LECTURE DEMONSTRATION VS. INDIVIDUAL LABORATORY WORK IN A NATURAL SCIENCE COURSE AT MICHIGAN STATE UNIVERSITY

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY

Robert L. Bradley

1962

This is to certify that the

thesis entitled

LECTURE DEMONSTRATION VS INDIVIDUAL LABORATORY
WORK IN A NATURAL SCIENCE COURSE
AT MICHIGAN STATE UNIVERSITY

presented by

ROBERT LINCOLN BRADLEY

has been accepted towards fulfillment of the requirements for

Ph.D. **degree in** Higher Education

Paul L. Dressel
Major professor

Date November 14, 1962

O-169

ABSTRACT

LECTURE-DEMONSTRATION vs. INDIVIDUAL LABORATORY
WORK IN A NATURAL SCIENCE COURSE
AT MICHIGAN STATE UNIVERSITY

by Robert L. Bradley

The major aim of the study was to determine the role of the individual laboratory and the lecture-demonstration methods of teaching natural science in achieving the objectives of general education as measured by a paper and pencil test. Specifically the problem resolved itself to comparing the relative merits of two instructional procedures in natural science. The treatment groups were the lecture-demonstration group and the individual laboratory group.

The aims and objectives of general education for the natural science course are:

I. To gain an understanding of science.

II. To gain an understanding of the products of science, mainly the subject matter of science.

III. To gain an understanding of the methods of science and the ability to carry out operations involved in these methods.

IV. To gain an understanding of the relationship of science to other areas such as social science and religion.

The experiment was carried out during the Spring quarters of

1960 and 1961. All subjects of the experimental study were students enrolled in Natural Science 183, a general education course in physical science at Michigan State University. Four sections of students each quarter were involved and were randomly assigned at the time of registration.

A factorial design involving five variables varying in two ways was used in this investigation. The experiment was carefully controlled with respect to the instructional time, laboratory apparatus and equipment, and the evaluation instruments.

Information on initial status of the students was obtained by the College Qualification Tests; sex, and previous college laboratory science experience was tabulated. The achievement criterion was the term-end final examination.

In the experimental design, the following six null hypotheses were tested:

1. There is no difference between the subject matter achievement of college students who undergo instruction in physical science for general education by either the lecture-demonstration method or the individual laboratory method,
2. there is no difference between sexes,
3. there is no difference between instructors,
4. there is no difference between the above median group and the below median group as measured by the CQT-T scores,
5. there is no difference between groups who had previous college laboratory science courses and

those who did not,

6. there is no interaction between method, sex, above median and below median groups, and those who have had previous laboratory science vs. those who have not.

The experimental data were subjected to analysis of variance by means of a 2^n factorial design.

The Findings

Five of the null hypotheses were accepted and the sixth--there is no difference between the above median group and the below median group as measured by the *College Qualification Tests*--was rejected. On examining in more detail the two groups for each instructor the sum of squares for the above median group for one instructor was found to be much larger than the sum of squares for the below median group.

Conclusions

The experimental evidence showed that neither method was superior for teaching natural science when the achievement of the aims and objectives of general education science are measured by a paper and pencil test.

The evidence also indicated that there was no significant interactions between teaching method, sex, instructor, and previous laboratory experience. There was evidence that one instructor was more effective with high ability students than below ability students.

On the basis of the experimental evidence, the experimenter recommends that the teaching method or methods be considered in view of each college's own resources and the method used be the one that is the most efficient for their situation.

LECTURE DEMONSTRATION vs. INDIVIDUAL LABORATORY WORK IN A NATURAL SCIENCE COURSE AT MICHIGAN STATE UNIVERSITY

By

Robert L. incoln Bradley

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

College of Education

1962

25591
6/12/63

ACKNOWLEDGMENTS

The writer wishes to acknowledge the very kind assistance and counsel provided by Dr. Paul Dressel and Dr. Chester Lawson.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION AND BACKGROUND OF THE PROBLEM

The number of high school graduates of Michigan high schools increased from 15,946 to 23,510 between 1955 and 1960 while the percentage of high school graduates going on to college increased from 32% in 1955 to 35% in 1960. While the number of students in college has been increasing the supply of available instructors with Ph.D.'s has been shrinking. The country as a whole produced 8,903 Ph.D.'s in the 1955-56 period and 8,380 in the 1957-58 period, and of each year's total only half have gone into college teaching. In certain fields, science and engineering for example, the proportion has been even less. Actual space per student is declining as a result of two main factors: first, the increasing college enrollment; and second, the failure of the college physical plants to keep pace with enrollment. The last factor results from the inavailability of funds to support needed expansion.

Despite all these pressures on higher education no one desires to lower the quality of instruction. Nevertheless, past practice in regard to class size, number of class meetings, and laboratory requirements should be subject to critical scrutiny. Wasteful and unnecessary requirements should be discarded or reviewed. Science laboratory requirements are of special concern because of the expensive nature of the space, equipment, and time involved. The role of laboratory in General Education Science for the non-science major has often been discussed and there have been some doubts expressed to its necessity and value. It is with this latter problem, specifically in reference to the natural science courses at Michigan State University, that this dissertation is concerned.

A partial answer to the question, "Is the laboratory really necessary in science for general education?", might be found in an historical examination of the role of the laboratory in the teaching of science.

Instruction in the natural sciences has been offered in American colleges and universities since the founding of these institutions. Astronomy and the "Nature of Plants" were offered at Harvard as early as 1690. From the latter part of the eighteenth century until the middle of the nineteenth, instruction in natural science became a very definite portion of the college curriculum. The various sciences were differentiated from each other and separate courses were offered in each. The first professor of chemistry in this country was Dr. James Smith (appointed in 1768) at the College of the Province of New York. Separate chairs of chemistry were established at Princeton in 1795, at Columbia in 1800, at Yale in 1802, and Williams in 1812. Geology was included in the early professorship of chemistry at Yale. Mineralogy, geology, and botany were introduced at Princeton in 1830.

Although science subjects were coming into prominence during this period, the instruction was almost entirely by lecture and demonstration on the part of the instructor. Very few laboratories were in existence and these were entirely for the use of the instructor.

One of the earliest attempts to include laboratory work by the students of natural science was in connection with the founding of Rensselaer Polytechnic Institute in 1824, although it was not till the latter part of the nineteenth century and the early part of the twentieth that individual laboratory work on the part of the student in science courses became at all common.

There were a number of factors that influenced the development of laboratories. They were: first, the influence of such men as Charles Eliot, Lewis Agassiz, Asa Gray, and others; second, the founding of technical schools among which were Rensselaer Polytechnical Institute (1824) and Massachusetts Institute of Technology (1862); third, the Morrill Act in 1862; fourth, the introduction of elective system; fifth, the increasing importance of science in industry and everyday life. As a result, we see at the beginning of the current century science courses reaching the height of their popularity.

More recently some reaction has set in. Laboratory facilities were becoming more and more inadequate, as regard to the space and equipment available per student and lack of skilled laboratory instructors. A part of this reaction the whole field of science instruction came into question. Aims, curricula, and methods were no longer unquestioned, and a period of critical study and evaluation of these elements of the science program began especially in the beginning course of science where enrollment was the largest.

The beginning college course of a science often plays two different roles: first, as an introductory course for students who are majors in the particular science area, and second, as a course for general education and cultural purposes. These two different roles are sometimes seen as incompatible. Hence in order to serve adequately the two very different groups of students there have been developed general education science courses for non-majors parallel to introductory courses for majors in the area.

The aims of the general education survey type course for non-majors generally include: (1) information, (2) development of interest in science, (3) understanding of relationships of science to environment and everyday life, (4) understanding of the relationships of the sciences, and (5) culture. The question then arises as to whether laboratory is essential to the achievement of these aims.

In regards to the role of laboratory in fulfilling the aims of the survey course, Hurd[1] lists 43 functions of laboratory work in science as proffered by 35 science instructors at the University of Minnesota. Among others there are listed the following: to develop manipulative skill (7 instructors); to aid memory (4); to give the scientific manner of thought and training in drawing conclusions (3); to give opportunity for developing sense of perception and acquisition of concepts (3); to develop powers of observation (3).

Of 47 secondary school science teachers polled as to whether laboratory was necessary to develop these abilities: 33 teachers thought that these abilities could not have been developed by any other procedure than the laboratory, 12 thought they could be developed without laboratory, and 2 had no opinion. It should be noted that these views were expressed some years before the integrated general education course with primary emphasis on understanding science came into wide usage.

There seems to be wide acceptance on the part of science

---

[1]Archer W. Hurd, Problems of Science Teaching at the College Level (University of Minnesota Press, 1929), pp. 9-10.

instructors that one of the principal functions of the laboratory is to develop manipulative skill. This manipulative skill as such is not an objective of general education science courses. Therefore, if the other aims and objectives can be mastered by the students in some other manner, the laboratory can be dispensed with in general education science.

Before recommending that the laboratory be dispensed with entirely, alternative approachs such as the lecture-demonstration method should be compared with the individual laboratory method to see if there is any evidence that the aims and objectives of general education science can be met by the lecture-demonstration method as effectively as by the individual laboratory method. The first step in this consideration is accomplished by a survey of related research in this area which compares the efficiency of the two methods in meeting the aims and objectives of science.

Cunningham[2] summarized the research done in the field of lecture-demonstration versus the individual laboratory method covering the period from 1930 to 1946 after examining 18 master's theses, 6 doctor's theses and 10 articles published in learned journals. He found one study which distinctly favored the individual laboratory method for developing laboratory resourcefulness. Studies that were concerned with manipulation of laboratory materials found in four cases that the individual laboratory method was preferable. This was evidence that supported the results of Hurd's[3] questionnaire addressed to secondary

---

[2]Harry A. Cunningham, "Lecture-Demonstration versus Individual Laboratory-A Summary," Science Education, Vol. 30, No. 2, pp. 70-82.

[3]Hurd, loc. cit., pp. 9-10.

school science teachers involving laboratory manipulation. However, none of these studies assumed that this ability in laboratory manipulation was transferable to other kinds of manipulation, thus leaving open the worth of these experiences for the non-science student.

Seventeen studies gave attention to one or more of the elements of scientific thinking but no one investigation made even a slight beginning in the study of this problem in all of its many aspects. The elements of the thinking process that were studied in some of the investigations were as follows:

(1) making proper conclusions to an experiment;

(2) application and interpretation of the results of an experiment;

(3) application of the principles learned;

(4) ability to think in terms of the science subject;

(5) ability to follow the steps in scientific procedure;

(6) per cent of thought questions answered correctly;

(7) method of attack on new problems;

(8) scientific attitudes;

(9) ability to observe;

(10) learning a scientific principle;

(11) greater carry over ability;

(12) ability to distinguish between fact and superstition;

(13) ability to generalize.

Of the 17 studies that gave attention to phases of this problem, 12 favored the demonstration method; four the individual laboratory method; and one came to the conclusion that students could learn to think equally well by either method. These thirty-four studies appear to justify the view that individual laboratory method does perhaps develop laboratory resourcefulness and manipulation of laboratory materials better than the lecture-demonstration method.

It is difficult to define and measure some types of objectives and therefore, results are not definitive. Knowledge and thinking abilities are not the sole concerns, but we do not do very well at measuring other objectives such as beliefs and attitudes. It is pertinent to examine the efficiency of the two teaching methods in regard to the utilization of space, time, and staff.

Cunningham[4] found that the 15 of the 34 reports which gave attention to the time required by each of the two methods to cover the same material favored the demonstration method. Two studies which examined the relative cost of the two methods in regard to apparatus, staff, time, and space found the demonstration method less costly than the individual laboratory method. To illustrate this, consider the following examples of the comparison of the cost of teaching a course in science by the two methods.

The science course considered is a four credit course. It meets four days a week for one hour each day. During the first two days the students, either individually or in pairs, conduct the week's experiment and record and interpret the data. The last two days are devoted to discussion, testing and other activities. One instructor has four sections of 40 students each making a total of 16 contact hours and 160 students. If this same course is taught by the demonstration method where the instructor demonstrates the experiment and

---

[4] Cunningham, loc. cit., pp. 70-82.

discusses the implications as it is being run, 60 students could be handled in the same room making a total of 240 students with the same number of contact hours for the instructor. Under this setup 50% more students could be handled in the same length of time, in the same space and with the same number of staff. This would be a tremendous saving in instructor salaries, space and in cost of equipment and supplies since less equipment and fewer supplies would be needed to demonstrate the experiment than to run it by the individual laboratory method.

The cost of operation could very well be the deciding factor in the method chosen to be used in general education science since manipulative skill is not, as pointed out previously, an objective and when the outcome of the two methods are compared on the basis of the same form of a standard final examination Cunningham[5] found that: 28 studies favored the lecture-demonstration method, 6 favored the individual laboratory method, and 2 found no difference.

The preceding discussion suggests that a closer look at general education science courses should be taken to see if the lecture-demonstration method is as efficient as the individual laboratory method. If so then this is one possible approach to handling the increasing number of students. The purpose of this dissertation is to consider this problem specifically in regards to Natural Science 183, the physical science portion of the general education science at Michigan State University.

The objectives of the Natural Science Department at Michigan State University, and the operational definitions that are used in

---

[5] Cunningham, loc. cit., pp. 70-82.

the study to attempt to answer the previous question as to the efficiency of the lecture-demonstration method versus the individual laboratory method of teaching general education science are as follows:

Objectives

I. To gain an understanding of science.

II. To gain an understanding of the products of science, mainly the subject matter of science.

III. To gain an understanding of the methods of science and the ability to carry out operations involved in these methods.

IV. To gain understanding of the relationships of science to other areas such as social science, religion, philosophy, and so forth.

Definitions

I. Individual laboratory method for the purpose of this study is defined as an individual student or team of two students performing the experiment according to the prescribed method, recording observations, and drawing conclusions with little or no help from the instructor.

II. Lecture-demonstration method for the purpose of this study is defined as: the instructor performs the experiment, the same directions being followed as were given in the individual laboratory method. The order of procedure was (1) statement of the problem; (2) performance according to method prescribed; (3) making the observations; and (4) formation of conclusions.

III. The experimental group was a section that had their work demonstrated to them by the instructor.

IV. The control group were the sections that did their own laboratory work.

V. CQT-T scores is the total score made by the student on the College Qualification Test.

The following chapter reviews the related research on similar courses or courses having analogous objectives and attempts to summarize the findings. This review also makes it possible to define the present research project more explicitly in relation to studies already made.

CHAPTER II

REVIEW OF RELATED RESEARCH

In Chapter I some of the conclusions regarding the effectiveness of lecture-demonstration method vs. individual laboratory method in achieving the aims and objectives of science were examined. It was further pointed out that manipulative skill was not an objective of general education science. Since this dissertation is concerned with the examination of the lecture-demonstration method vs. individual laboratory method in general education science, specifically in the third term of Natural Science at Michigan State University, a more detailed review of the research conducted in this area is in order.

Before attempting to determine relative effects on achievement of different procedures there has to be some basis for evaluating the reports of research in this area. Monroe and Engelhart[6] stated seven requirements or criteria for determining the relative effectiveness of research studies in the area of instructional procedures. Cunningham[7] expanded this list to twelve requirements or criteria.

Ten criteria in the form of questions were synthesized from these statements. The ten criteria provide a basis for appraising

---

[6] W. S. Monroe and Max D. Engelhart, "Experimental Research in Education," University of Illinois Bulletin, Vol. 27, No. 32, Bureau of Educational Research Bulletin No. 48 (Urbana: University of Illinois Press, 1930) pp. 18-58.

[7] Cunningham, *loc. cit.*, pp. 70-82.

research on instructional procedures as well as for planning research. The desirable answer to some of the questions is clearly implied although the actual answer in a given case must be relative and qualified rather than absolute.

(1) Were experimenters and agencies, to which the research work on this problem was submitted, dependable?

(2) Have the problems of these studies been definitely and precisely stated at the beginning of each undertaking?

(3) Is there clear specification of experimental factors?

(4) Is there control of instructor factors?

(5) Is there control of student factors?

(6) Is there control of general school factors?

(7) What is the duration of the experiment?

(8) What kind of data were obtained and how were they measured?

(9) How is the experimental data interpreted?

(10) What results did the experimenter report?

On examining the literature, no studies were found that concerned themselves with the lecture-demonstration vs. individual laboratory method specifically in the area of general education science. However, three studies that concerned themselves with lecture-demonstration method vs. some phase of laboratory work in general education science were examined. Additional studies in the area of lecture-demonstration method vs. individual laboratory method in science in general were also examined in order that the effectiveness of the lecture-demonstration method vs. individual laboratory method could be compared in science as a whole.

These studies, divided into two groups, will be discussed and analyzed by reference to the ten criteria. The first group of studies includes those which deal with the question of lecture-demonstration method vs. laboratory work as it is concerned with general education in science. The second group includes three studies involving lecture-demonstration method vs. individual laboratory method as found in science courses generally. Finally there is a summary of conclusions drawn from the studies that dealt with general education science and the other studies that dealt with science in general.

The three studies which were related to general education are those of Barnard,[8] Kruglak,[9] and Ward.[10] After analyzing the three studies with reference to the ten requirements or criteria, the following analyses were prepared.

Barnard's study was concerned with six classes of the biological portion of an orientation science course in the School of Education, New York University. Three sections consisting of 145 students were taught by a problem solving method that consisted of a laboratory-discussion setup. The other three sections consisting of 137 students were taught by the lecture-demonstration method. This method consisted of formal lectures on the subject matter of the course, supplemented by

---

[8] J. Darrel Barnard, "The Lecture-Demonstration vs the Problem Solving Method of Teaching a College Science Course," Science Education, Vol. 26, No. 3-4 (Oct.-Nov., 1942), pp. 121-132.

[9] Haym Kruglak, "A Comparison of the Conventional and Demonstration Methods in the Elementary College Physics Laboratory," Journal of Experimental Education, Vol. 20, pp. 293-300.

[10] John N. Ward, "Group-Study vs Lecture-Demonstration Method in Physical Science Instruction for General Education College Students," Journal of Experimental Education, Vol. 24, pp. 197-210.

instructor demonstrations which could be used to illustrate important concepts covered in the lectures. There were no laboratory exercises as in the problem-solving method. The study was conducted under the auspices of the School of Education at New York University. Barnard stated his problem as follows:[11]

> This study compared the relative effectiveness of a lecture-demonstration method and a problem-solving method of teaching the biological portion of an orientation science course at the college level with respect to achievement on tests constructed to measure (1) recall of specific information, (2) understanding of generalizations, (3) abilities in problem solving and (4) scientific attitudes.

The specifications of the experimental factors were very carefully laid out since it was pointed out that the study was not concerned with a comparison of the specific teaching procedures used in any one of the two methods nor with an evaluation of any plan of organizing the subject matter but only with the outcomes of achievement on tests to measure (a) recall of specific information, (b) understanding of generalizations, (c) abilities in problem solving and (d) scientific attitudes. The retention of final achievement in the four outcomes was not determined.

The two groups were equated by pairing the students in the two methods upon two bases, scores on the American Council on Education Psychological Examination for College Freshmen and on pretests constructed by the experimenter administered at the beginning of the study. The scores on these two measures were equated with one another to render the two scores comparable.

The school factors were rather carefully controlled. Instructional materials and time devoted to learning activity were nearly the

---

[11] Barnard, loc. cit., p. 131.

same for both groups. Since the experiment was carried on in one department in a particular college for one semester, such factors as size of college, college administration and so forth did not need to be considered as a variable but this of course greatly limits the generality of the results.

The two groups had such common elements as: same tests, same bibliography, same instructor and assistants for all six groups, the requirement of a written report of each unit, and the same general problem areas for the course. The statistical analysis appeared adequate.

Barnard reached the following conclusions about the effectiveness of the two methods:

1. The lecture-demonstration method has some advantages over the problem-solving method with respect to achievement on tests covering specific information, although the results in all cases are not significant.

2. Neither method has statistically significant advantages over the other with respect to achievement on tests covering the understanding of generalizations.

3. The problem-solving method has statistically significant advantages over the lecture-demonstration method with respect to achievement on test covering certain abilities in problem solving.

4. The problem-solving method has statistically significant advantages over the lecture-demonstration method with respect to achievement on test dealing with scientific ideas.

In conclusion, the chief criticisms of the experiment would be these: (1) the groups were not large enough to be comparable when broken down into three ability levels to justify the determination of the significance of obtained differences. (2) One instructor only was used for both methods. (3) The tests used were not standardized by any outside criterion although they had acceptable intercorrelation.

Kruglak[12] conducted a study of the lecture-demonstration method vs. individual laboratory method in an area where laboratory work is commonly thought of as indispensible. This experiment was conducted at the University of Minnesota in the Department of Physics and dealt with 194 students enrolled in a non-technical mechanics course Fall quarter 1950. This course could be used to satisfy the science requirements for non-science majors and it is on this basis classified as a general education science course. The two methods studied were the lecture-demonstration method and the individual laboratory method. In the lecture-demonstration method group the instructors did the experiment under study. The instructor assembled the apparatus, made adjustments, etc. The individual laboratory method on the other hand gave this responsibility to the students.

Kruglak[13] stated his problem thus:

> The objective of this experiment was to compare the learning outcomes of two instructional methods in the elementary college physics laboratory.

Equivalence of groups was secured by using 108 male Minnesota high school graduates who had not previously taken physics or its equivalent and then by means of random numbers were assigned to the 4 sections of the experimental group and the 4 sections of the control group. As a result of withdrawals and incomplete data the original number was reduced to 87. This caused an unequal number in the groups. Consequently, a number of cases were rejected by random number technique until equal groups resulted. The final number of cases studied was 64.

---

[12] Kruglak, loc. cit., pp. 293-300.

[13] Ibid., p. 293.

There was no need to consider general school factors as a variable since the study was conducted in one specific department at a single university. The instructor factor was randomized by tossing a coin to see which graduate assistant would teach which pair of sections, one from each of the two method groups, experimental and control. Due to the size of the school the three lectures per week were the same for all 8 groups and all the laboratories met on the same day. The textbook, assignments and the tests as well as the duration of the experiment--one quarter, the first of the physics sequence--were the same.

The achievement in the course was measured by tests that had been carefully reviewed and approved by prominent physics professors.

The interpretation of the experimental data by Kruglak[14] was:

> I. On the basis of the results of this study it was valid to conclude that the conventional method in the physics laboratories was more effective than the demonstration method for the teaching of instrumental situations, simple measuring techniques and problems involving apparatus. It was also concluded that neither method was superior for the more complex laboratory problems. The experimental evidence appeared to justify the conclusion that neither method influenced measurably the scores on pencil-paper tests based on the material of the lectures and laboratory work.
>
> II. No statistically significant differences in the means of the four criteria could be attributed to the differences between the individual instructors, nor did any one method appear to give better results for one instructor than for any other instructor.

This experiment is highly commendable in several respects. The groups may be considered to have been of satisfactory equivalence, the school factors were the same, the instructor factors were satisfactorily randomized and the tests used were reasonably reliable and valid. The groups were several in number, which is a definite aid in forming

---

[14] Ibid., p. 299.

conclusions. Yet with all the definitely controlled conditions there were some points of weakness in this study: (1) the size of the groups were small; (2) they were all male; (3) all the instructors were relative inexperienced graduate students.

Ward's[15] experiment was conducted at the Pennsylvania College for Women (now Chatham College) as a Ph.D. requirement under Palmer O. Johnson at the University of Minnesota. This study is the only one of the three major studies reviewed that actually was conducted with a group enrolled in a course planned especially for general education purposes.

The study examined the effects of two methods of presenting the material in a course in Science for General Education students. The two methods employed were the lecture-demonstration method and the group method. In the lecture-demonstration method the instructor always treated in class the topics to be covered, performing the experimental exercises and manipulating all the audio-visual aids. The students themselves introduced the topics, did the experimental exercises and used the audio-visual aids in the group method. Obviously this procedure was somewhat different from the usual individual laboratory effort.

Ward stated his problem at the beginning of his paper thus:

> The purpose of this investigation was to determine whether subject matter in physical science would be learned as well under a group method of instruction as under a lecture-demonstration method in general education college course. Specifically the problem resolved into comparing the relative effectiveness of the two objectives of general education: (1) recall and recognition of facts, principles, and symbols, and (2) more understanding of implications of facts and

---

[15]Ward, <u>loc. cit.</u>, pp. 197-210.

principles, of pertinent reading material, and problem situations.

Randomization in this experiment was based upon the fact that the samples of the students were not influenced by the experimenter in any way regarding their selection. They were assigned on the basis of alphabetical registration into sections that fit their schedule. The claim of randomization was verified by analysis made of the means and variances of the ACE-Q scores of the five groups studied plus those of the preceeding two years classes. The random processes used in subdividing the samples were also verified by analysis of the means and variances.

The students involved were all women, all non-science majors, and all sophomores or above. There was no need to be concerned about school factors or instructor factors as variables since this experiment was conducted by one instructor in a specific department of one college. This of course limits the generality of the study and is one of its principal weaknesses.

The experiment which lasted two semesters employed two criterion measures, Test I (reliability, .804), given in the middle of the first semester and six months later as Retest I (reliability .77), and Test II (reliability, .57) given at the end of the first semester. Analyses of variance and covariance were used to test the null hypotheses for the scores achieved on each of the measures, Test I, Retest I, and Test II, by the total samples.

The results of the study were: (1) In more understanding type of learning situation, the group-method was favored for the high ability group while the lecture-demonstration method was favored for the low ability group. (2) When the learning situation was in terms of the longer retained recall-recognition type of learning, both methods were of equal value for the high ability group while the lecture-demonstration method was favored for the low ability group.

In conclusion this experiment is highly commendable in several respects. The groups may be considered to have been satisfactorily equivalent, the college and instructor factors were satisfactorily controlled and the tests used were reasonably reliable.

The chief points of weakness of this experiment would be these: (1) the size of the groups were small and consisted of all women, (2) one instructor was used to teach both methods, (3) no test of known high validity was used.

The three studies dealing with lecture-demonstration method vs. individual laboratory method in non-general education science courses are those of Knox,[16] Johnson,[17] and White.[18]

Two of the studies, those of Knox[19] and White,[20] stated their problem very specifically at the beginning of their report while Johnson[21] discussed the history and background of the lecture-demonstration

---

[16]W. W. Knox, "The Demonstration Method versus the Laboratory Method of Teaching High School Chemistry," The School Review, Vol. 25, (May, 1927), pp. 376-386.

[17]Palmer O. Johnson, "Comparison of the Lecture-Demonstration, Group Laboratory Experimentation, and Individual Laboratory Experimentation Methods of Teaching High School Biology," Journal of Educational Research, Vol. 18 (Sept. 1928) No. 2, pp. 103-111.

[18]John R. White, "A Comparison of the Laboratory and the Lecture-Demonstration Methods in Engineering Instruction," An unpublished Doctor's Dissertation, School of Education, New York University, 1943.

[19]Knox, loc. cit., p. 376.

[20]White, loc. cit., p. 1.

[21]Johnson, loc. cit., p. 105.

method vs. the individual laboratory method first. He stated his problem as an examination of three methods: lecture-demonstration method, group laboratory method, and individual laboratory method.

All three of the studies were very specific about the experimental factors and all had adequate control of student, instructor, and general school factors. The study of Knox had two instructors, White's study involved but one instructor, while that of Johnson involved six instructors.

The studies all used acceptable modern statistical methods based on data that were extensive and accurate. The duration of the studies varied from one semester for White's to one year for those of both Knox and Johnson.

The experimental data were very adequately interpreted. Knox found the laboratory method to be superior for the inferior students and the lecture-demonstration method superior for the better students. Johnson found that although the lecture-demonstration method out-ranked both of the laboratory methods in five of six sets of experiments none were statistically significant. White found no significant difference between the two methods.

The chief points of weakness of these studies seem to be: (1) the sizes of the groups in all three cases were relatively small, the largest only involved 48 students. One study, involved only males. (2) the validity of the tests used, while all seemed to be reliable, seem questionable in light of present day validation procedures. (3) The results are not conclusive because the differences in achievement were not consistently favorable to any one method.

A Summary of Conclusions

The experimental evaluations of the lecture-demonstration method versus some form of the laboratory method of teaching science exhibit great variation in the conclusions reported by the various investigators. It also very quickly becomes evident that all of the possible desirable outcomes of any one method were not all tested by the various investigators.

Although most of the data seems valid, the diversity of findings appear to cast some doubt on the validity of the tests, the adequacy of controls of such factors as instructor conditions, and the use of small unrepresentative groups and no retrial of experiments. There also seems to be no standard lecture-demonstration or laboratory method. The conclusions reached by the various studies offer support for both methods and are scarcely definitive.

I. Conclusions contending that the laboratory method is superiod:

(1) There is a statistically significant advantage for the laboratory method in areas dealing with scientific attitudes and certain abilities in problem solving for general education in biology. (Barnard)

(2) The order of preference for methods involving instrumental situations in physics for non-majors places the laboratory method first. (Kruglak)

(3) For permanent long range learning in situations involving more understanding type of learning for high ability students the lecture-demonstration method seems least effective. (Ward)

(4) The individual laboratory method seems to provide greater opportunity for the exercise of individual differences. (Johnson)

(5) The laboratory method appears to be superior for providing knowledge and method of attack. (Knox)

(6) For laboratory manipulation and skills and understanding of apparatus, the laboratory method appears superior. (White)

II. Conclusions claiming that the Lecture-Demonstration Method is superior:

(1) The lecture-demonstration method has some advantages over the laboratory problem solving method with respect to achievement on tests covering specific information although the results in all cases are not statistically significant. (Barnard and Kruglak)

(2) Lecture-demonstration method was statistically significant at the 5% level of confidence for lower ability sub-group for expression of individual differences in longer retained more-understanding type of learning and also for the longer retained recall-recognition type of learning. (Ward)

(3) For the average or superior student in providing knowledge for both immediate and permanent retention and for purpose of providing techniques for handling new problems, the lecture-demonstration method is much to be preferred. (Knox)

(4) With respect to long range retention of material the lecture-demonstration method seems to be better for the above average students. (Johnson and White)

III. Conclusions contending that students achieve equally well by either method:

(1) Immediate retention is about equal in both lecture-demonstration method and the laboratory method. (Kruglak, White and Ward)

(2) Neither method has statistically significant advantages over the other with respect to achievement on tests covering the understanding of generalizations. (Barnard, Kruglak, Ward and Knox)

(3) No statistically significant differences in means of the four criteria could be attributed to the differences among the four instructors. (Kruglak)

From these studies, one can only conclude that:

(1) No one method can be considered superior in all cases. The objectives of science teaching, the ability level of the students, and the facilities available should largely determine the method used.

(2) Where costs per student is a major concern, the lecture-demonstration method seems to offer the best advantages.

(3) The problems of the lecture-demonstration method versus some kind of laboratory method still seems unsolved and as complex as ever. It appears that there should be more careful experimentation involving careful control of non-experimental factors. More reliable testing is needed before any definitive answer can be given. When experimentation has indicated that a particular method is superior in outcomes, the method must still be examined in terms of the values of these outcomes relative to the costs involved.

CHAPTER III

DESIGN OF THE STUDY

I. Introduction

This study examines an alternative method of teaching the physical science portion of the science program for general education at Michigan State University. The science program consists of three quarters. The first deals with reproduction and the cell theory and simple genetics. The second quarter examines some basic geological principles and the evidence that supports organic evolution. The third quarter (physical science) is concerned with the nature of matter and energy as explained by the concept of heat, atomic-molecular theory, and is approached on the basis of methods involved in their development.

The two methods which were examined are the lecture-demonstration method and the individual laboratory method. The lecture-demonstration method is the experimental part of the study. It consisted of the instructor performing the experiments for the quarter in a rather rigid manner. In the individual laboratory method the students did their own experiments. In both groups the same reading assignments were made from the same textbook, the lectures were common to both groups, and the tests were the same since they were given in the common lecture.

The lecture-demonstration method and the individual laboratory method were based upon certain assumptions, among which the following were pre-eminent: (1) course objectives were the same for all students, and were the responsibility of the instructor, (2) course subject matter should be selected by the instructor, (3) classroom activities should be determined by the instructor in order to motivate and stimulate learning, and (4) evaluation of each individual student's achievement in the course was the responsibility of the instructor, and should be made on the basis of scores obtained on valid and reliable measuring instruments. The only difference in the two groups was the manner in which the data were gathered and interpreted for the laboratory exercises.

Scientific experimentation is concerned with the empirical testing of hypotheses. In order to place the burden of showing any significant difference between methods of instruction directly upon the evidence obtained from them, the following null hypotheses were adopted:

(1) There is no difference between the subject matter achievements of college students who undergo instruction in physical science for general education by either the lecture-demonstration method or the individual laboratory method,

(2) there is no difference between sexes,

(3) there is no difference between instructors,

(4) there is no difference between the above median group and the below median group as measured by the College Qualification Test, total scores (CQT-T).

(5) there is no difference between groups who had previous college laboratory science courses and

those who did not,

(6) there is no interaction between method, sex, above median and below median groups, and those who have had previous laboratory science versus those who did not.

The first hypothesis was tested by univariate analysis of variance. The other five hypotheses were tested by factorial analysis. Due to the natural limitations on the sizes of the samples of students investigated, a five per cent level of significance was adopted; thus to reject the null hypothesis that any observed difference in samples' means would be due to chance factors rather than to other causes, the observed difference would have to be large enough to be attributable to chance factors in five per cent or less trials only.

The cirterion measure employed was the final examination for the course in Natural Science. The assumption is that final examinations adequately measure the objectives of Natural Science stated previously in this study and that the final examination letter grades based on a fifteen point scale are comparable from year to year for the same quarter.

II. The Population

The students included in the study were selected from those enrolled in Natural Science 183 at the time. The population consisted of two groups inasmuch as the students studied were selected at two different times. One experimental group of 80 taught by instructor "B_1" was studied during the Spring term of 1960 and the other experimental group of 82 taught by instructor "B_2" was studied during the

Spring of 1961. Three regular sections for each instructor served as the control. With rare exception, Natural Science 183 was the first physical science course taken by the students in the university. Natural Science 183 is the third and final term of a three-term sequence required of all freshman and all transfer students who do not have previous credit for the science. As a consequence, practically all students participating were freshmen in their third quarter of college work.

III. Sampling Procedure

The admission of students to both the experimental and control sections of the two instructors were not influenced by either the instructors or the experimenter. This coupled with alphabetical admission to registration, and the distribution of available space evenly throughout the period of registration was considered to be an adequate randomization procedure. Subsequent tests verified this to be true.

IV. Description of the Treatments

There were two experimental groups studied in two different quarters, one large group in the Spring of 1960 taught by instructor "B_1" and the second large group in the Spring of 1961, taught by instructor "B_2". There were two basically different treatment patterns in methods of teaching and size of class. These were as follows: the standard method; this is the normal class of Natural Science 183 and acted as the control. The class size was 35 to 38 per section. The teaching procedure was one two-hour laboratory, one one-hour discussion, and two one-hour lectures. The students did their own laboratory work and recorded and interpreted their own data. The lecture-demonstration method was the large group. This was the experimental group. The populations in the large group were 80 and 82 at the beginning of the quarter and 77 in each at the end of the quarter due to drop-outs. The teaching procedures were quite different from the regular method. In this group all of the laboratory exercises were demonstrated by the instructor with or without the aid of a member of the class.

V. Assignment of the Treatments

Each instructor had four sections. Three of these sections had their laboratory work in the usual fashion, namely, performing their own experiments, gathering and interpreting their own data for each exercise. In the fourth section laboratory work was demonstrated to them by the instructor. The time of meeting of this latter group, designated as the experimental group, was picked before registration on the basis of availability of a room to hold a large group. The total contact hours of both the control and the experimental groups were held constant, the only difference being in the second hour of the two-hour laboratory period of the experimental group where attendance was not compulsory. The experimental group had the option of going to the laboratory during the second hour to examine or rerun on their own, the experiment that had been previously demonstrated to them. Approximately one-fourth of the students availed themselves of this opportunity.

VI. Administration of Treatments

The course in Natural Science 183 meets twice a week for one hour in a lecture session including all the sections assigned to one instructor. The sections meet separately in one two-hour laboratory session, and in one one-hour discussion session which meets at other times through the week. The only difference between the control group and the experimental group was in the two-hour laboratory period where the experimental group was handled as mentioned under the assignment of treatments.

In the lecture-demonstration group, that is, the experimental group, the equipment was oversized for most of the exercises in order that the students have a greater opportunity to see the apparatus used in the experiment. All the tests were administered during the common lecture period. This meant that both the control and the experimental group had the same test at the same time.

There was no effort on the part of the instructors to follow the same pattern of testing. One gave 10 minute weekly quizzes in the last lecture period of the week and the other gave five one-hour tests in the lecture and no quizzes. The instructor who gave weekly quizzes, gave only one, one-hour lecture test.

VII. Data Available and Methods of Analysis

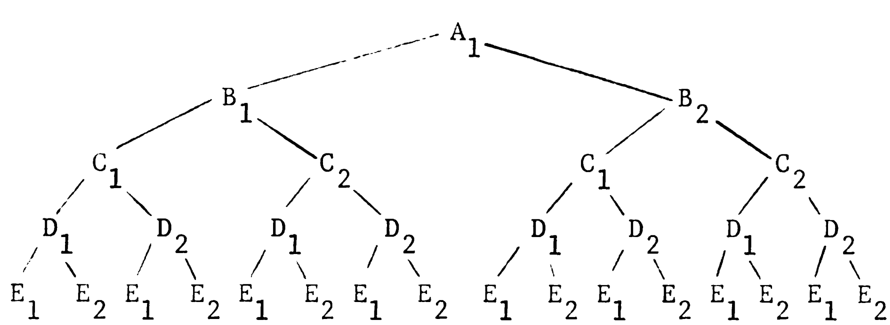
A. Data Available.--The following data were tabulated for each student:

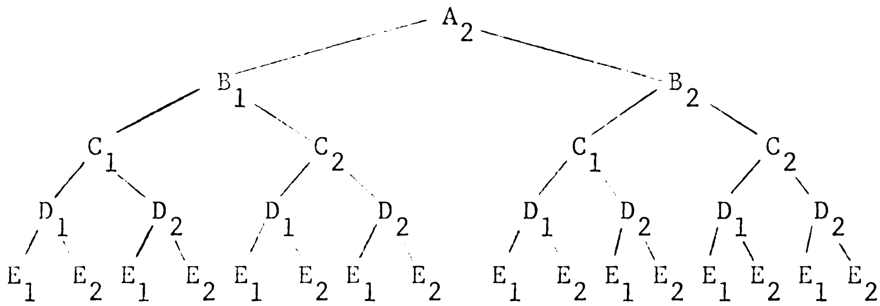
1. CQT-T scores made at the time of entrance to the university.
2. College laboratory courses previous to Natural Science 183, disregarding Natural Science 181 and 182.
3. Sex
4. Raw score on final examination in Natural Science 183.
5. Derived score on final examination in Natural Science 183 on the basis of a 15 point scale.
6. Name of the instructor.

Students not completing the final examination were discarded from the roster of each group and statistical analysis was carried out on the remaining 77 students in both of the experimental groups.

B. Methods of Analysis.--The available data were treated as follows:

1. The CQT-T scores for the treatment groups under each instructor were subjected to analysis of variance to establish whether selection procedures resulted in groups that were reasonably equivalent random samples of the same parent population.
2. The criterion measure (term-end derived examination scores) for the treatment groups under each instructor were subjected to analysis of variance to establish whether there were any difference between treatment groups.
3. The five variables: A(A_1 lecture-demonstration group and A_2 individual laboratory group); B (instructor "B_1" and instructor "B_2"); C (C_1 above median on the CQT-T scores group and C_2 below median on the CQT-T scores group); D (D_1 male and D_2 female); E (E_1 previous college laboratory science and E_2 no previous college laboratory science course groups) were analyzed by a 2^n factorial design of the model:





to establish whether there were any differences or interactions among the five variables.

CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

Johnson cites three requirements to be satisfied by the self-contained experiment.[22] The first is randomization. This is essential in statistical experimentation for tests of significance to validate the estimate of treatment effects to be unbiased. This is done by assuring that whatever source of error may affect the experimental results, also, with equal probability, may affect the estimate of error.

The second requirement of a self-contained experiment is replication. Precision of the experiment depends upon replication. Replication provides the only means of estimating the experimental error. This experimental error decreases in size as the number of replications increases, providing, of course, that there is no increase in heterogeneity of the experimental groups or that there is no greater carelessness in the use of techniques.

The third requirement of a self-contained experiment is control or controls. The control allows the comparison of experimental groups. The control may be another experimental group. All treatments directly compared, including the control, are specified and must be compared upon the same experimental material.

In this third requirement, the control factors were:

---

[22] Palmer O. Johnson, Statistical Methods in Research (New York: Prentice Hall Inc., 1949), p. 282.

1. The two lectures per week were identical.
2. The textbook, the assignment, and the tests were the same.
3. An equal number of clock hours were available for all students.
4. The same experiments with the same type of equipment and laboratory manual were performed.
5. One experimental and three control groups were taught by each of the two instructors.
6. The initial status of the students were examined by analysis of variance and found to be homogeneous.
7. The term-end final, that is the criterion, while not the same, was converted to a derived score on the same basis, so in effect they were equivalent.

Lindquist[23] lists three conditions to be met if a significant F ratio is to be interpreted as evidence that the experimental treatments have different effects. They are:

1. All treatment groups were originally drawn at random from the same population.
2. The variances of the criterion measure are the same for each of these populations.
3. The distribution of criterion measure for each treatment population is normal.

The extent to which these conditions have been met will be discussed in the following analysis of data.

---

[23] E. F. Lindquist, Design and Analysis of Experiments in Psychology and Education (Boston, Houghton Mifflin Co., 1956), p. 73.

I. Analysis of Data

The procedure of sampling in the present study, discussed above, was designed to assure random selection of subjects from the population, thus meeting Lindquist's Condition 1. The original intention was to carry out analysis of variance using the CQT-T scores obtained for each student on entrance to the University. These had been tabulated for each student participating and analysis of variance of these scores for the treatment groups and subgroups under each instructor gave the following results:

TABLE 1.--Analysis of variance of CQT-T scores for each treatment group under each instructor

Instructor	Source	df	SS	Variance	F
"B_1"	Between Groups	1	342	342	.6
	Within Groups	139	77798	560	
	Total	140	78140		
"B_2"	Between Groups	1	64	64	.12
	Within Groups	161	85456	531	
	Total	162	85520		

On examining the above data, there seems to be no difference as to ability level between the control and experimental groups for either instructor. Therefore, in light of the absence of significant differences among groups, no adjustment was considered to be necessary between subgroups for each instructor.

The second condition of Lindquist cited above is "The variance of the criterion measures be the same for each treatment population." He comments as follows[24] "While statistical tests of heterogeneity of variance are available--there will be relatively few situations in

[24] Lindquist, loc. cit., p. 86.

which any such test is required", and again[25] "Fortunately the form of the sample distribution of the mean square ratio is not very markedly affected by the moderate degrees of heterogeneity of variance and hence, the F test may still be satisfactorily used in many experimental situations."

After consultation with David Krathwohl and John Paterson, of Michigan State University Bureau of Educational Research it was decided that a test of homogeneity was then unnecessary because of the manner in which the cases were assigned to the different groups.

No test was made of the normality of the distribution of the criterion scores (equated final examination scores) for each treatment population, since detailed investigation of the influence of non-normality on significance level of the F test indicates that the influence is not great. In this regard Lindquist comments[26] "In general, the F distribution seems so insensitive to the form of the distribution of the criterion measure that it hardly seems worth while to apply any statistical test to the data to detect nonnormality, even though such tests are available."

Taking all of the above into consideration, it was felt that the data met the necessary criteria and an analysis of variance was carried out on the criterion measure (term-end--derived test scores) of each instructor's treatment groups. These results are summarized in Table 2.

On examining the table, it is seen that none of the F values

---

[25]Lindquist, loc. cit., p. 86.

[26]Lindquist, loc. cit., p. 86.

are significant. Therefore, it is concluded that there is no difference as measured by the final examination between the students doing their own laboratory work and the lecture-demonstration for both instructors. The null hypothesis is accepted.

TABLE 2.--Analysis of variance of the criterion measure for the treatment groups under each instructor

Instructor	Source	df	SS	Variance	F
"B_1"	Between Groups	1	7	7	.93
	Among Groups	138	1032	7.5	
	Total	139	1039		
"B_2"	Between Groups	1	6	6	.61
	Among Groups	161	1575	9.8	
	Total	162	1581		

In an attempt to investigate further whether the ability of the students, instructor, sex, or previous college laboratory science courses have any bearing on their performance in Natural Science 183, a factorial design was worked out. This design met the requirements of a self-contained experiment according to Johnson and Lindquist as set forth earlier. The factorial design of five variables: A, B, C, D, and E that was set up was a 2 by 2 by 2 by 2 by 2. Each main variable is varied in two ways making a total of 32 subgroups. The five variables that were studied are: A, the two methods, experimental and control groups. B, the two instructors. C, the above median on the CQT-T scores and the below median on the CQT-T scores. D, sex; E, previous laboratory science courses and no previous laboratory science course, prior to Natural Science 183. The factorial design has a number of advantages. Edwards[27] stated them as follows:

---

[27] Allen L. Edwards, Experimental Design in Psychological Research (New York: Rinehart and Co., Inc., 1950), pp. 232-233.

1. The full numbers of subjects enter into every comparison.
2. The sum of squares which is used as an estimate of uncontrolled variation of subjects treated alike is based on a large number of degrees of freedom.
3. Information is obtained not only about several variables but also thc interaction between variables.
4. The over-all advantage is that it provides a sounder basis for generalizing about the effectiveness of the experimental variables since they are tested not only in isolation, but in conjunction with the effects of the other variables.

The assumptions involved in pooling higher order interaction are, according to Edwards:[28]

1. Each of the mean squares corresponding to the order interaction is an estimate of the same common variance.
2. The common variance would not differ significantly from the variance obtained with replication.
3. The number of cases in each class is equal.

Since the number of cases in each class was not equal, it was necessary to adjust the data. This was accomplished as follows: Assume that the mean ($\overline{X}$) and the variance (s^2) are good estimates, then solve for a harmonic mean of the number of cases and use this as the N for each group. To solve for the ΣX_a adjusted, use the formula, $\Sigma X_a = N_a\overline{X}$ where $\overline{X}$ is the mean. The ΣX_a^2 is found by substituting in the formula $\Sigma X_a^2 = N_a(s_x^2 + \overline{X}^2)$. The above method of adjustment for unequal N's was suggested by John Paterson, the research associate of

---

[28] <u>Ibid</u>., pp. 254-256.

the Bureau of Educational Research, Michigan State University.

The analysis given in Table 3 is on the basis of adjusted values. There were 32 groups and 202 cases on the adjusted basis.

TABLE 3.--Two part analysis of variance of 32 groups of subjects tested under different experimental conditions

Source of Variation	Sum of Squares	df	Variance	F
Between Groups	521	31	16.81	1.96**
Among Groups	1469	171	8.59	
Total	1990	202		

The hypothesis of random sampling from a normal population is rejected. There are group differences which may be explained by more detailed analysis.

Before proceeding with the partitioning, the sum of squares between groups in the experiment, the homogeneity of the variances combined into the within group variance need to be tested to see if the hypothesis of homogeneity was tenable.

TABLE 4.--Bartlett's test of homogeneity for 32 groups

X^2 Calculated	X^2 tab. .05, 31 df
.061	44.000

Since chi square (X^2) has a value much less than that associated with the probability of .05 for 31 degrees of freedom, it is not regarded as significant. Hence, it was concluded that the variation values of the sum of squares is within the limits of random sampling from a population with the common variance and that the hypothesis of homogeneity of variance is therefore tenable. The within group sum of squares was then usable as the error term.

II. Partitioning the Sum of Squares Between Groups

In analysis of this sort, the sum of squares between groups may be partitioned into as many component parts as there are degrees of freedom associated with it. In this experiment, the sum of squares between groups is based on 31 degrees of freedom and may be analyzed into 31 meaningful parts. Each part is based on a single degree of freedom. Actually, the partition is similar to the following model.

SOURCE OF VARIATION

Main Variable	A	1
	B	1
	C	1
	D	1
	E	1
Simple Interaction	AxB	1
	AxC	1
	AxD	1
	AxE	1
	BxC	1
	BxD	1
	BxE	1
	CxD	1
	CxE	1
	DxE	1
Second-order Interaction	AxBxC	1
	AxBxD	1
	AxBxE	1
	AxCxD	1
	AxCxE	1
	AxDxE	1
	BxCxD	1
	BxCxE	1
	BxDxE	1
	CxDxE	1
Third-order Interaction	AxBxCxD	1
	AxBxCxE	1
	AxBxDxE	1
	AxCxDxE	1
	BxCxDxE	1
Fourth-order Interaction	AxBxCxDxE	1

TABLE 5.--Analysis of variance of criterion scores for 32 groups of subjects tested under different conditions (interactions). Partitioned sum of squares of between groups--Table 3

Source of Variation	Sum of Squares	df	Mean Square	F
A-Exper. & Control Groups	2.00	1	2.00	--
B-Two Instructors	40.00	1	40.00	4.65*
C-Above & Below Median CQT-T	218.30	1	218.30	25.38**
D-Sex	21.60	1	21.60	2.51
E-Prev. vs. No Prev. Lab Courses	49.50	1	49.50	5.76*
Combined Interactions	189.8	26	7.3	--
Among Groups	1469.00	171	8.59	

Edwards[29] gives the following conditions for use of an interaction or pooled interaction mean square instead of the usual residual mean square among groups, they are:

1. The interaction mean square must be larger than the residual mean square.

2. The categories of one or more of the variables in the experimental design must be a random selection from the population.

All the variables but the instructor are random selection from the population because of the way the students were assigned at registration as was stated earlier.

When the pooled simple and higher order interaction mean squares was examined, it was found to be smaller than the residual mean square. This meant that the main variables could not be examined on the basis of the combined interaction mean square rather than the residual mean square because the first condition as stated by Edwards above is not

[29] Ibid., p. 251.

satisfied. On further examination it was found that the combined interaction mean square of the simple and higher order interactions along with the mean square of the main variables that were not significant is still smaller than the residual mean square. If however, we pool each of the significant variables in turn with the rest of the non-significant main variables and the pooled interactions of the second-order and higher, Edward's first condition is met. This serves in place of replication and permits the assumption that this common variance would not differ significantly from the variance obtained by replication. It permits the examination for significance of the three pairs of main variables in turn using the combined interactions as an error term. The results appear in the table below.

TABLE 6.--Analysis on the basis of pooled interactions

Source	SS	df	Mean Square	F
C	218.30	1	218.30	25.97**
E	49.50	1	49.50	5.66
Pooled Int. + ABD	253.40	29	8.74	
B	40.10	1	40.10	4.43*
C	218.30	1	218.30	24.09**
Pooled Int. + ADE	262.80	29	9.06	
B	40.10	1	40.10	2.69
E	49.50	1	49.50	3.33
Pooled Int. + ACD	431.60	29	14.88	

With the information from Table 6 the main variables B, C, and E are examined individually in Table 7. Variable C is the only one which remains significant.

TABLE 7.--Analysis of three main variables individually of the basis of pooled interactions

Source	SS	df	Mean Square	F
B	40.10	1	40.10	2.50
Pooled Int.	481.10	30	16.04	
C	218.30	1	218.30	21.61**
Pooled Int.	309.90	30	10.10	
E	49.50	1	49.50	3.15
Pooled Int.	471.70	30	15.72	

Variable C involved the CQT-T scores. Since it was highly significant when examined against the pooled interactions of the other variables it was examined further in Table 8.

TABLE 8.--Analysis of variance of the two groups above-median CQT-T and below-median CQT-T on basis of the criterion scores (final examination derived scores) for each instructor

Source		SS	df	Mean Square	F
Instructor "B_1"	Total	716			
	AM-BM				
	Between	4	1	4	.56
	Among	712	100	7.12	
Instructor "B_2"	Total	1017			
	AM-BM				
	Between	170	1	170	20.00**
	Among	847	100	8.47	

On re-examining Table 1, page 35, the analysis of variance of the CQT-T scores for both instructors, it is found that there was no significant difference between treatment groups. Table 8, however, seems to indicate that for instructor B_2 the above-median group (CQT-T) did better for both treatments (experimental and control) on the final examination than did the below-median group (CQT-T) when measured against the criterion (final examination). This does not explain the highly significant result for the C variable (CQT-T) in Tables 6, 7, and 8.

However, the SS for instructor B_2 the A-M group SS was 486 and the B-M group SS was 355. This appears to indicate that the A-M group of instructor B_2 learned a great deal more than the B-M group as measured by the term-end examination. Tables 7 and 8 then would seem to indicate that the highly significant F found in Table 3, page 39 was due almost entirely to the interactions between the A-M group and the B-M group of instructor B_2 with the other main variables.

III. General Summary

The following general conclusions appear to be supported by the data:

I. For randomly selected groups under instructors B_1 and B_2:

1. The data satisfies all conditions necessary for valid use of the F ratio as a test of the null hypothesis.
2. For the unpartitioned and partitioned treatment groups under any one instructor, differences in taking the laboratory failed to produce statistically significant differences in student achievement as measured by the term-end examination.

II. For randomly selected groups combined and adjusted for unequal N's in each sub-class analyzed by the factorial design:

1. The data satisfies all conditions necessary for valid use of the F ratio as a test of the null hypothesis.
2. The partitioned main variables and higher order interactions failed to produce statistically significant differences in the two main variables--method and sex while the three main variables--instructor, CQT-T, and previous vs. no-previous laboratory courses produced statistically significant differences.
3. On pooling interactions and testing each main variable, no statistically significant differences were found for:

   (a) Method-A

   (b) Instructor-B

   (c) Sex-D

   (d) Previous vs. no previous laboratory courses-E

   but a statistically significant difference was found for the CQT-T--C.
4. On testing the above median (A-M) against the below-median (B-M) for CQT-T for both instructors, it was found that there was no statistically significant difference for instructor B_1 while for instructor B_2 it was statistically significant at the 1% level of confidence.

The significance of these findings for the originally stated problem will be considered in the following chapter.

CHAPTER V

SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

I. General

As all present day knowledge of the natural sciences is due to experimentation, it seems natural that the students should have first hand experience designed to promote an understanding of what science is and what science does. This seems to require some form of laboratory method. But with pressures building up in the colleges and universities, such as expanding enrollment, lack of instructors, and failure of the physical plant to keep pace with enrollment, it is necessary to examine alternate methods of teaching science other than the laboratory method; a method which will save instructor time, space, and cut down on the overhead of servicing and maintaining laboratories.

Since introductory science courses and courses used for general education have large enrollments they should be the place to conduct method studies which evaluate desired outcomes in terms of the different methods. This study is one attempt to gather pertinent data on the efficiency of lecture-demonstration method vs. individual laboratory method.

II. Summary

When the study reported here was designed, it was felt that the below median ability group would benefit the most by having their laboratory exercises demonstrated to them, especially the women who had no previous college laboratory science. While this expectation was not supported by previous studies it wasn't denied. There seemed about as much evidence in favor as against. However, when statistical analyses consistently showed lack of significant differences among means of treatment groups, it became apparent that further examination of the literature was in order. As a result all references cited were read with rather a different and more critical viewpoint. The studies that reported no differences were re-examined in light of the number of variables held constant and were found to be much more detailed and consistent as to the number of these variables than did the studies favoring the laboratory method. The studies favoring either the lecture-demonstration method or the individual laboratory method were also re-examined more critically. The studies favoring the laboratory method did so primarily on the basis that it was more efficient for teaching various phases of laboratory manipulation and problem solving situations involving apparatus. The studies that consistently found in favor of the lecture-demonstration method were characterized as a group by inadequate controls.

Thus it was seen after a more critical examination of the research that there was no support for the expectation that below median ability group of women with no previous laboratory science experience would benefit by the lecture-demonstration method. The research studies did however support the conclusions of this study.

Since manipulative skills are not an objective of general education science the evidence for supporting the use of individual laboratory does not seem at all tenable. Therefore an analysis of variance was run on the data and later the data were further analyzed by means of a factorial design consisting of five main variables varied two ways to provide more evidence in the drawing of conclusions.

III. Findings

1. After verifying that there was no difference between the ability level of the two instructors' groups--lecture-demonstration and individual laboratory--as measured by the "T" score of the College Qualification Tests it was found that the two instructors' lecture-demonstration groups did equally as well on the term end test as did their individual laboratory groups.

2. Further analysis was made by a factorial design that measured the interactions of five variables: teaching method, instructor, ability level, sex, and previous laboratory science experience. Each variable had two values. Homogeneity of the 32 groups was found to be tenable so the highly significant difference found for the between groups was broken down into 31 meaningful parts.

(a) There was interaction between groups for the variables: instructor, ability level, and previous college laboratory experience.

(b) There was no interaction between groups for sex and teaching method variables.

(c) None of the simple or higher order interactions were significant.

(d) When the variance due to four of the variables was pooled in turn with the simple and higher order interactions only the instructor variable was significant.

IV. Conclusions

On the basis of the results of this study it was valid to conclude that: first, both the lecture-demonstration method and the individual laboratory method were just as effective means of teaching Natural Science 183 for the aims and objectives of General Education Science as measured by the term-end examination; second, the experimental evidence appeared to justify the contention that there was no interaction between methods, sex, above median on CQT-T group and the below median on the CQT-T group, and those having previous college laboratory science compared to those that had no previous laboratory science courses.

Although neither method appeared to give better results for either instructor on the basis of the criterion measure, the null hypothesis of no interaction between instructors and the other variables was rejected. On further examination and as a result of analysis of variance, the null hypothesis of no difference between the instructor's above median and below median group on the CQT-T scores with the criterion as the term-end examination was rejected for one instructor and accepted for the other. Examining the sums of squares for the rejected instructor's two groups it was concluded that the above median group on the CQT-T scores achieved more than was expected of them for the one instructor.

As an interesting side light the two instructors were asked to state their philosophy of teaching as regard to ability levels of students as measured by the CQT. The instructor whose above median group was significant stated:

"I tend to teach only for the above average group."

while the other instructor said that he tended to teach for the "dull average" student. The evidence of the study seems to indicate that the results they achieve with their classes agree with their philosophy.

V. Implications

The outstanding result of this study is that the particular methods of teaching differ very little as evidenced by student learning and retention, when the aims and objectives of General Education Science are examined on the basis of a paper and pencil test. It may well be that many methods have their place in science teaching. If so, the proper functions of each should be determined.

It may be on the other hand, that the failure to reveal a clear superiority of either method was due to the limitations in experimental techniques employed. Possibly achievement of objectives specific to the laboratory was not measured with sufficient accuracy to reveal tendencies.

The seeming lack of difference in the two methods may however, mask their true significance. Since the lecture-demonstration method, which held its own with the individual laboratory method as a means to immediate learning, makes possible a considerable savings in apparatus, physical plant and instructor time, this may offset any advantages of the individual laboratory method not revealed or considered in this study.

VI. Suggestions for Further Research

There are many aspects of science teaching that were not included in the scope of this study. For instance no attempt was made to evaluate in any form the understanding of the relationship of science to other areas.

Following are typical problems left unanswered by this study.

(1) For what particular laboratory objectives is the lecture-demonstration method superior?

(2) Would the result of an experiment involving science oriented students be the same?

(3) Would the difference between instructors become more pronounced through special training in demonstration techniques?

(4) How can the reliability and validity of the term-end examination be raised in terms of measuring the aims and objectives of general education science?

(5) Might some other form of testing such as tests giving immediate knowledge of results be superior?

(6) Would some form of the problem solving method be superior to both the lecture-demonstration method and the individual laboratory method?

These and other questions deserve further study. However, in so far as this study provides evidence, individual laboratory experience for the attainment of general education goals does not appear to be necessary.

BIBLIOGRAPHY

Barnard, J. David. "The Lecture-Demonstration vs The Problem Solving Method of Teaching a College Science Course," Science Education, Vol. 26, No. 3 + 4, 121-132.

Cunningham, Harry A. "Lecture Demonstration vs. Individual Laboratory Method in Science Teaching-A Summary," Science Education, Vol. 30, No. 2, 70-82.

Edwards, Allen L. Experimental Design in Psychological Research. New York: Rinehart and Co., 1950.

Hurd, Archer W. Problems of Science Teaching at the College Level. University of Minnesota Press, 1929.

Johnson, Palmer O. "Comparison of the Lecture-Demonstration, Group Laboratory Experimentation, and Individual Experimentation Method of Teaching High School Biology," Journal of Educational Research, Vol. XVIII, No. 2, 103-111.

Johnson, Palmer O. Statistical Methods in Research. New York: Prentice-Hall, 1949.

Knox, W. W. "The Demonstration Method vs. the Laboratory Method of Teaching High School Chemistry," The School Review, Vol. 25, (May 1927), pp. 376-386.

Kruglak, Hayam. "A Comparison of the Conventional and Demonstration Methods in the Elementary College Physics Laboratory," Journal of Experimental Education, Vol. XX (March 1952), 293-300.

Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Boston, Mass.: Houghton Mifflin Co., 1956.

Monroe, W. S. and Max D. Engelhart. "Experimental Research in Education," University of Illinois Bulletin, Vol. 27, No. 32, Bureau of Educational Research Bulletin No. 48 (Urbana: University of Illinois Press, 1930) pp. 18-58.

Ward, John. "Group-Study versus Lecture-Demonstration Method in Physical Science Instruction for General Education College Students," Journal of Experimental Education, Vol. 24, (March, 1956), 197-210.

White, John R., "A Comparison of the Laboratory and the Lecture-Demonstration Methods in Engineering Instruction," an unpublished Doctor's Dissertation, New York University, 1943.

APPENDIX

TABLE 9.--Primary data for treatment group under instructor "B_1"

Experimental					Control				
CQT-T	Term-end Raw Score	Term-end Equated Score	Sex	PLC*	CQT-T	Term-end Raw Score	Term-end Equated Score	Sex	PLC*
77	51	8	F	+	137	49	8	F	+
140	51	8	F	-	118	39	5	F	-
87	33	3	F	-	114	40	5	F	-
130	64	12	F	-	91	34	3	F	-
84	46	8	F	-	97	53	9	F	-
116	37	4	F	+	130	31	2	F	-
134	40	5	F	-	118	47	8	F	-
93	40	5	F	-	143	62	11	F	-
127	48	8	F	-	94	43	7	F	-
134	51	8	F	-	65	25	2	F	-
109	47	8	F	-	81	42	6	F	-
127	75	14	F	-	150	59	11	F	-
110	42	6	F	-	114	52	9	F	-
121	50	8	F	+	107	56	10	F	-
90	39	5	F	-	132	37	4	F	-
83	33	3	F	-	117	35	4	F	-
101	33	3	F	+	149	45	7	F	-
107	49	8	F	-	80	40	5	F	-
107	28	2	F	-	85	48	8	F	-
110	40	5	F	-	108	43	7	F	-
124	52	9	F	-	141	57	11	F	-
139	41	6	F	-	97	45	7	F	-
100	38	5	F	-	142	37	4	F	-
88	35	4	F	-	127	61	11	F	-
77	30	2	F	+	96	32	2	M	-
127	40	5	F	+	136	60	11	M	+
139	54	9	F	+	156	58	11	M	+

87	35	4	F	-	116	47	8	M	-
155	52	9	F	-	115	47	8	M	-
133	41	6	F	-	159	50	8	M	-
141	50	8	M	-	110	32	2	M	-
121	54	9	M	-	89	37	4	M	-
119	41	6	M	-	116	29	2	M	-
131	59	11	M	-	118	57	11	M	-
123	59	11	M	+	144	48	8	M	+
56	46	8	M	-	160	62	11	M	+
169	66	12	M	-	176	69	14	M	+
134	61	11	M	+	120	54	9	M	+
100	34	3	M	-	75	39	5	M	-
97	44	7	M	-	111	63	12	M	-
146	66	12	M	+	100	54	9	M	+
75	47	8	M	-	111	43	7	M	-
128	54	9	M	+	137	66	12	M	+
127	47	8	M	-	119	45	7	M	-
93	48	8	M	-	138	54	9	M	+
111	37	4	M	-	158	48	8	M	-
148	49	8	M	-	130	51	8	M	-
126	58	11	M	-	138	75	14	M	-
112	41	6	M	-	100	34	3	M	-
89	39	5	M	-	145	61	11	M	+
111	53	9	M	-	118	41	6	M	-
118	54	9	M	-	176	61	11	M	-
138	43	7	M	-	95	57	11	M	-
137	62	11	M	+	109	46	8	M	-
157	71	14	M	+	147	50	8	F	+
89	53	9	M	-	121	27	2	F	-
149	62	11	M	-	100	44	7	M	+
151	44	7	M	-	128	39	5	M	-
108	43	7	M	-	142	58	11	M	-
113	40	5	M	+	163	53	9	M	+
121	40	5	M	+	107	51	8	F	+
159	38	5	M	-	99	57	7	F	+
134	43	7	M	-	109	62	8	F	+

TABLE 9.--Continued

Experimental					Control				
CQT-T	Term-end Raw Score	Term-end Equated Score	Sex	PLC*	CQT-T	Term-end Raw Score	Term-end Equated Score	Sex	PLC*
110	48	8	M	+					
150	64	12	M	+					
160	43	7	M	+					
107	70	14	M	+					
84	27	2	M	-					
122	37	4	M	+					
89	46	8	M	-					
135	51	8	M	-					
117	46	8	M	-					
91	48	8	M	+					
103	55	10	F	-					
122	29	2	F	-					
140	56	10	M	-					
124	40	5	F	-					

TABLE 10.--Primary data for treatment group under instructor "B_2"

Experimental					Control				
CQT-T	Term-end Raw Score	Term-end Equated Score	Sex	PLC*	CQT-T	Term-end Raw Score	Term-end Equated Score	Sex	PLC*
132	69	10	M	-	141	84	14	M	-
116	61	8	M	-	122	70	10	F	-
144	68	9	M	-	83	79	12	F	+
115	68	9	F	+	89	51	5	F	-
132	58	7	F	-	83	54	6	M	-
157	76	12	F	-	108	78	12	M	+
103	60	7	F	-	136	58	7	M	-
120	63	8	M	-	77	63	8	F	+
126	81	13	M	+	89	51	5	M	-
126	52	5	M	-	98	45	3	F	+
119	52	5	M	-	128	48	4	F	-
164	68	9	M	+	113	65	9	F	+
144	77	12	F	-	96	69	10	F	-
96	66	9	M	-	157	53	5	F	-
109	66	9	F	-	141	65	9	F	-
93	61	8	M	+	111	65	9	F	-
139	59	7	F	-	145	82	13	F	-
99	58	7	M	-	128	59	7	F	+
87	56	6	F	-	99	59	7	F	-
147	66	9	F	-	114	47	4	F	-
130	59	7	F	-	136	67	9	F	-
93	63	8	F	-	142	62	8	F	+
113	68	9	F	-	117	57	7	F	-
135	58	7	F	-	141	76	12	F	-
125	69	10	F	-	125	73	11	M	+
171	86	14	F	-	103	60	7	M	-
112	59	7	F	-	104	68	9	M	+

131	76	12	F	+	157	73	11	M	+
139	65	9	F	+	100	59	7	M	+
120	67	9	F	-	172	51	5	M	-
169	93	15	F	-	145	66	9	M	+
103	56	6	F	-	135	68	9	M	+
142	79	12	F	-	124	64	8	M	+
79	59	7	F	-	128	74	11	M	+
132	72	11	F	+	157	69	10	M	-
126	69	10	M	-	140	65	9	M	+
134	66	9	M	+	112	73	11	M	+
103	65	9	M	-	152	73	11	M	+
121	67	9	M	+	96	36	1	M	-
153	78	12	M	+	88	48	4	M	-
125	62	8	M	-	161	65	9	M	+
132	36	1	M	+	95	45	3	M	-
106	41	1	M	+	120	61	8	M	+
144	82	13	M	+	169	74	11	M	+
165	81	13	M	-	86	50	4	M	-
124	64	8	M	+	148	73	11	M	+
77	64	8	M	-	148	77	12	M	-
130	79	12	M	+	99	66	9	F	-
136	53	5	M	+	151	70	10	F	-
114	70	10	M	+	87	68	9	F	-
91	50	4	M	-	138	64	8	F	-
139	70	10	M	+	127	51	5	F	-
121	58	7	M	-	118	57	7	M	+
124	86	14	M	-	148	90	15	M	+
122	63	8	M	+	102	62	8	M	+
150	58	7	M	+	136	75	12	M	+
123	59	7	M	-	100	59	7	M	-
115	57	7	M	-	94	72	11	M	-
161	94	15	F	+	103	48	4	F	-
131	51	5	F	-	103	40	1	F	-
117	59	7	F	+	99	50	4	F	+
116	62	8	F	+	164	83	13	F	-
96	47	4	F	-	115	60	7	F	-

TABLE 10.--Continued

Experimental					Control				
CQT-T	Term-end Raw Score	Term-end Equated Score	Sex	PLC*	CQT-T	Term-end Raw Score	Term-end Equated Score	Sex	PLC*
141	75	12	F	-	128	59	7	F	-
144	63	8	F	-	95	70	10	F	+
131	70	10	F	-	109	59	7	F	-
91	56	6	F	-	167	96	15	F	+
121	63	8	F	+	93	48	4	F	-
160	81	13	F	-	117	61	8	F	-
159	81	13	F	-	156	74	11	F	+
86	46	3	F	-	142	86	14	F	+
120	76	12	F	-	106	53	5	F	-
105	65	9	F	-	142	55	6	F	-
145	66	9	F	-	141	70	10	F	-
99	63	8	F	-	117	46	3	F	-
114	55	6	F	+	131	66	9	F	-
82	56	6	F	-	159	82	13	F	-
					109	68	9	F	-
					140	75	12	M	-
					155	72	11	M	-
					119	54	6	M	+
					128	77	12	M	-
					155	59	7	M	-
					131	76	12	M	+
					113	53	5	F	-
					99	67	9	F	+