S_{zy} - S_{xz}S_{xy}}{d''}$	.656
$S'_{yy} = S_{yy} - b''_{yx}S_{xy} - b''_{yz}S_{zy}$	2031.69

APPENDIX C

STATISTICAL EQUATIONS RELATED TO TUKEY'S PROCEDURE FOR COMPARING INDIVIDUAL MEANS

1. Control Group: Values

$E_{xx} = \sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n_1}$	656.40
$E_{zz} = \sum Z_i^2 - \frac{\left(\sum Z_i\right)^2}{n_1}$	457.70
$E_{yy} = \sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{n_1}$	604.00
$E_{xy} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n_1}$	344.30
$E_{zy} = \sum Z_i Y_i - \frac{\sum Z_i \sum Y_i}{n_1}$	489.70
$E_{xz} = \sum X_i Z_i - \frac{\sum X_i \sum Z_i}{n_1}$	368.00
$E'_{yy} = E_{yy} - b_{xy} E_{xy} - b_{yz} E_{zy}$	207.50
$\overline{Y}_{A1} = \overline{Y}_1 - b''_{yx} (\overline{X}_1 - \overline{X}) - b''_{yz} (\overline{Z}_1 - \overline{Z})$	20.59

2. Lecture Textbook Group:

$$E_{xx} = \sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n_2} \qquad 876.8$$

$$E_{zz} = \sum Z_i^2 - \frac{\left(\sum Z_i\right)^2}{n_2} \qquad 1549.8$$

$$E_{yy} = \sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{n_2} \qquad 1766.0$$

$$E_{xy} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n_2} \qquad 731.4$$

$$E_{zy} = \sum Z_i Y_i - \frac{\sum Z_i \sum Y_i}{n_2} \qquad 1303.9$$

$$E_{xz} = \sum X_i Z_i - \frac{\sum X_i \sum Z_i}{n_2} \qquad 690.0$$

$$E'_{yy} = E_{yy} - b_{yx} E_{xy} - b_{yz} E_{yz} \qquad 751.0$$

$$\overline{Y}_{A2} = \overline{Y}_2 - b''_{yx} (\overline{X}_2 - \overline{X}) - b''_{yz} (\overline{Z}_2 - \overline{Z}) \qquad 24.26$$

3. Program Lecture Discussion Group:

$$E_{xx} = \sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n_3} \qquad 735.0$$

$$E_{zz} = \sum Z_i^2 - \frac{\left(\sum Z_i\right)^2}{n_3} \qquad 967.0$$

$$E_{yy} = \sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{n_3} \qquad 734.4$$

$$E_{xy} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n_3} \qquad 595.4$$

$$E_{zy} = \sum Z_i Y_i - \frac{\sum Z_i \sum Y_i}{n_3} \qquad 784.4$$

$$E_{xz} = \sum X_i Z_i - \frac{\sum X_i \sum Z_i}{n_3} \qquad 414.4$$

$$E'_{yy} = E_{yy} - b_{xy} E_{xy} - b_{zy} E_{zy} \qquad 89.63$$

$$\overline{Y}_{A3} = \overline{Y}_3 - b''_{yx} (\overline{X}_3 - \overline{X}) - b''_{yz} (\overline{Z}_3 - \overline{Z}) \qquad 25.40$$

4. Lecture Program Discussion Group:

$$E_{xx} = \sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n_4}$$ 1598.3

$$E_{zz} = \sum Z_i^2 - \frac{\left(\sum Z_i\right)^2}{n_4}$$ 1542.3

$$E_{yy} = \sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{n_4}$$ 1265.0

$$E_{xy} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n_4}$$ 860.0

$$E_{zy} = \sum Z_i Y_i - \frac{\sum Z_i \sum Y_i}{n_4}$$ 960.0

$$E_{xz} = \sum X_i Z_i - \frac{\sum X_i \sum Z_i}{n_4}$$ 1092.3

$$E'_{yy} = E_{yy} - b_{xy} E_{xy} - b_{zy} E_{zy}$$ 456.17

$$\overline{Y}_{A4} = \overline{Y}_4 - b''_{yx} (\overline{X}_4 - \overline{X}) - b''_{yz} (\overline{Z}_4 - \overline{Z})$$ 26.76

5. Mean Square of the Error (Significant Gap Test)

$$S_{\bar{x}} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} + \frac{S_3^2}{n_3} + \frac{S_4^2}{n_4}} \quad ; \qquad 1.24$$

S_i^2 = mean square of error for the i group

6. Mean Square of the Error (Straggler Test)

$$S_{\bar{x}} = \sqrt{\frac{S_2^2}{n_2} + \frac{S_3^2}{n_3} + \frac{S_4^2}{n_4}} \qquad .82$$

7. Mean Square of the Error (Test for Excessive Variability)

$$S_{\bar{x}} = \sqrt{\frac{S_3^2}{n_3} + \frac{S_4^2}{n_4}} \qquad .55$$

VITA

Philip Edward Gibbons

Candidate for the Degree of

Doctor of Education

Thesis: A COMPARATIVE ANALYSIS OF THE IMPACT OF VARIOUS METHODS OF INSTRUCTION ON ACHIEVEMENT AND UNDERSTANDING IN MATHEMATICS FOR ELEMENTARY TEACHERS

Major Field: Secondary Education

Biographical:

Personal Data: Born in Kingston, Pennsylvania, December 19, 1932, the son of Philip Lloyd and Helen Rosaria Gibbons.

Education: Attended grade school and high school in Forty Fort, Pennsylvania; graduated from Forty Fort High School in 1950; received the Bachelor of Arts degree from King's College, Wilkes Barre, Pennsylvania, in June, 1955; received the Master of Arts degree from Boston College, Chestnut Hill Massachusetts, in June, 1961, with a major in mathematics education; completed requirements for the Doctor of Education degree at Oklahoma State University in July, 1967.

Professional Experience: Employed as a mathematics teacher at Candor Central High School, Candor, New York from September, 1955 to June, 1959; attended the National Science Foundation Institute in Mathematics at Boston College from September, 1959 to August, 1960; employed as a high school mathematics teacher at Maine-Endwell Senior High School, Endwell, New York in September, 1960; became department chairman in June, 1962; became coordinator of mathematics, K-12, in June, 1963; attended the National Science Foundation Institute in Mathematics at Oklahoma State University, Stillwater, Oklahoma from June, 1965 to August, 1966; employed as mathematics instructor at Northern Oklahoma College from September, 1966 to May, 1967; employed as mathematics instructor at Southwestern State College, Weatherford, Oklahoma since June, 1967.

Professional Organizations: National Council of Teachers of Mathematics, National Education Association, Oklahoma Education Association, and Phi Delta Kappa.