1953

Determination for the modal age level for grades seven and nine of the difficulty of the principle: gases, liquids and solids expand when heated and contract when cooled.

Creighton, William Andrew
Boston University

https://hdl.handle.net/2144/14985
Boston University
BOSTON UNIVERSITY
SCHOOL OF EDUCATION

Thesis

DETERMINATION FOR THE MODAL AGE LEVEL FOR
GRADES SEVEN AND NINE OF THE DIFFICULTY OF
THE PRINCIPLE: GASES, LIQUIDS AND SOLIDS
EXPAND WHEN HEATED AND CONTRACT WHEN COOLED

Submitted by

William Andrew Creighton
(A.B., Boston University 1950)

In Partial Fulfillment of Requirements for
the Degree of Master of Education

1953
First Reader: John G. Read
Professor of Education

Second Reader: James F. Baker
Assistant Professor of Education
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Justification</td>
<td>4</td>
</tr>
<tr>
<td>Design of the Experiment</td>
<td>14</td>
</tr>
<tr>
<td>Scope and Limitations</td>
<td>14</td>
</tr>
<tr>
<td>Discussion of Procedure</td>
<td>16</td>
</tr>
<tr>
<td>II. SURVEY OF RELATED LITERATURE</td>
<td>28</td>
</tr>
<tr>
<td>Teaching By the Use of Principles</td>
<td>28</td>
</tr>
<tr>
<td>The inductive method</td>
<td>28</td>
</tr>
<tr>
<td>The deductive method</td>
<td>29</td>
</tr>
<tr>
<td>Some results</td>
<td>32</td>
</tr>
<tr>
<td>The Lecture-Demonstration Method of Teaching</td>
<td>33</td>
</tr>
<tr>
<td>The Effectiveness of Lecture-Demonstrations</td>
<td>33</td>
</tr>
<tr>
<td>The Criteria for a Good Demonstration</td>
<td>48</td>
</tr>
<tr>
<td>Statement of the problem</td>
<td>48</td>
</tr>
<tr>
<td>Need for research</td>
<td>49</td>
</tr>
<tr>
<td>Definition of demonstration</td>
<td>50</td>
</tr>
<tr>
<td>Review of the literature</td>
<td>53</td>
</tr>
<tr>
<td>Treatment of the data</td>
<td>59</td>
</tr>
<tr>
<td>Selected criteria</td>
<td>60</td>
</tr>
<tr>
<td>The Test Technique</td>
<td>65</td>
</tr>
<tr>
<td>Structure of the Test</td>
<td>65</td>
</tr>
<tr>
<td>Multiple choice items</td>
<td>65</td>
</tr>
<tr>
<td>Levels of difficulty</td>
<td>66</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>68</td>
</tr>
<tr>
<td>The test tryout</td>
<td>69</td>
</tr>
<tr>
<td>The test period</td>
<td>70</td>
</tr>
<tr>
<td>Aims and Use of the Test</td>
<td>71</td>
</tr>
<tr>
<td>Employing statistics</td>
<td>71</td>
</tr>
<tr>
<td>Test-retest method</td>
<td>72</td>
</tr>
<tr>
<td>What the test endeavors to determine</td>
<td>73</td>
</tr>
<tr>
<td>Characteristics of the Test</td>
<td>74</td>
</tr>
<tr>
<td>Reliability</td>
<td>74</td>
</tr>
<tr>
<td>Validity</td>
<td>75</td>
</tr>
</tbody>
</table>
### Chapter III: Experimental Procedure

#### Introduction

- Page 76

#### Description of Schools

- Town A: Page 76
- Town B: Page 76
- Distribution of pupils in Cooperating Schools: Page 77

#### Classroom Procedure

- Page 79

#### The Test

- Page 80

#### The Demonstration

- Page 80

#### Obtaining and Recording Test Results

- Record Card: Page 81
- Modal Mental Age: Page 82
- Recording of Data: Page 89
- Code used in reading the charts: Page 89

### Chapter IV: Summary and Findings

- Otis Normal Percentile Charts: Page 91
  - Chart I: Page 91
  - Chart II: Page 92
  - Chart III: Page 91
  - Chart IV: Page 91
  - Chart V: Page 92
  - Chart VI: Page 92
  - Chart VII: Page 92
- Recapitulation: Page 92

### Appendix

- Appendix A - Otis Normal Percentile Charts and Worksheets: Page 94
- Appendix B - Testing Materials: Page 109
  - The Test: Page 110
  - The Answer Sheet: Page 117
  - The Record Card: Page 119
  - The Script: Page 120
Appendix C - Bibliography................. 125
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Distribution of Pupils According to Schools in the Control and Experimental Groups</td>
<td>78</td>
</tr>
<tr>
<td>2.</td>
<td>Distribution of Pupils According to Sex in the Control and Experimental Groups</td>
<td>78</td>
</tr>
<tr>
<td>3.</td>
<td>Distribution of Time during the Classroom Period</td>
<td>79</td>
</tr>
<tr>
<td>4.</td>
<td>Comparison of Mental Ability in Control and Experimental Groups of the Seventh Grade</td>
<td>83</td>
</tr>
<tr>
<td>5.</td>
<td>Comparison of Mental Ability in Control and Experimental Groups of the Ninth Grade</td>
<td>85</td>
</tr>
<tr>
<td>6.</td>
<td>Number of Pupils in the Sub-groups of the Seventh Grade Control and Experimental Groups</td>
<td>86</td>
</tr>
<tr>
<td>7.</td>
<td>Number of Pupils in the Sub-groups of the Ninth Grade Control and Experimental Groups</td>
<td>86</td>
</tr>
<tr>
<td>8.</td>
<td>A Comparison of the Modal Chronological Ages and the Modal Mental Ages of the Three Sub-groups in the Seventh Grade</td>
<td>88</td>
</tr>
<tr>
<td>9.</td>
<td>A Comparison of the Modal Chronological Ages and the Modal Mental Ages of the Three Sub-groups in the Ninth Grade</td>
<td>88</td>
</tr>
<tr>
<td>TABLE</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>I. The Frequency of Mention of the Selected Criteria by the Sources Consulted.</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1.</td>
<td>Modal Mental Age Group Results on the Pre-test for Both the Seventh and Ninth Grades</td>
<td>95</td>
</tr>
<tr>
<td>2.</td>
<td>Modal Mental Age Group Results on the Post-test for Both the Seventh and Ninth Grades</td>
<td>97</td>
</tr>
<tr>
<td>3.</td>
<td>Modal Mental Age Experimental Group Results on the Pre-test and Post-test for Both the Seventh and Ninth Grades</td>
<td>99</td>
</tr>
<tr>
<td>4.</td>
<td>Underage Deviate Experimental Group Results on the Pre-test and Post-test for Both the Seventh and Ninth Grades</td>
<td>101</td>
</tr>
<tr>
<td>5.</td>
<td>Overage Deviate Experimental Group Results on the Pre-test and Post-test for Both the Seventh and Ninth Grades</td>
<td>103</td>
</tr>
<tr>
<td>6.</td>
<td>Seventh Grade Experimental Group Results on the Post-test for Each Mental Age Level</td>
<td>105</td>
</tr>
<tr>
<td>7.</td>
<td>Ninth Grade Experimental Group Results on the Post-test for Each Mental Age Level</td>
<td>107</td>
</tr>
<tr>
<td>9.</td>
<td>Record Card</td>
<td>117</td>
</tr>
<tr>
<td>Plate</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Lecture-Demonstration Apparatus</td>
<td>81</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1. Introduction

The grade placement of scientific principles has been undertaken by the Boston University School of Education, Science Department, as a project to be carried on through the succeeding years under the direction of Dr. John G. Read. The experiment will be concluded when enough pupils and schools have been included in the study to make the conclusions statistically significant. As more data are accumulated, the extent to which the study will have to be carried should become apparent. When a sufficient number of principles has been tested a complete overall report can be written giving the conclusions. That is, there will be a "percentage of learning" index for each principle for each modal mental age level. From this index it should be possible to ascertain that if a certain principle is taught to a certain modal mental age level then a certain percentage of the pupils can be expected to learn the principle.

This year both the background for the study and the design of the experiment are being done under the guidance of Dr. Read and Mr. Herbert Oxendine. Also, the first group will start the experiment with each experimenter
taking one principle. The teaching method will be a lecture-
demonstration. The amount of learning will be measured by
identical tests given before and after the lecture-demonstra-
tion. Data to be gathered will consist of the test scores,
the pupil's I.Q., M.A., C.A., sex, previous science instruc-
tion and science background. The results to be found by
each individual participating in this study will be for only
one or two grades in several schools. This will give an in-
dication of the next grade that should be tested depending
upon the "percentage of learning" for that modal mental age
level.

As the study proceeds investigation into these prin-
ciples will be continued and others started until there is
an index of the "percentage of learners" for each modal
mental age level for each principle. Each experimenter
starting a new principle will leave his material for many
teachers-in-service to use when he is finished. This will
include the demonstration, a copy of the script, a tape
recording of one of the actual lecture-demonstrations and
the test. The same material will be used throughout the
study for the same principle. The sample of schools will be
chosen each year so that complete coverage may be made of
each socio-economic level for grades 3-12.

It is assumed for this study that there will exist a
difference in the percentage of learners at different mental age levels. It is also assumed that the time spent in a good demonstration with a carefully prepared talk would produce a small increment of learning.

The committee whose responsibility it was to compile the data included in the literature as background for this study consisted of the following members under the chairmanship of Norman G. Mills:

Isabel L. Bouin
John T. Callahan
James Creighton
Wallace J. Gleekman
Eugene H. Goldrick
George F. Griffin
Robert H. Jackman
Eleanor Kanevitch
John G. Minot
Henning A. Sahlberg
Vincent J. Silluzio
Schuyler G. Slater
Virginia M. Wilson
2. Justification

Very little scientific evidence is available on the grade placement of science principles. Because of the great increase of scientific knowledge, educators emphasize the need for research that will determine the age levels at which science concepts, principles and skills may be introduced into the curriculum with optimum effectiveness.

Beck\(^1\) states that because scientific knowledge is accumulating at such a rapid rate, there is neither time nor excuse for teaching the elementary scientific concepts in the higher grades. He points out that the scientific background and foundation prerequisite to an understanding of the individual science courses offered in the high schools are lacking in the beginning students. To find a solution to this problem, he suggests that research be started to determine, "...what fundamentals of science can we expect most children of similar ability and cultural background to master at each maturity level".\(^2\)

From a Progress Report of the Committee on Research in Elementary Science for the National Association for Research in Science Teaching, Venill\(^3\) believes that with

\(^1\) Alfred D. Beck, "Some Unanswered Questions Pertaining to the Organization of a Twelve Year Science Sequence", *Science Education* (April, 1948), 34:176-177.

\(^2\) Ibid., p. 177.

the great expansion of scientific knowledge, concepts which previously have been reserved for high school science courses will have to be taught in the junior high schools. He summarizes that, "...studies should be made on pupil readiness for more advanced science concepts." 1/

In the Thirty-first Yearbook the National Society for the Study of Education 2/ suggests that a twelve-year sequence of science be taught, based on the broad generalizations of science. As an outgrowth of this plan, many problems for research were recognized. Morrison 3/places the selection and sequences of courses within the curriculum and the grade placement of topics at the top of a list of needed research in science teaching.

However, research on the location of curricular material in science classes is complex. In order to make such studies objective and meaningful, educators, says Bellack, must take into consideration the basic findings from the fields of educational philosophy, sociology, child growth and development and psychology of learning. 4/

1/ John Venill, op. cit., p. 175.
3/ Ibid., p. 354.
The writer believes that the aim of education is to give some meaning, some security and purpose in life. Ideally, education should provide an understanding of the diversity and richness of the present-day world and take into account our uncertainty on ideas of life and the universe. More specifically, science education seeks to teach effectively those principles and skills of science which touch so largely upon everyday life. The aims of science teaching are contributory to the aims of education, mainly, as Bellack says, "life enrichment."\(^1\)

In our society great emphasis is placed on education. Laws compel schooling up to a certain age, and all children are assured a free education. It is the school's responsibility both to society and to the children to present those activities which will prepare the individual student to participate intelligently in our democratic society.

Bellack\(^2\) further says, "In planning the sequence and placement of school experiences, then, consideration must be given at every stage to the demands of society in regard to both the important responsibilities of citizenship and

\(^1\) Arno A. Bellack, op. cit., p. 42.
\(^2\) Ibid., p. 623.
the great variety of learnings and adjustments occasioned by circumstances peculiar to our culture."

In part, grade placement of curricular material is a matter of providing experiences at each grade level which are suitable to the maturity level of the students and are designed to achieve the objectives of the program.\(^1\)

Kingsley defines maturation as "...the normal physical growth of the physiological functions. If these physiological structures have not developed to the point where the child can carry on the activity essential for a particular kind of learning, it is quite obvious he will be unable to achieve success in this direction."\(^2\)

In order to obtain the maximum efficiency in learning, maturation of the child must be considered carefully. Hilbreth\(^3\) points out that if a child is presented with a problem which is beyond his maturity level, he will reduce or simplify the problem to his own realm of understanding which may lead to misconceptions and make learning more difficult when the proper maturation level is reached.

\(^1\) Ibid., p. 625


\(^3\) Gertrude Hilbreth, "The Difficulty Reduction Tendency in Perception and Problem Solving", *The Journal of Educational Psychology* (April, 1941), 32:305-313.
Washburne\(^1\) points out that if a child is presented with a problem above his maturity level with the implication that he should succeed, it will give him a feeling of failure and undermine his security. "Instead we must guide him into those learning situations that he can attack effectively and with sufficient success to yield satisfaction, encouragement and growth."\(^2\)

Many of the studies that have attempted to assign learning experiences to definite maturity levels have been concerned with motor-skill development in pre-school children. \(^3\) But a number of studies have been made to determine the role that maturation plays in the development of various concepts, among them that of Pistor\(^4\) who conducted an experiment to determine how time concepts are acquired by children. Two groups of 320 children were used in the study. In grades four and five, traditional separate courses in geography and history were taught to one group while the second group was taught geography as a major


\(^2\) Ibid., p. 3.


course and history incidentally. In the sixth grade the first group was given instruction with special attention placed on time charts, time lines and other teaching aids. The other group had regular instruction with no special importance placed on time concepts. Through analysis of test results at the completion of the sixth grade, it was found that the group with special instruction gained slightly, but not significantly, in time-concept understanding, over the group without special instruction. Pistor concludes that "...evidence points heavily in favor of maturation rather than training as the dominating factor in time-concept development."¹/

Piaget²/ attempted to assign stages in the child's thought development to maturity levels. Through personal interviews, questions were asked relating to the child's ideas of the causes of natural phenomena. The responses were then placed in categories developed by Piaget. For example, Piaget traced three steps in concept development relating to the origin of the sun and the moon. The first step was that of artificiality of that the sun and moon were made by some human being. The second step was a belief that the sun

¹/ Frederick Pistor, *op. cit.*, p. 111.

and moon were developed by a combination of artificial and natural causes. The third stage in this development was the belief in a completely natural origin of the sun and the moon. The implication was that the child passed from one stage to another only when he had reached the proper maturity level. Due partly to the complexity and the subjective nature of interpreting the responses of the children, Piaget's method has been criticized and his conclusions challenged.

Deutsche conducted a study at the University of Minnesota in another attempt to trace the development of concepts of causal relations in children. Identical demonstrations and tests were given to children in grades three through eight. Three experts familiar with Piaget's work attempted to classify the answers to the test items into Piaget's categories. There was little agreement among the jury as to where each response should be placed. It was found that there was a great deal of overlapping, that most kinds of answers were found over the entire age group and that the answers of children of a given age group could not be classified into a single type. Deutsche concluded that "Causal thinking apparently does not develop by stages

but by a gradual process.\textsuperscript{1} She also found that the adequacy of the answers to the test questions increased with age, and the greatest increase noted was between the ages of 11 and 12 years.\textsuperscript{2}

Haupt\textsuperscript{3} sought to gather evidence to find out if young children were capable of the mental activities associated with the "large generalization" type aim. His study was limited to grades one through six. Haupt found that the ability to generalize prevailed at all grade levels, but that this ability was limited by the complexity of the concepts studied.

Croxton's\textsuperscript{4} study also indicates that children in the higher primary, the intermediate and the junior high school are capable of generalizing.

However, grade placement of curricular material is not entirely a matter of maturation. It is a matter of learning readiness which includes maturation, experiential levels, interest and attitudes, social pressures and training. But

\textsuperscript{1} Jean Deutsche, \textit{op. cit.}, p. 93.
\textsuperscript{2} Ibid., p. 29-42.
\textsuperscript{4} W. C. Croxton, "Pupils Ability to Generalize", \textit{School Science and Mathematics} (Jan. 1936), 36:627-634.
these factors are extremely difficult to separate for study. In the human body every organ is an integrated part of the whole body. If one organ is malfunctioning, it will affect the normal activity of the whole organism. Similarly, the child is a composite of many factors, each affecting the functioning of the other.

All experiences, according to Dewey,¹/ both take up something from those which have gone before and modify in some way the quality of those which come after. West²/ says that it is useless to show that a given volume of warm air is lighter than the same volume of cold air before the concept that air is something that has weight and occupies space is understood. So while the pupil may be at the maturity level for understanding a certain concept, if the necessary background is lacking, he will not learn effectively. Even if the maturity and experiential levels are adequate for learning, lack of interest or proper attitudes, inadequacy of teaching method and materials may account for unprofitable learning.


²/ Joe Young West, "Do We Expect Too Much or Too Little of Children from Their Experiences in Science?" Science Education (Oct. 1944), 33:298.
This study is designed to establish a learning index of the various scientific principles. This learning index will indicate the approximate mental age level at which these principles can be taught effectively to children of similar ability and background. Because of the complexity of the learning process, absolute values as to where each principle should be included in the curriculum is not expected, but the results may prove helpful to curriculum planners in determining the grade placement of these principles. The results of this study will be of importance to classroom teachers, textbook writers, standardized-test makers, and producers of visual-aids. It may, moreover, help bring about an orderly, systematic teaching of science, resulting in more and efficient learning.
3. Design of the Experiment

A. Scope and Limitations

The objective of this study is to establish a learning index for a number of scientific principles both in the elementary and secondary grades, the total study being made over a period of approximately ten years. An index of learning is to be assigned to each level at which the experiment is carried out.

A beginning has been made here by twenty-eight students working with different principles. Subsequent investigators using identical techniques with the same or other principles at different age levels may, after having secured data on a large number of pupils, predict with some accuracy where a certain principle might be taught with knowledge of its being understandable to the majority of pupils at that age level.

The procedure to be described is essentially the same in all the studies made by these twenty-eight investigators. However, since certain of these persons could not, of necessity, meet all of the conditions here set down because of their own teaching duties, the procedure has had to be slightly varied in such cases. Whenever any changes have been expedient, it will have been noted in subsequent chapters.
The population used in the study made by this first group of investigators is composed of pupils from the third to the twelfth grades. They are a stratified sampling of the school population of several New England states.

Each pupil's mental age is known through the use of chronological age, as furnished either by the pupil himself or the teacher, and the I.Q. obtained from the results of the administration of the Otis Quick-Scoring intelligence test. This enabled the experimenter to establish the mental modal age for each grade division of pupils tested.

The pupils whose test scores are included in the study all have mental ages within the limits of one year from the highest to the lowest. Once the modal mental age had been established, only the scores of those pupils with mental ages of plus or minus six months from the mode were selected to be included in the subsequent analysis.

Each investigator has examined two class divisions in five schools. Of the total of ten groups included, five are samples of the same grade level and the other five are samples from a different grade level which are separated from the first five samples by two years; that is, if a particular investigator chose five tenth-grade divisions, he will also have chosen either five eighth- or five twelfth-grade divisions.
Two examinations have been given to all pupils included in this study. The first will be known as the pre-test and the second as the post-test. They were identical. The post-test was given within an hour after the administration of the pre-test. The time lapse between the two was occupied either by the presentation of a demonstration serving to illustrate the particular principle being tested, in the case of what is known as the experimental group, or by reading non-relevant material by the group to be known as the control.

Strict discipline was maintained in each group in order that the pupils might not communicate with one another or be distracted from the examination or the demonstration given.

The demonstration was of large enough size, and was well lighted, so as to be seen by all the pupils in the classroom easily.

B. Discussion of Procedure

The first step involved the selection by each investigator of a principle to be demonstrated and tested. Such accepted lists as that compiled by Robertson were consulted.

A review of the literature established that the teaching of principles is an effective method for teaching science. It was found that facts were retained better when pupils were taught by principles. Also, relationships in applied learning were perceived more easily. Further findings on science teaching by principles are discussed in greater detail in a subsequent section.

The second step consisted in devising one or more demonstrations which illustrated the chosen principle. The time allotted for this teaching material was in most cases approximately 15 minutes. These demonstrations were necessarily simple, large, and contained as nearly as possible the "purity of concept" which has been interpreted by Nichols\(^1\) to mean that the demonstration illustrates one and only one principle. But if all other principles could not be eliminated, they were judged not to detract from the principle which the demonstration illustrated; they were judged not to lead to a misconception of the material taught.

The demonstration material of each experimenter was decided upon and the apparatus set up after having been presented to and passed on by a board consisting of a small group of investigators who, in turn, held their demonstration

material up to scrutiny by fellow board members.

In a subsequent section of this thesis are discussed the criteria for a good demonstration. The eleven pertinent points are summarized as follows:

1. The demonstration should illustrate a basic principle.
2. The demonstration should illustrate one principle only.
3. The action of the demonstration should be clearly visible to all.
4. The apparatus should be on a large scale.
5. The demonstration should work; it should be as infallible as possible.
6. The demonstration should be simple and the speed of action suitable.
7. The demonstration should be dynamic.
8. A slight dramatic element is sometimes useful.
9. An element of the unexpected is sometimes effective.
10. The apparatus should be of easily available and inexpensive material.
11. The apparatus used in a given demonstration should be stored away intact until it is to be used again.

All members of the group have adhered rigidly to these criteria.

It might be mentioned here that research, which will be
described in Chapter 2, on the idea that demonstrations are effective brought to light the fact that the demonstration is equal to or better than any other method of teaching science. Thus it is seen that if a particular scientific principle can be taught at a certain age level, the demonstration method is as good a way known to aid in the teaching of it.

After having perfected the demonstration a third step in the procedure was followed. Each investigator devised a test of the four-answer multiple choice type to be administered in not over 15 minutes time. This type consisted of approximately thirty items divided into three groups. The first ten items were based directly on the demonstration to be given. The second group consisted of items which involved transference; that is, these items did not test an understanding of the demonstration directly but tested the ability to apply the scientific principle involved to other simple nearby situations. The last ten items were more difficult; they involved an application of the principle but were of such a nature that correct answers might be made by the pupils who had gotten the most from the demonstration.

All of the items were so worded that the pupil could be given this test before the demonstration had been seen and yet answer the questions if he understood the principle.
For example, a question might be begun with a phrase such as "If a tight wire is plucked,......", etc.

In order to establish a suitable vocabulary for the items on the test, Thorndike's Teacher's Word Book was used. This volume lists words used most often in standard English reading material. Words used in the items were compared with the list to suit either the elementary or the secondary grades. If the particular words were not mentioned, others had to be substituted. The final form of the test contained a vocabulary which was suitable to the level at which each investigator was working. A copy of the writer's test is included in the appendix.

The test items were put in the interrogative form whenever practical with the answer to each consisting of one correct response and three distractors.

When the test was completed, it was presented to the same board which had previously judged the quality of the demonstration material. The items were passed if, in the opinion of the board, they were valid. An answer sheet for the test was devised whereby an enclosed space was left after the number of each item for the letter of choice.

The fourth step in the procedure involved the administration of the test to approximately one hundred pupils of the same grade level as the pupils for which it was finally

\[^{1/}\text{Edward L. Thorndike and Irving Lorge, The Teacher's Word Book of 30,000 Words, Bureau of Publications, Teacher's College, Columbia University, New York, 1944.}\]
intended. The results were incorporated in an item analysis which is described in a later section of this thesis. Any items which were shown not to be serving especially well were left on this final form of the test but only those items which were functioning well were used in subsequent compilations.

As the fifth step, a script to accompany the demonstration was written by each investigator using a suitable vocabulary selected from Thorndike's word list. \(^1\) This was not to be read to the experimental group while the demonstration was shown but served as a guide for the demonstration lecture, key points of the written procedure having been committed to memory by the experimenter. This minimized the probability of the individual lectures varying widely from day to day.

The investigator then presented his lecture and demonstration to a few pupils inviting comments after the presentation. In this manner both the script and demonstration were refined.

At this point in the procedure, the test and the demonstration were ready to be given. Each investigator had written to superintendents of schools, receiving permission to test pupils of two particular grades in each school.

\(^{1}\) Edward L. Thorndike and Irving Lorge, \textit{op. cit.}\)
Altogether five schools were selected and the pupils of two grade divisions in each school were chosen as subjects for the experiment. In some cases, investigators chose the elementary grades and in others, the secondary.

The sixth step involved the administration of the Otis Quick-Scoring intelligence test, by the investigator or the teacher of each particular division, during a period within two weeks of the demonstration.

The largest part of the experimental work is contained in the seventh step. On a prearranged date at a prescribed hour all the students of one class were pre-tested at the same time; that is, the examination was presented to the pupils before the principle was demonstrated. They were first given a test booklet and an answer sheet marked Test 1 on which there was a place for the filling in of the following information: name, sex, date of birth, name of school and town, and the previous training each pupil had in science. With regard to some of these items, in the lower elementary grades the information noted had to be checked and, many times, supplied by the teacher.

Each answer sheet contained a random number in the upper right hand corner and also a place for the investigator to later fill in any information he desired such as socio-economic background, I.Q., etc. A sample answer sheet is shown in the
appendix.

The time allotted for the pre-test was approximately fifteen minutes. At the end of this time the answer sheets were collected and half of the pupils in the class were sent to another room, after handing their test booklets to the demonstrator or the teacher in charge. The remaining half kept their booklets and stayed in the room to see the demonstration.

Half of the pupils were randomly selected according to a method used by Lindquist ¹. A table in his book was consulted and utilized. To explain the use of the table, it is perhaps expedient to use a hypothetical class in a single run of the experiment. Since there are 36 pupils in this class and half are to be selected at random, 18 pupils must be chosen arbitrarily. The first step is to assign numbers from 00 to 35 to the 36 answer sheets. This may be done in any order. Then it is necessary to select a starting point on the table by referring to a column and row number. As Lindquist ² states,

"This starting point should be determined before looking at any number in the table. Once having selected the starting point and direction, no peculiarity in the numbers read should be permitted to cause one to disregard the results and start anew at another point."

¹/ Everett F. Lindquist, Statistical Analysis in Educational Research, Houghton Mifflin Company, Boston, 1940, table 18, p. 262

From the starting point and reading in the chosen direction, the first 18 unlike numbers below 36 are taken and the pupils previously assigned these numbers are then one of the halves of the class.

After the class was divided, the answer sheets for the pre-test were collected and half the class was removed, as stated above. This half was designated as the control group. They spent the next 15 minutes reading silently some non-science material in another room.

Up to this time, the demonstration apparatus, which had been previously placed in the room where the pre-test had been given, was kept covered with a cloth. With only half the original group present, these demonstration materials were uncovered and the investigator began his experiments with his accompanying remarks of explanation.

When the demonstration, having lasted approximately 15 minutes, was over, a post-test answer sheet marked Test 2 was distributed to each pupil. This sheet was the same as that for the pre-test with the exception of the identifying test number. The original closed test booklet, which was to be used for the post-test, had been placed in the upper right hand corner of each pupil's desk. The group had previously been cautioned by the demonstrator not to open the booklets or talk among themselves while he was experimenting.
The post-test, being identical to the pre-test since the test booklet contained only the one test, was then administered to this experimental group. At the end of the allotted 15 minutes, the booklets and both answer sheets were collected.

Meanwhile in the room to which the other half of the class, the control group, had moved, the same post-test was given as was administered to the experimental group by a teacher who also supplied the pupils with an answer sheet marked Test 2 and a test booklet. After about 15 minutes had elapsed, the papers and booklets were collected.

In the cases where the investigator was working with elementary grade school pupils, the demonstration was given to the control group after they had taken the post-test because of the interest they undoubtedly had, because of administrative reasons, and, more important, because the time element was not such an important factor as it was in the secondary school where the control group was not given the demonstration.

This same procedure was repeated with individual divisions in each school until, as mentioned above, data on a total of ten divisions in five schools was collected.

It has been found that a reliable method of measuring the amount of learning of some specific activity, is by means of the test-retest method. By using the test-retest
method, the level of previous knowledge concerning the activity may be established. Using this information any gain in knowledge can be easily established. A detailed section on the test technique will be found in the next chapter.

The eighth step in the experimental procedure involved the compilation of statistics using the scores on both the pre-test and the post-test, the group modal mental age which had been computed from the I.Q. and the chronological age of each pupil. As was stated above only the scores of those pupils with mental ages of plus or minus six months from the mode were included in the statistical analysis. If a pupil of the experimental group showed a lack of understanding of the questions relating directly to the demonstration on his post-test, his scores were excluded from the analysis. A score which was less than 80 per cent correct on this part of the post-test was not used. The second chapter of this thesis contains a detailed explanation of how the scores were handled statistically.

The ninth and final step of the procedure was the making of a tape recording using the previously refined script for the demonstration lecture. When this had been done and the resulting recording found satisfactory, it was packaged along with the test booklets, sample answer sheets, and the
demonstration material. In this way, all necessary information and equipment will be ready for future investigators using the same principles.
CHAPTER II
SURVEY OF RELATED LITERATURE

1. Teaching by the Use of Principles

The teaching of science by principle rather than by extraneous collections of facts has been generally accepted by educators. The Thirty-first Yearbook of the National Society for the Study of Education, says that life enrichment, the aim of education, can best be achieved if the schools activities are "of the kind from which ideas may be developed and if the ideas may in turn be associated into principles and generalizations that are interwoven into human experience. Functional learning is conditioned upon attainment of some such integration." ¹

Hoban says: "Education is not simply the accretion of information. It involves the fundamental knowledge and the understanding of the basic principles of the universe, of which man is a part." ²

The inductive method.--Here the learner arrives at a general conclusion, e.g. certain laws of physical sciences,

by examining a number of individual cases. The weakness in this method is that there is a possibility of too general a conclusion, as the enumeration of particulars can never be totaled. For example after several enumerations of plants having flowers such as, the cactus has a flower; the buckwheat has a flower; the stringbean has a flower; we might conclude all plants have a flower. This is too general a conclusion as there are active fungi which do not possess flowers. Induction is thus essentially imperfect as a mode of reasoning, though invaluable as a means of fixing general principles and laws amid the succession of particularities given in experience.¹

The deductive method.--The learner reasons from a principle to a particular. It is in this method that we shall be mainly interested, for we are basing our whole experiment on the reasoning powers of the learners to go from the principle to a particular inference to the principles in their learning process. For example: If the learner understands the principle of friction he can deduce that heat is released and wear between the surfaces takes place when one body is rubbed over another.

A large amount of our teaching attempts to pupils to see

the implication of the laws, principles and rules that they may have learned. As contrasted with induction, deduction is a much simpler and shorter process. It is an unusual situation when a bit of deductive teaching lasts longer than a few minutes.

Advantages of deductive educative teaching:

1. Much more simple than the inductive method
2. Results in very desirable outcomes
3. Introduces factors of organization
4. Makes meaningful the principles that have been mastered already
5. Arouses puzzle or questioning instinct, a very valuable aid
6. Helps pupils to derive their principles from books or demonstrative techniques.

Jones, Leonelli, Martin and others have emphasized the value of teaching science by principle, and have listed hundreds of principles.

However, there is some disagreement as to what constitutes a principle. Heinmann defines a principle as "a statement of relationship between two or more facts."


Wilbur's definition as stated by Martin is much more precise and makes a principle a very specific kind of generalization. His criteria state that a principle——

"Is stated positively and definitely
Is true but with rare exceptions within the limitations set up by the statement
Clearly states or implies a dynamic process or interaction
Is demonstratable experimentally
Is clearly not a part of a larger principle which can be clearly stated
Is not merely a definition or description
Has wide application in the natural environment and is not ruled out by any of the preceding criteria."

Robertson's definition of a principle was the result of many weeks of consideration by a seminar in science teaching under F. D. Curtis at the University of Michigan:

"a. To be a principle a statement must be a comprehensive generalization
b. It must be true without exception within limitations specifically stated
c. It must be a clear statement of a process or an interaction
d. It must be capable of illustration so as to gain conviction
e. It must not be a part of a larger principle
f. It must not be a definition
g. It must not deal with a specific substance"

With this definition, Robertson sought to determine a comprehensive list of principles suitable as goals of instruction for elementary schools. He evaluated nine separate studies


listing principles found in textbooks, arranged according to frequency and stress, by a jury of three science teachers and several subject matter specialists. A list of the 243 principles found was sent to fifteen elementary school science teachers and from their ratings 113 principles were chosen. These are the principles used in the present study.

Some results.—There is considerable evidence that scientific principles can be taught effectively to students at the secondary level. Freud and Cheronis readministered a comprehensive test to students of a survey course in physical science one year after the course had been completed. They found that principles and the ability to apply such principles were retained much better than were unrelated facts.¹

Babitz and Keyes paired eight classes in chemistry in two California High Schools. Four of the classes, designated as the control groups, received standard instruction; the other four designated as the experimental groups, had direct and intensive training on the application of principles. The tests administered at the end of the experiment required the solution of problems in chemistry and the identification of scientific principles related these two. All the experimental groups showed superiority over the control groups in

¹/ Henrietta Z. Freud, and Cheronis, N. D., "Retention in the Physical Science Survey Course", Chemical Education Journal (June, 1940), 18:288-293
the same schools. The differences however were not statistically significant.\textsuperscript{1} Kilgore paired 120 students in high school physics with respect to their previous experience in science courses studied and I.Q. He found at the end of his study that students of both high and low ability were significantly better in making applications of principles of physics when the instructor placed emphasis on such application.\textsuperscript{2}

The evidence from these studies seems to indicate that the learning of principles of science, and the ability to apply them, may be attainable objectives of the teaching of science at the secondary level provided such objectives are emphasized in instruction.

2. The Lecture-Demonstration Method of Teaching

A. The Effectiveness of Lecture-Demonstrations

The areas which will be treated in this section are to define and describe the term lecture-demonstration, and then to quote freely the written opinions of science

\textsuperscript{1} Babitz and Keyes, "An Experiment in Teaching Pupils to Apply Scientific Principles", Science Education (December, 1939), 23:367-370.

\textsuperscript{2} W. A. Kilgore, "Identification of Ability of Apply Principles of Physics", Teacher's College Contribution to Education, No. 340, Columbia University, New York, 1941, p. 34.
educators with regard to the use of demonstrations in science teaching, describing the psychological and logical basis for the use of demonstrations in teaching. Then, a review of the research in which the lecture-demonstration is compared with other methods of science teaching will be presented.

Before discussing desirable qualities in a demonstration, Mack, in describing and defining a demonstration, says in part:

"Inherent in the concept of demonstration is the factor of movement of a material thing, not a static condition or display. A demonstration is an appeal through the senses of sight and of hearing, and less frequently through the other senses. Results must follow the purpose; there must be conviction, compelling to an inescapable conclusion."

Regarding lecture-demonstrations, Stuit and Englehart express their definition by stating:

"The term lecture-demonstration is used to describe a method of teaching in which the teacher carries out a demonstration for the entire group and lectures in parallel with it. The students observe the demonstration and ask any questions which they desire about the demonstration or theory involved."


Any discussion of the use of the demonstration in science teaching should be related to certain principles of learning. Potthoff\(^1\) has expressed awareness of such a relationship in the following writing:

"The use of the concrete, particularly where it deals with the unfamiliar, can provide an experimental basis for learning, whether that learning be remembering facts, understanding processes, seeing relationships, or getting an idea of how motor skills are executed. Direct experience, especially if it is with the unfamiliar, may motivate the learner, attract his attention, stimulate his interest, and arouse his curiosity. Demonstrations can be helpful also in facilitating comprehension of the abstract, giving reality to the spoken word, and reinforcing it by providing impressions through several sense avenues. In general, learning may be more meaningful, more accurate, more complete, and more permanent if it is based upon actual experience with that which is being studied."

Additional emphasis on the importance of the real or direct experience in learning has been made by Richardson and Cahoon in Methods and Materials for Teaching General and Physical Science\(^2\). They stated that:

"Probably the most usual use of the demonstration is for illustrating and explaining scientific principles and their applications. For most students seeing the real thing is much more helpful than reading about it or looking at a picture of it."

---


Whether the demonstration precedes or follows activities such as discussion, reading, films, and laboratory work, it may not automatically provide an understanding; but it furnishes a real experience upon which the teacher may build, along with other well-chosen procedures and activities.

Demonstrations can be used for providing pupil experiences in thinking. Cahoon's views on this topic are, in part, these:

"The demonstrations, laboratory experiments, directed studies, pupil projects, motion pictures, textbook statements, and pupil-teacher discussions are teeming with possibilities for pupil experience in thinking. It is largely a matter of utilizing these appropriately as one goes about teaching science facts and principles to pupils.

Like any other teaching aid or pupil activity, a particular exercise or experience in thinking may or may not be appropriate to use with a particular class at a given time.

A certain demonstration for one class may be given to help obtain a particular fact of science, at another time as an experience in accurate observation, at another to utilize previous knowledge by predicting 'what will happen', at another as an application of a recently studied principle."

However, it must be added here, that "A demonstration performed by a teacher who points out what is happening and indicates the conclusion which should be drawn or how it illustrates a particular principle may furnish little experience..."

in thinking." However, "thinking" comes when the principle is applied.

Before presenting a review of the research in which the lecture-demonstration is compared with other methods of science teaching, the evolution of the popularity of the demonstration method should be mentioned. Webb states that it developed in this manner:

"The growth of the demonstration method as a substitute for the individual experiment was accelerated by the depression, during which time funds for operating the schools were much reduced. It was argued that if the course in science be given by demonstrations, only one set of apparatus need be procured; whereas if it were given by student individual or group experiments a considerable number of duplicate sets must be purchased."

Cunningham's summary of "Lecture Demonstrations Versus Individual Laboratory Method in Science Teaching" covers a twenty-five year period. The field of research includes eighteen Master's Theses, six Doctorate Studies, and other studies. All of the reports were published in such professional periodicals as: Journal of Educational Psychology, School Science and Mathematics, School Review, Journal of Educational Research, and Pennsylvania School Journal. From the results reported by the experimenters, Cunningham


4/ Ibid., p. 76.
states that:

"Twenty-eight studies gave specific attention to the general outcome - immediate recall or immediate results. Twenty gave results favoring the demonstration method; six favored the individual laboratory method; and two said that there was no difference between the two methods."

Of the twenty-four studies that gave specific attention to delayed results, then favored the demonstration method, eleven the individual laboratory method, and three reported no difference.

The interest stimulated in the pupils by the two methods was studied in seven of the enterprises. The majority of the pupils in three of the enterprises favored the demonstration method; and in four of the enterprises favored the individual laboratory method.

All of the studies - fifteen - that gave attention to the time required by each of the two methods reported a saving of time under the demonstration method. The time saved varied from one-fifth to one-half."

Later in the summary, Cunningham tells of the treatment of scientific thinking in these studies by these comments:

"Seventeen studies gave attention to one or more of the elements of scientific thinking but no one undertaking 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 undertakings were as follows: amount retained in thought work; making proper conclusions to an experiment; application of principles learned; ability to think in terms of science subject; ability to follow the steps in scientific procedure; per cent of thought questions answered correctly; method of attack on new problems; scientific attitude; ability to

\[1\] Op. cit., p. 76
observe; learning a scientific principle; greater carry-over ability; ability to distinguish between fact and superstition; and ability to generalize.

Of the seventeen studies that gave attention to some phases of this big and very important problem, twelve favored the demonstration method; four the individual laboratory method; and one came to the conclusion that the pupil could learn to think about equally well by either method."

This comprehensive statement is part of the concluding remarks made by Cunningham:

"Our decision, as to what to do in practice, is made easier when we realize that all of our laboratory teaching need not - should not be done by one method. It is possible that we may be ignoring a whole continuous series of possibilities between these two extremes. In many cases it may be found best to use both methods in teaching a given idea in science."

The studies presented in the summary of Cunningham were ranked according to the criteria presented in an article by Stuit and Englehart by Keiser as to their superior or inferior value. It is well to note here that Keiser used only the first six of the seven criteria to determine the value of these studies. The seven criteria, as established by Stuit and Englehart, are as follows: (1) specification of experimental factors; (2) control of

1/ Ibid., p. 79
pupil factors; (3) control of teacher factors; (4) control of general school factors; (5) duration of experiment; (6) measurement of achievement; and (7) interpretation of experimental data.

For comparative purposes the writer has used the studies of Anibel¹, Knox², and Wiley³ in this discussion because each study is partly concerned with the demonstration method versus the laboratory method of teaching high school chemistry. The problem of the research as stated by each author and the significant conclusions, in part, will be related.

The study of Fred G. Anibel⁴, ranked superior, is as follows:

**Problem:** To determine scientifically through objective data how the results of teaching high-school chemistry by lecture-demonstration method compared with the individual laboratory method.

**Conclusions, in part:**

1. The immediate retention is as adequate when material is presented by the lecture-demonstration method as when the class is taught by


⁴ Loc. cit.
the regular individual laboratory procedure. Indications are that the lecture-demonstration procedure would result in better immediate retention.

2. The delayed retention is so little different that one method may be considered as good as the other. There was a slight indication that the material was better remembered when taught by the individual laboratory procedure.

3. The brighter students are likely to profit more by the lecture-demonstration method than are the others.

The study by W. W. Knox, which was ranked superior, is as follows:

Problem: To establish the relative value of the demonstration and laboratory methods of science instruction.

Conclusions, in part:

1. The demonstration method is superior to the laboratory method in teaching mentally heterogenous groups of pupils for the purpose of immediate retention and relatively permanent retention of subject matter in high school chemistry.

2. For the purpose of imparting to a group of pupils a scientific attitude and training in a method of attack on new problems, the demonstration method is equal, if not superior, to the laboratory method of instruction.

3. From the standpoint of the coefficients of correlation, it appears that the demonstration method provides superior opportunity for adaptation to individual differences in mental ability so far as teaching for immediate retention, delayed retention, and method of attack are concerned.

4. So far as providing knowledge and method of attack are concerned, the laboratory method is slightly superior to the demonstration method in the case of the average inferior pupil.

1/ Loc. cit.
5. For the purpose of providing knowledge for both immediate retention and relatively permanent retention, and for the purpose of providing a technique for handling new problems, the demonstration method is much to be preferred to the laboratory method in the case of the average superior pupil.

Before stating the problem and significant conclusions of a study ranked inferior by Keiser\(^1\) according to the first six of the seven criteria developed by Stuit and Englehart\(^2\), it should be recognized that the study made by Wiley\(^3\) was a pioneer enterprise, being published in 1918. Probably it has been ranked as of inferior value because of the following factors: no mention is made of any attempt to measure the mental abilities of the pupils\(^4\); the tests to measure immediate and delayed retention were of doubtful validity; the method of scoring the tests was highly subjective; and there was no mention made of statistical treatment of the data found.

The study made by William H. Wiley\(^5\) is as follows:

Problem: To determine the best of the three methods of teaching chemistry, the textbook recitation method, the so-called lecture demonstration method, and the laboratory method.

Conclusions, in part:
1. There is not as great a difference

\(^1\) Loc. cit.
\(^3\) Loc. cit.
\(^4\) Note the date of publication of the first group intelligence tests.
\(^5\) Loc. cit.
as is ordinarily supposed in the value of the three methods, lecture, demonstration, textbook, and laboratory, so far as imparting knowledge is concerned.

2. For immediate learning the textbook method is unquestionably superior.

3. For permanent learning the laboratory method is perhaps slightly superior.

4. In every respect the lecture demonstration method is the least effective in imparting knowledge to high school students.

5. The rate of forgetting is greatest with the textbook method and least with the laboratory method.

6. The different methods show decided individual differences both for immediate and delayed reproduction.

7. Probably a combination of the three methods will give the best results in teaching high school chemistry.

Stuit and Englehart have also made an excellent critical analysis of the lecture demonstration versus the individual laboratory method of teaching high school chemistry.

A summary of their report, which consisted of the combined conclusions of various investigators, is as follows:

Conclusions contending that the laboratory method is superior:

1. There is a slight indication that material was better retained when taught by the individual laboratory method - Anibel.

2. The order of preference of the methods studied places the individual laboratory method before the demonstration method - Horton.

3. In every respect the lecture method is least effective in imparting knowledge to high school students - Wiley.

4. For permanent learning the laboratory method is perhaps slightly superior - Wiley.

5. For providing knowledge and method of

attack, the laboratory method is superior for the inferior pupil - Knox.

Conclusions claiming that the demonstration method is superior:
1. Bright pupils are more likely to profit by the lecture-demonstration method than are the others - Anibel.
2. Dull pupils profit more from demonstration than from individual laboratory work - Carpenter.
3. The lecture-demonstration takes less time and costs less - Anibel.
4. The teacher (demonstration) method is best - Nash and Phillips.
5. Lecture-demonstration method gives better control over the individual since all are under teacher guidance - Pugh.
6. For purpose of providing knowledge for both immediate and permanent retention and for the purpose of providing technique or handling new problems, the demonstration method is much to be preferred to the laboratory method in case of average superior pupil - Knox.

Conclusions contending that the students achieved equally well by either method:
1. Immediate retention is about equal in both lecture-demonstration and individual-laboratory methods - Anibel.
2. There is not as great a difference as is ordinarily supposed in the value of the three methods, lecture, textbook, and laboratory, so far as imparting knowledge is concerned - Wiley.
3. The results of this experiment point to the conclusion that the majority of students in high-school, laboratory-chemistry classes, taught by the demonstration method, succeed as well as when they perform the experiment individually, if success is measured by instruments which measure the same abilities as are measured by these tests, namely, specific information and ability to think in terms of chemistry - Carpenter.

General conclusions based on evaluation of the reported research:
After considering the above conclusions the writers have arrived at a few ideas which seem justifiable in the light of the evidence
given by this study.

1. No method can be considered to be the best in every case. The objectives of chemistry teaching, the preference of the teacher, the nature of the pupil, and the facilities of the schools will largely determine which method should be used.

2. In small schools where money and space are not plentiful the lecture-demonstration method seems to be most practicable.

3. The written test cannot be used to test all the outcomes of a course in high school chemistry. Some sort of manipulative tests seem necessary to test the laboratory skills.

4. The problem of the relative merits of the lecture-demonstration and individual-laboratory methods still seems unsolved and as complex as ever. More careful experimentation, involving careful control of non-experimental factors and reliable testing, is needed in order to justify any definite and final conclusions. When experimentation has shown the relative superiorities of the methods in terms of outcomes, the methods should be evaluated in terms of the values attached to these outcomes.

Evidence of the evolution of teaching methods particularly by the visual method is apparent in the next study to be considered, that of Smith. The visual method has been long recognized by leading educators as one of the most valuable ways of training pupils in all stages of learning. It is difficult, as a result, to find a school that does not, in one way or another, make use of visual aids in teaching. The alert teacher and administrator are constantly seeking suggestions and illustrations by means of which the vague conceptions of the pupils may be made into real facts and parts of their experience.

---

1/ Herbert A. Smith, "A Determination of the Relative Effectiveness of Sound Motion Pictures and Equivalent Teacher Demonstration in Ninth Grade General Science", Science Education (April, 1949), 33:214-221.
Any method will not be overlooked if it can provide both clarity and simplicity combined.¹/

In the study done by Smith²/ the problem involved was:

1. What is the relative effectiveness in ninth grade general science classes of experimental demonstrations performed by the teacher and equivalent demonstrations presented through the medium of educational sound motion pictures.
2. The determination of the relative effectiveness of these two instructional techniques with pupils of different levels of intelligence.

In the plan of study three methods of presentation were used: (1) teacher demonstration, (2) use of films, and (3) a combination of teacher demonstrations and the use of films. The conclusions made, as a result of this study, were as follows:

1. Educational sound motion pictures and teacher demonstration are of equal merit as instructive devices in ninth grade general science when they include essentially the same materials in so far as merit can be determined by the techniques employed in this investigation. The use of either method singly is as effective as the combination of the two.
2. There is a tendency for increased intelligence as expressed in terms of an intelligence quotient to be accompanied by increased learning where learning is represented by the gain of final over initial test scores on the objective tests used in this investigation. The degree of relationship is independent of the method of instruction

utilized indicating the same relative value for sound motion pictures and teacher demonstrations portraying essentially identical materials regardless of the level of intelligence of the students.

Certain similarities exist between the problem of this science seminar and the problem investigated in the study by Croxton.\(^1\) His problem was stated thusly:

"Is the failure of children to generalize due primarily to lack of power or tendency, or is it simply due to want of sufficient experience...."

In this study most of the experiments tested the pupils' ability to formulate and apply a principle after eight minutes exposure to the essential experimental basis in the form of a demonstration or directed play. The tentative conclusions made by Croxton\(^2\) are as follows:

1. The data indicates that many children in the higher primary, the intermediate, and the junior-high school grades are capable of generalizing.
2. While the experiments do not prove that most pupils in the kindergarten and lower primary grades could not generalize if a more adequate experience basis was provided, the data together with the evident obsession manifested by these children for obtaining emotional satisfaction do suggest that early childhood is preeminently a period for satisfying reactions.
3. There is little in these experiments to suggest that junior-high school pupils possess markedly superior ability to generalize than intermediate grade pupils possess, the difference in the scores in favor of the former being

\(^1\) W. C. Croxton, "Pupils' Ability to Generalize", School Science and Mathematics (June, 1936), 36:627-634.

\(^2\) Ibid., p. 634.
little more than might reasonably be credited to added experience.

In summary, therefore, of the research and studies compiled by investigators on the value and the effectiveness of the lecture-demonstration, as compared to other methods of science teaching such as the individual laboratory method, the textbook method, and sound motion pictures, it can be concluded that the lecture-demonstration method of science teaching is equal to, if not better than, any other method of teaching. It incorporates direct experience of the pupil, pupil experience in thinking, utilization of the senses, understanding processes, application of scientific principles, and ability of the pupil to generalize. Through the use of the lecture-demonstration method most, if not all, of these above qualities are satisfied. Again, the writer would like to cite the fact that the lecture-demonstration method is equal to, if not better than, any other method of teaching science.

2. Criteria for a Good Demonstration

Statement of the problem. -- There are two problems involved in developing a list of criteria for a good demonstration: (1) to define clearly the word "demonstration" as it is to be used in this experiment; (2) to evolve, through reference
to the literature, the criteria.

Need for research.--Since the demonstration is the instructional procedure selected for use in the experiment, it is necessary to clarify the meaning of the demonstration method.

Noll\(^1\) has pointed out the fact that investigators seldom define teaching methods carefully and minutely enough. Various writers have recognized the need in research for accurate definition of terms. Like Noll, Riedel\(^2\) has made a plea for clear definition of teaching methods and experimental procedures. Mack\(^3\) has stated that there are as many definitions of "demonstration" as there are authors treating the subject. Preston\(^4\) also realized this and called for clarification and unification of terminology.

Most of the literature on the demonstration method fails to recognize the difference between the lecture-demonstration, the class experiment, and the illustrated lecture.


\(^4\) Carleton E. Preston, "Is the Debate in Common Terms?" Science Education (February, 1935), 19:14-16.
Preston attributes much of the success of lecture-demonstrations to their actually being class experiments.

**Definition of demonstration.---**The demonstration is in this experiment actually a lecture-demonstration. Reference to the literature will help to clarify the meaning. First, the "demonstration" is defined by the Dictionary of Education as follows:

"(1) The method or process of presenting or establishing facts; (2) the procedure of doing something in the presence of others either for means of showing them how to do it themselves or in order to teach a principle."

The same source defines the lecture-demonstration thus:

"An instructional procedure in which the verbal message is accompanied by use of apparatus to illustrate principles, determine or verify facts, clarify different parts, or test for comprehension of material under discussion."

Preston further clarifies the concept of the lecture-demonstration as distinct from the class experiment:

"In true lecture-demonstration the teacher shows everything, explaining or interpreting each point as he, or some pupil, performs the work. In true class experimentation the teacher endeavors, by well-directed questions, to get the members of the class to observe or come to conclusions themselves as to the proper interpretation, and perhaps to plan further steps or procedures. Thus,

1/ Loc. cit.


3/ Ibid., p. 238

in the lecture-demonstration the flow of information and explanation is from teacher to pupils; in the class experiment it is exactly the opposite."

Elsewhere, in defining lecture-demonstration, Preston makes the point that "no questions interrupt the speaker and he asks his audience none, other than for rhetorical effect."

Preston, however, does not distinguish the lecture-demonstration from the illustrated lecture as does the Encyclopedia of Modern Education:

"The lecture-demonstration differs from the illustrated lecture in that the latter focuses attention on the screen and shows the relationships by means of pictures, slides, moving pictures or specimens while the lecture-demonstration focuses attention on the lecturer who shows the relationships through the use of manipulation of physical material, machines or appliances."

The meaning of "demonstration" is further expanded by the following observation made by Mack: "Inherent in the concept of demonstration is the factor of movement of material things, not a static condition or display." This so-called dynamic quality of the demonstration leads Mack to exclude from the demonstration procedure certain standard teaching materials:


4/ Loc. Cit.
"objects, unless they can be operated ... so also, specimens, samples and parts ... Likewise models, as such, are barred unless they are working models; so also, miniatures and enlargements."

Although micro-projection techniques are gaining increasing favor in demonstration work,\(^1\) it would seem that this method should also be excluded on the same basis as the other visual aids.

Further, Mack\(^2\) states that the demonstration is "an appeal through the senses of sight and hearing and less frequently through the other senses." He would, therefore, exclude from demonstration work materials that appeal to only one sense; such as, transparencies, pictures, charts, recordings and radio reproductions.

Thus certain characteristics of the demonstration have been determined by definition. These are:

1. The demonstration is an instructional procedure.
2. It is frequently used to teach principles.
3. It differs from the class experiment.
4. It differs from the illustrated lecture.
5. Movement and action are essential.
6. It is an appeal through two senses: sight and hearing.

The necessary implications of each of these statements have already been suggested.


\(^2\) Op. cit., p. 21
Review of the literature.--A review of the literature was made in order to discover those basic principles which might be used as a guide in doing demonstrations.

First, a search was made to locate any previous studies that paralleled this investigation. The Bibliographic Index provided the necessary references. It was found that many investigators had subjectively listed criteria in one form or another. However, only one study, documented with references, proved similar to this one. Mack\(^1\) covered many of the same sources in developing his checklist for evaluating desirable qualities of demonstration apparatus. He lists as "factors" those conditions inherent in the physical surroundings and in good techniques and as "qualities" those conditions inherent in the apparatus. Much of his research had to be duplicated in this review, but for a different purpose which called for more complete and descriptive statements.

A working bibliography was developed consisting of five types of sources: (1) professional journals and science publications, (2) methodology textbooks, (3) teaching science textbooks, (4) audio-visual texts and (5) books on experiments. The following reference sources were consulted: Bibliographic Index, Encyclopedia of Educational Research, Bibliographies and Summaries in

There was great variety in the nature of the material covered which included such items as:

1. Steps to follow
2. Desirable qualities
3. Desirable characteristics
4. Points to keep in mind
5. Rules for demonstrating
6. Suggestions for making demonstrations effective
7. Criteria
8. General discussions of the demonstration method

Works included in this study fall into four categories:
(1) Those which deal with the demonstration in a general sense; (2) those from the field of biology (3) those from the field of physics (4) those from the field of chemistry.

Several of the authors in the first category, the "general", emphasize only one or a few aspects of the use of demonstrations. In discussing the presentation of example demonstrations, Cahoon indicates certain steps taken to insure effectiveness of the demonstration and emphasizes only visibility and size of apparatus.

Colvin\textsuperscript{1} offers three cautions to be observed in class demonstrations. Hoff\textsuperscript{2} emphasizes only visibility and planning. Pinkus\textsuperscript{3} suggests the need for apparatus especially designed for demonstration purposes and stresses the factor of visibility.

A few in this same group attempt more detailed coverage. Potthoff\textsuperscript{4}, for example, offers several suggestions for performing demonstrations effectively and contributes many excellent ideas. In discussing the art of lecture table demonstration, Davison\textsuperscript{5} mentions several rules to follow in demonstrating. Rakestraw\textsuperscript{6} touches on six different aspects of the good demonstration in his extensive discussion of lecture-demonstration.

\textsuperscript{1} Stephen Sheldon Colvin, \textit{An Introduction to High School Teaching}, The Macmillan Company, New York, 1924, Ch. 12.


\textsuperscript{3} L. F. Pinkus, "Some Suggestions in Demonstrations", \textit{Science} (October 20, 1933), 78:364.


\textsuperscript{5} H. F. Davison, "The Art of Lecture Table Demonstration", \textit{Journal of Chemical Education} (June, 1927), 2:443-47.

Still others in the "general" group have systematically attempted to list criteria in some form. Billinger¹ lists five requirements for a successful demonstration. Dale² offers fourteen suggestions for improving demonstrations and eleven questions for evaluating them. Under "demonstration techniques", Haas³ lists ten steps to be completed before conducting the experiment and five suggestions for conducting it. Heiss⁴ elaborates on seven excellent rules for demonstrating. Holley⁵ lists seven things a teacher can do to insure successful demonstrations. Mack⁶ developed a lengthy checklist of desirable qualities in demonstration apparatus. In a group thesis edited by Murray⁷, five criteria for a demonstration were listed.


which had been developed in a seminar discussion. Richardson and Cahoon\(^1\) list five criteria for a good demonstration. Selberg\(^2\) lists sixteen common errors in demonstration techniques (actually class experiment techniques) and offers an excellent plan to follow in doing classroom demonstrations. In the second category, the works from the field of biology, only one study was found. Gramet\(^3\) lists eight characteristics of the good demonstration.

In the third category, works from the field of physics, the same breakdown can be made as for the first category. Among the few who emphasize only one aspect, Coyle\(^4\) stresses the value and importance of vertical mounting of apparatus on special boards. Also, Sutton\(^5\) stresses the need for simplicity and originality. Among his suggestions for improving physics teaching, Weaver\(^6\) stresses visibility and size of apparatus.


Hitchcock\textsuperscript{1/} emphasizes \textit{action} as the essential quality of good demonstrations and includes, as he elaborates this theme, many other criteria.

Duff\textsuperscript{2/} is the only one in the field of physics to make a systematic listing. He enumerates nine desirable qualities in demonstration experiments. The fourth and final category, the works from the field of chemistry, may be similarly analyzed. Arthur\textsuperscript{3/} presented a lengthy discussion on visibility including many excellent suggestions. Reed\textsuperscript{4/} discusses in some detail four aspects of good demonstrations and techniques. Wiles\textsuperscript{5/} also deals only with a few aspects of successful demonstrations.

Dunbar\textsuperscript{6/} lists eleven desirable characteristics in demonstrations. His list is based on Duff's and includes

\textsuperscript{1/} Richard C. Hitchcock, "I Like Action in Physics Demonstrations", \textit{School Science and Mathematics} (December, 1941), 41:832-839.


\textsuperscript{4/} Rufus D. Reed, "High School Chemistry Demonstrations", \textit{Journal of Chemical Education} (November, 1929), 16:1905-9.

\textsuperscript{5/} L. A. Wiles, "The Value of Lecture Table Demonstrations in the Teaching of Chemistry", \textit{Journal of Chemical Education} (September, 1928), 5:1109-1111.

specific examples in chemistry, Frank\(^1\) provides twelve suggestions regarding use of class demonstrations which he believes to be justified by the experience of a number of teachers. Gould\(^2\) enumerates on eight to consider in planning and performing demonstrations. Van Horne\(^3\) offers five suggestions for the preparation of apparatus and materials and four rules to follow in conducting demonstrations.

**Treatment of the data.**—From the literature selected for inclusion in the study, each separate statement which seemed to form the basis of a possible criterion was noted on an individual card. Items were accepted for consideration if they were mentioned once. They were rejected on the basis of (1) inconsistency with the definition of a demonstration developed, or (2) inconsistency with the design of the experiment. The items thus selected for inclusion in the list of criteria were organized into the outline which appears below. The criteria themselves are listed as major statements. Suggestions for implementing them are listed as sub-topics under the criterion to which they seem best to apply. This arrangement is entirely ar-

---


bitrarry. Many of the items organized as sub-topics were in some studies listed as a "major criterion". Various items included as sub-topics were mentioned by different authors in different places. However, an arbitrary organization seemed justified on two bases: (1) no two of the studies duplicate each other; and (2) the nature of the material covered is of such diverse nature. The criteria thus organized were submitted to the seminar for criticism and revision. This list was accepted as it appears below.

Selected criteria.--The criteria for a good demonstration as used in this experiment are as follows:

**CRITERIA FOR A GOOD DEMONSTRATION**

I. **THE DEMONSTRATION SHOULD ILLUSTRATE A BASIC PRINCIPLE.**

II. **THE DEMONSTRATION SHOULD ILLUSTRATE ONE PRINCIPLE ONLY.**

III. **THE ACTION OF THE DEMONSTRATION SHOULD BE CLEARLY VISIBLE AND AUDIBLE TO ALL.**

A. Remove all the audio-visual distractors.

B. Make sure the lighting facilities are adequate. Spotlight or otherwise sufficiently illuminate the thing being demonstrated.

C. Adjust window shades so that students can see from all parts of the room.

D. If necessary, rearrange the seating so that everyone has an unobstructed view.

Dunbar's list of desirable characteristics is based on the list developed by Duff.
E. Be sure that those with poor hearing and vision are seated appropriately.

F. Have the demonstration table arranged so that all pupils can see the demonstration.
   1) Vertical mounting of apparatus is especially effective.
   2) Place the apparatus well forward on the desk, facing out toward the pupils.
   3) Place demonstration table in best position for all to see from all angles.

G. Wherever possible, make use of color contrast to make the apparatus or materials stand out.

IV. THE APPARATUS SHOULD BE ON A LARGE SCALE.

A. The apparatus must be clearly visible from the furthest corner of the room.

B. Where a thermometer (or other meter) is essential to the demonstration, use a mock-up or working model to help the class visualize this part of the procedure.

C. Large signs and diagrams may be used to supplement the spoken word.
   1) They must be previously prepared.
   2) They must be clearly visible to all.
   3) Green print on yellow is preferable to black on white.
V. THE DEMONSTRATION SHOULD WORK: IT SHOULD BE AS INFALLIBLE AS POSSIBLE.

A. Apparatus should be in sound working condition.

B. Apparatus should be as simple as possible.
   1) Simplicity of operation.
   2) As few parts as possible.
   3) Avoid crowding, overlapping and masking of the parts.

C. The demonstration should be rehearsed in advance.

D. The demonstration should be well-planned and prepared.
   1) Set up apparatus and have all materials carefully arranged on the demonstration table before the class meets.
   2) All the necessary measuring and weighing should be done before class.
   3) Scales and graduates should be placed away from the demonstration table when no longer in use.

VI. THE DEMONSTRATION SHOULD BE SIMPLER AND THE SPEED OF ACTION SUITABLE.

A. Use simple setups and place the equipment in order on the table so that the action can proceed logically.

**--These might well be separate criteria.
B. Talk while you work. Be sure to:
   1) Emphasize the main points; do not digress.
   2) Keep summarizing as you go along.
   3) See to it that the demonstration moves on quickly to a conclusion; do not hurry or drag.

C. Use a simple vocabulary.

VII. THE DEMONSTRATION SHOULD BE DYNAMIC.
   A. By definition, movement and action are essential to the demonstration.
   B. Positive effects of motion are more impressive than null effects of static display.

VIII. A SLIGHT DRAMATIC ELEMENT IS SOMETIMES USEFUL.
IX. AN ELEMENT OF THE UNEXPECTED IS SOMETIMES EFFECTIVE.
X. THE APPARATUS SHOULD BE OF EASILY AVAILABLE AND INEXPENSIVE MATERIAL.
XI. THE APPARATUS USED IN THE GIVEN DEMONSTRATION SHOULD BE STORED AWAY INTACT UNTIL IT IS TO BE USED AGAIN.

The frequency with which the above mentioned criteria were mentioned by the sources consulted is indicated by the chart below. The count was made merely for general interest. It has, however, certain obvious values. The frequency of mention of the various criteria provides means of establishing their validity. The table shows the relative importance of the criteria as recognized by these authorities.
Table 1. The Frequency of Mention of the Selected Criteria by the Sources Consulted.

<table>
<thead>
<tr>
<th>Author-Source</th>
<th>Criterion Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1. Arthur.....</td>
<td>x x x</td>
</tr>
<tr>
<td>2. Billinger...</td>
<td>x</td>
</tr>
<tr>
<td>3. Cahoon......</td>
<td>x</td>
</tr>
<tr>
<td>4. Coyle......</td>
<td>x</td>
</tr>
<tr>
<td>5. Colvin......</td>
<td>x x x x x</td>
</tr>
<tr>
<td>6. Dale......</td>
<td>x</td>
</tr>
<tr>
<td>7. Davison.....</td>
<td>x</td>
</tr>
<tr>
<td>8. Duff........</td>
<td>x x x</td>
</tr>
<tr>
<td>9. Dunbar......</td>
<td>x</td>
</tr>
<tr>
<td>10. Frank.......</td>
<td>x x x x x</td>
</tr>
<tr>
<td>11. Gramet......</td>
<td>x</td>
</tr>
<tr>
<td>12. Gould.......</td>
<td>x</td>
</tr>
<tr>
<td>13. Haas.......</td>
<td>x</td>
</tr>
<tr>
<td>14. Heiss.......</td>
<td>x</td>
</tr>
<tr>
<td>15. Hitchcock...</td>
<td>x</td>
</tr>
<tr>
<td>16. Hoff.........</td>
<td>x</td>
</tr>
<tr>
<td>17. Holley......</td>
<td>x</td>
</tr>
<tr>
<td>18. Mack.........</td>
<td>x</td>
</tr>
<tr>
<td>19. Murray......</td>
<td>x</td>
</tr>
<tr>
<td>20. Pinkus.......</td>
<td>x x x x x</td>
</tr>
<tr>
<td>21. Potthoff....</td>
<td>x</td>
</tr>
<tr>
<td>22. Rakestraw...</td>
<td>x x x x x x x x x</td>
</tr>
<tr>
<td>23. Reed.........</td>
<td>x</td>
</tr>
<tr>
<td>24. Richardson...</td>
<td>x</td>
</tr>
<tr>
<td>25. Riedel......</td>
<td>x</td>
</tr>
<tr>
<td>26. Selberg......</td>
<td>x</td>
</tr>
<tr>
<td>27. Sutton.......</td>
<td>x</td>
</tr>
<tr>
<td>28. Van Horne...</td>
<td>x</td>
</tr>
<tr>
<td>29. Weaver.......</td>
<td>x</td>
</tr>
<tr>
<td>30. Wile s.......</td>
<td>x</td>
</tr>
</tbody>
</table>

Frequency....... 11 7 25 19 22 14 4 7 6 6 2 2

Criterion number 1 2 3 4 5 6 7 7 9 10 11
3. The Test Technique

A. Structure of the Test

Multiple choice items.--The test is composed of approximately thirty multiple choice items. Each item is in the form of an incomplete sentence, or a question, referred to as the stem, accompanied by three or more possible responses. Of the possible responses presented to the examinee, one is the best response. The examinee is also presented an answer sheet upon which he checks in the parenthesis the response he has selected. The construction of the response items, in accordance with suggestions made by Ross\(^1\), has been grammatically consistent, approximately of equal length, and plausible, in so far as possible. The writer has endeavored to make the type of responses for each item homogeneous in nature, in order to detect higher levels of understanding and discrimination.

This "best-answer" variety of the multiple choice test means simply that one response best refers to the stem of the item. Each item provides "a response that competent critics can agree upon as best."\(^2\) The competent critics in this specific situation compose a group of in-service


science teachers. The writer has taken great care to "make all distractors plausible and attractive to examinees who lack the information or ability tested by the item",¹ and "avoid highly technical distractors".² In reference to the multiple choice type test, Odell³ states that "they may be used to test not only knowledge of facts and amount of acquired information, but also knowledge of cause and effect relationships, ability to make comparisons, to evaluate, to apply, to illustrate, to define, and so forth. They are easier to prepare, and also to score, than some of the other types." He further adds⁴ "almost all kinds of multiple answer tests can be constructed so that they possess practically perfect objectivity." The scorer is not faced with the problem of partial credit on this type of an examination. Either the response that is checked upon the paper is correct, or it is not correct, with no qualifications.

Levels of difficulty.--The writer is making an attempt, to determine to what probable extent application and recognition, as well as understanding of a scientific principle

²/ Ibid., p. 235.
have been gained through the demonstration activity. For this reason, it is necessary for the examiner to approximate the difficulty range of the test items which he has prepared. It is well recognized that there are various levels of learning. In order to measure these levels of learning, a testing device of various levels of difficulty must be constructed. The actual judgment of item difficulty must be left up to the subjective judgment of the test constructor. "The use of subjective judgment in estimating item difficulty at the stage of item construction is to be encouraged. Such judgments, when based on all available experience, are distinctly helpful in leading to the construction of items of the desired difficulty." The constructor has ample opportunity to construct the items of various degrees of difficulty by using more remote subject matter applications, or by including unusually good distractors in the test items. Odell states that, in reference to good distractors, "their selection will depend to some extent upon how difficult it is desired to make the test. Incorrect answers should, however, never be obviously incorrect to a pupil who knows little or


3/ Ibid., p. 286.
nothing of the matter dealt with...."

The various levels of learning may be broken down to three broad categories. The first level of learning may be labelled, or described as mere factual retention. The second level employs enough understanding of the factual retention so that the learner can recognize and apply, in simple situations, the principles or concepts which he has retained. The third level of learning is reached when the learner can recognize and apply the understanding of the factual material to more complex, unfamiliar, and difficult situations. The test has been constructed with these three levels of learning in mind. The first third of the test is concerned with items of the first level of learning, and so on. Thus, the test can be said to measure three levels of learning, all concerned with the same demonstration, and the same scientific principle. This method of testing tells the examiner to approximately what extent the pupil can recall, understand, or apply the principle.

Vocabulary.--It is only logical for one to assume that the vocabulary used throughout the experiment must be consistent, or at least on the same level. Vocabulary comprising the test must, of necessity, be equivalent to that used during the demonstration. Inconsistent vocabulary is one of the factors which could unfavorably affect the reliability of
the testing program. If the vocabulary within the testing device is inconsistent with that of the oral demonstration, one can expect a low reliability of the whole testing procedure. Reliability, itself, is the consistency with which a test measures "what it measures".

The vocabulary of the testing device has been amended by the critic-jury to establish consistency of vocabulary throughout the experiment and vocabulary comprehension at the grade level at which the test is used.

The test tryout.---"After a set of test items has been written, criticized by subject matter experts, and revised on the basis of their criticisms, it must ordinarily be tried out experimentally on a sample of examinees." 1

Prior to any experimentation, the test was subjected to a tryout on at least one hundred pupils of equivalent age and grade level, but are not included in the experiment. This independent tryout tended to expose any unusually poor items, or poor distractors among the possible responses. Such items could be dropped completely from the test, or eliminated in the final tabulation of the total results.

As was stated previously, the total number of items in the test approximates thirty, but some may be dropped due to the discretions of the critic-jury, or as a result of the test tryout.

The test period.—The length of the testing period for both the pre-test and the post-test has been indefinite, in so far as no specific time limit has been set for either of the tests. The test period may continue on until every pupil has completed the test, in so far as possible. Each pupil is allotted sufficient time to at least read all of the items presented him. A multiple choice test of thirty items can be approximated as requiring about ten minutes to be read through completely. Odell 1/ has recommended that "on the average elementary-school pupils be expected to respond to three or four such exercises [multiple choice items] per minute."

By allotting sufficient time for all examinees to attempt all the items, the influential factor of time itself is eliminated. As stated by Lindquist, 2/ "The most common way of reducing or eliminating the influence of time on tests is to set the time limits so liberally that all, or nearly all, pupils are able to consider or attempt all the items in the test." Pupils are told to complete all items, and are watched to see that they keep at this task until finished.


B. Aims and Use of the Test

Employing statistics.--The test is an instrument devised to obtain statistics for measuring growth of learning, due to a specific educational experience, namely a scientific demonstration. Every effort has been made in the construction of the test to measure as precisely as possible, the "meaningful learning" that has been grasped by each pupil subjected to the demonstration and the test-retest procedure. The only descriptions of the learning and understanding that have taken place are the statistics which can be applied to the results of the tests taken by the examinees. In accordance with Guilford, it appears obvious that "statistics enable us to summarize our results in meaningful and convenient form". The summaries of the test results will enable educators in the field of science education to make general conclusions and predictions concerning the presentation of the particular scientific principle that has been demonstrated. Experimental and statistical methods cannot be divorced from each other, in so far as, "The experiment directs our observations and yields data. By means of statistical methods, we can summarize those data, interpret them, and determine their

reliability.\footnote{1}

In this respect, Brownell\footnote{2} has stated that "Altogether too commonly understandings are disregarded in evaluation (and in teaching) in favor of outcomes which are more easily measured (and achieved)."

**Test-retest method.**—If the educator is to measure growth, or learning, due to some specific learning activity, he cannot overemphasize the "importance of knowing initial status with respect to understanding."\footnote{3} The writer is convinced that the only reliable method of measuring the amount of learning, due to some specific activity, is by means of the test-retest method. That is, by administering identical tests prior to, and after the learning situation. It is conceded that "not all the gain found can be correctly attributed solely to the remedial program \footnote{7} the demonstration period\footnote{7}. Some of it is doubtless due to the practice effect or to familiarity with the test itself, part of it to teaching received outside of school, and part of it to natural growth."\footnote{11} For purposes of predicting this "probable gain", the writer has made use of a control group in the experiment.


What the test endeavors to determine.--Any increase in scores of the control group on the post-test (the same test that has been given the second time) may be labelled as the probable gain that can be attributed to familiarity, or external factors concerning the test. The writer has sought to determine the significant increase of the scores on the post-test of the experimental group, and compare this increase with any possible increase made by the control group on the post-test. By knowing approximately what percentage gain on the test scores may be attributed to "chance", as determined by the control group, the writer is able to conclude in this instance, that any significantly larger gain in the scores of the experiment group has been due to learning gained during the demonstration process.

Assuming that the constructed test is both reliable and valid, statistics applied to the results emanating from the test will yield invaluable data in predicting at what grade, or grades this specific scientific principle can be presented with predictively good results. Statistical interpretations of the test results are the means to these predictions. This is stated in essence by Guilford\(^1\) who states that "statistical reasoning is basic to all predictions".

\(^1\) Ibid., p. 176.
C. Characteristics of the Test

Reliability of the test.--The reliability, being the precision and consistency with which the test measures "what it measures", is a most important characteristic of the test. In this specific testing situation, the scores on the pre-tests and post-tests given to the experimental group cannot be correlated for purposes of determining reliability since the material being tested has been presented to the examinees in the period intervening the two tests.

All external factors concerning the test have been kept as consistent as possible. The element of time does not detract from the reliability, because provisions have been made for each pupil to at least consider all the test items. The influential time factor has been kept at a minimum. Lindquist concurs in stating that "The procedures become entirely unsatisfactory particularly in any test in which speed is a significant element in the score".

The sampling of the material has been adequate, since all the test items have been constructed on the basis of a single scientific demonstration. A test of high reliability is further assured in the length of the test. It is generally conceived that the longer the test, the higher the

\[\text{E. F. Lindquist (Editor), Educational Measurement, op. cit., p. 617.}\]
reliability. The test in consideration contains approximately thirty items, measuring the understanding derived from a single scientific principle.

Validity.--Validation of the test items has been by jury, as mentioned previously. The jury was composed of in-service science teachers.
CHAPTER III

EXPERIMENTAL PROCEDURE

1. Introduction

The principle used in this part of the group study was the determination for the modal age level for grades seven and nine of the difficulty of the principle: gases, liquids and solids expand when heated and contract when cooled.

2. Description of Schools

Two representative communities located a short distance to the South of Boston were used to carry out this study. The towns will not be mentioned by name, but will be referred to as A and B.

Town A.-- In this town, two Junior High Schools were used. The first school was built only a year ago. It is of the latest architectural design. No expense was spared in order to assure the pupils of every educational opportunity. Two seventh grade classes and two ninth grade classes participated. The students expressed a great deal of interest in the demonstration. This school will be referred to in the following tables as school number 1.

The second school in this town was located in the same building as the Senior High School. However, the seventh and
eighth grades were separated from the upper four grades. Two seventh grade classes and two ninth grade classes were obtained. The seventh grade classes were very interested in the demonstration. The ninth grade classes, however, were rather indifferent. This school will be referred to in the following tables as school number 2.

**Town E.**— Two seventh grade classes were tested in this town. Both of these classes were in the same school building, which was of post Civil War construction and is now in the process of being completely renovated. There were only seventh and eighth grades in this school. This was a result of overcrowding in the lower grades. The pupils exhibited a great deal of enthusiasm towards the demonstration. This school will be referred to in the following tables as school number 3.

**Distribution of pupils in Cooperating Schools.** — A total of two hundred and six pupils were tested in this study. The control group consisted of one hundred and five pupils and the experimental group consisted of ninety-one. The following table will show the distribution of these pupils not only by groups, but by schools as well. In recording the pertinent data for this chapter and the next, all seventh grade classes are considered as one class. This applies also to the ninth grade classes.
Table 1. Distribution of Pupils According to Schools in the Control and Experimental Groups.

<table>
<thead>
<tr>
<th>School</th>
<th>Grade 7</th>
<th></th>
<th>Grade 9</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Con.</td>
<td>Exp.</td>
<td>Con.</td>
<td>Exp.</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>(2) 14</td>
<td>(3) 14</td>
<td>(4) 16</td>
<td>(5) 15</td>
<td>59</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>(2) 33</td>
<td>(3) 31</td>
<td>(4) 22</td>
<td>(5) 21</td>
<td>107</td>
</tr>
<tr>
<td>3</td>
<td>(2) 20</td>
<td>(3) 20</td>
<td>(4) --</td>
<td>(5) --</td>
<td>40</td>
</tr>
<tr>
<td>Totals</td>
<td>(2) 67</td>
<td>(3) 65</td>
<td>(4) 38</td>
<td>(5) 36</td>
<td>206</td>
</tr>
</tbody>
</table>

Note that the following table will show that the sampling in the seventh grade according to sex was very good for the control group. The experimental group had an approximate ratio of two boys to one girl. In the ninth grade this ratio was considerably reversed in both groups in favor of the girls. The ratio is approximately 2.6 girls to 1 boy. This can be explained by the fact that two of the ninth grade classes tested were all girl classes.

Table 2. Distribution of Pupils According to Sex in the Control and Experimental Groups.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Control</th>
<th></th>
<th>Experimental</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2) 34</td>
<td>(3) 33</td>
<td>(4) 25</td>
<td>(5) 40</td>
<td>132</td>
</tr>
<tr>
<td>Seven</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nine</td>
<td>(2) 27</td>
<td>(3) 11</td>
<td>(4) 26</td>
<td>(5) 10</td>
<td>74</td>
</tr>
<tr>
<td>Totals</td>
<td>(2) 61</td>
<td>(3) 44</td>
<td>(4) 51</td>
<td>(5) 50</td>
<td>206</td>
</tr>
</tbody>
</table>
3. Classroom Procedure

The class was first divided into two groups by random selection. The experimental group remained in the room. The control group left the room in charge of the classroom teacher.

A pre-test was administered to the experimental group. A time limit was set for ten minutes. The demonstration was then shown. Immediately following the demonstration a post-test was given. The time limit was also set for ten minutes.

The same procedure as described above was followed with the control group. The one exception being that the demonstration was not shown. During that period the control group read magazines.

An answer sheet was passed out with the test booklet. After the pre-test, the answer sheet was collected. The test booklet was turned upside down on the pupils' desks. At the end of the demonstration, another answer sheet was passed out for the post-test. The same test was used for the pre-test and the post-test. The entire period was forty minutes long.

Table 3. Distribution of Time during the Classroom Period

<table>
<thead>
<tr>
<th>Minutes</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>10</td>
<td>Pre-test</td>
<td>Pre-test</td>
</tr>
<tr>
<td>20</td>
<td>Demonstration</td>
<td>Read Magazine</td>
</tr>
<tr>
<td>10</td>
<td>Post-test</td>
<td>Post-test</td>
</tr>
</tbody>
</table>
4. The Test

The test was divided into three sections. Each section contained ten questions. The first section consisted of questions based on the demonstration. The second section contained questions based on applications of the principle. The questions in the third section were of greater difficulty than those asked in either section one or two and required a greater amount of thought in order to answer them correctly. Each question had four possible answers, only one being correct. A copy of the test may be found in the appendix.

5. The Demonstration

The following equipment was used in the demonstration:-

A. Solids
   1. Classic ball
   2. Ring

B. Gases
   1. "U" tube
   2. Ring stand and clamps
   3. 1000cc flask
   4. Rubber stoppers
   5. Glass and rubber tubing
   6. Alcohol lamp

C. Liquids
   1. 500cc flask
   2. Long glass tube with stopper
3. Pan of ice
4. Colored water

This is a picture of the apparatus set up for the demonstration.

Plate 1. Demonstration Equipment

The script which accompanies this demonstration may be found in the Appendix.

6. Obtaining and Recording Test Results

Record Card. -- A separate report card was made for each and every pupil tested. The information recorded on this card was very informative. It was as follows:

1. Name
2. School
3. Town
4. Sex
5. Chronological age
6. I.Q.
7. Group
   A. Experimental
   B. Control
8. Questions correctly answered
   A. Pre-test
   B. Post-test
9. Total Scores of both tests
10. Grade level
11. Mental age

This card saved considerable amount of time in compiling the data. For example, the modal mental age for an entire group was obtained by placing a mark in the appropriate block on the edge of the card. The cards were then spread out until all the marks of the other cards could be seen. The modal age could then be easily found by inspection. The same procedure was followed in collecting other data from these cards. A sample of this card may be found in the Appendix.

**Modal Mental Age.**—The modal mental age may be defined as that mental age which is most prevalent in a group. The mode is defined as that point where the greatest number occurs in a distribution.

In order to determine the I.Q. and the mental age of each pupil, the Otis Quick-Scoring Mental Ability Test was used.
Specifically, the Beta Test Form Cm with separate answer sheets was used.

With this information the modal mental age for both grades was found. When the mode was established, the experimental and control groups were each divided into three sub-groups as follows:

1. Overage deviate group
2. Modal mental age group
3. Underage deviate group

The next two tables show that the random sampling done in the seventh grade was effective. The modal mental age group, as well as the overage and underage deviate groups, are exactly the same in both the experimental and control groups. The range in the various groups differ by only two or three months.

Table 4. Comparison of Mental Ability in Control and Experimental Groups of the Seventh Grade

<table>
<thead>
<tr>
<th>Grade and Group</th>
<th>Sub-group</th>
<th>Modal Mental Age for Each Group</th>
<th>Range in Mental Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Seventh Control</td>
<td>Overage deviate</td>
<td>14-1</td>
<td>17-7</td>
</tr>
<tr>
<td></td>
<td>Modal Mental Age</td>
<td>13-1</td>
<td>13-4</td>
</tr>
<tr>
<td></td>
<td>Underage deviate</td>
<td>11-10</td>
<td>12-4</td>
</tr>
</tbody>
</table>
Table 4. (concluded)

<table>
<thead>
<tr>
<th>Grade and Group</th>
<th>Sub-group</th>
<th>Modal Mental Age for Each Group</th>
<th>Range in Mental Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Seventh</td>
<td>Overage deviate</td>
<td>14-1</td>
<td>17-7 to 13-10</td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td>13-1</td>
<td>13-4 to 12-7</td>
</tr>
<tr>
<td></td>
<td>Modal Mental Age</td>
<td>11-10</td>
<td>12-4 to 9-7</td>
</tr>
<tr>
<td></td>
<td>Underage deviate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The random sampling done in the ninth grade was not as effective as in the seventh. This may be explained by the fact that twice as many seventh grade pupils were tested as were ninth grade pupils. Only in the overage deviate groups are the modal mental ages equal. The range in the modal mental age groups are similar. In the range of the underage deviate groups, there is the greatest difference in the mental ages. However, when the total range in the experimental and control groups are considered, the mental ability of both groups is approximately equal.

The following two tables will show the same comparison with the ninth grade groups as was previously shown with the seventh grade groups in Table 4.
Table 5. Comparison of Mental Ability in Control and Experimental Groups of the Ninth Grade

<table>
<thead>
<tr>
<th>Grade and Group</th>
<th>Sub-group</th>
<th>Modal Mental Age for Each Group</th>
<th>Range in Mental Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Ninth Control</td>
<td>Overage deviate</td>
<td>15-7</td>
<td>17-7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to 15-4</td>
</tr>
<tr>
<td></td>
<td>Modal Mental Age</td>
<td>15-1</td>
<td>15-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to 14-1</td>
</tr>
<tr>
<td></td>
<td>Underage deviate</td>
<td>12-7</td>
<td>13-7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to 10-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade and Group</th>
<th>Sub-group</th>
<th>Modal Mental Age for Each Group</th>
<th>Range in Mental Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Ninth Experimental</td>
<td>Overage deviate</td>
<td>15-7</td>
<td>17-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to 15-7</td>
</tr>
<tr>
<td></td>
<td>Modal Mental Age</td>
<td>14-10</td>
<td>15-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to 14-1</td>
</tr>
<tr>
<td></td>
<td>Underage deviate</td>
<td>13-4</td>
<td>13-10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to 12-7</td>
</tr>
</tbody>
</table>

The following two tables show the number of pupils in the three sub-groups by grade, school and control or experimental groups.
Table 6. Number of Pupils in the Sub-groups of the Seventh Grade Control and Experimental Groups

<table>
<thead>
<tr>
<th>Grade and School</th>
<th>Group</th>
<th>Overage deviate</th>
<th>Modal Mental Age</th>
<th>Underage deviate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Seven</td>
<td>Control</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>14</td>
<td>13</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Seven</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>11</td>
<td>8</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>49</td>
<td>38</td>
<td>45</td>
<td>132</td>
</tr>
</tbody>
</table>

The next table will show a similar picture for the ninth grade control and experimental groups.

Table 7. Number of Pupils in the Sub-groups of the Ninth Grade Control and Experimental Groups

<table>
<thead>
<tr>
<th>Grade and School</th>
<th>Group</th>
<th>Overage deviate</th>
<th>Modal Mental Age</th>
<th>Underage deviate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Nine</td>
<td>Control</td>
<td>4</td>
<td>9</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Nine</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>7</td>
<td>4</td>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11</td>
<td>10</td>
<td>15</td>
<td>36</td>
</tr>
</tbody>
</table>


The tables that follow are intended to show the relationship of the modal mental age to the modal chronological age of each sub-group. The modal chronological age was obtained by taking each sub-group and finding the age of most common occurrence; i.e. the mode.

In the seventh grade, the overage deviate groups have a modal mental age of one and a half years older than their modal chronological age. The modal mental age of the seventh grade overage groups is superior to the ninth grade underage deviate groups.

In the ninth grade there is relatively little difference in the modal mental age and the modal chronological age of the average group.

The seventh and ninth grades had in their modal mental age groups an average of a half year increase over their modal chronological age. The underage groups in both grades have a marked decrease in the modal mental age when compared to their modal chronological age. It is interesting to note that the average difference of the sub-groups in the seventh and ninth grades is 1 - 6 mental age years.
Table 8. A Comparison of the Modal Chronological Ages and the Modal Mental Ages of the Three Sub-groups in the Seventh Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>Group</th>
<th>Sub-group</th>
<th>Modal Chronological Age</th>
<th>Modal Mental Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Seven</td>
<td>Control</td>
<td>Overage</td>
<td>12-7</td>
<td>14-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modal Mental Age</td>
<td>12-7</td>
<td>13-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Underage</td>
<td>12-5</td>
<td>11-10</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>Overage</td>
<td>12-5</td>
<td>14-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modal Mental Age</td>
<td>12-5</td>
<td>13-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Underage</td>
<td>13-9</td>
<td>11-10</td>
</tr>
</tbody>
</table>

The next table will show a similar picture for the ninth grade control and experimental groups.

Table 9. A Comparison of the Modal Chronological Ages and the Modal Mental Ages of the Three Sub-groups in the Ninth Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>Group</th>
<th>Sub-group</th>
<th>Modal Chronological Age</th>
<th>Modal Mental Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Nine</td>
<td>Control</td>
<td>Overage</td>
<td>15-6</td>
<td>15-7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modal Mental Age</td>
<td>14-8</td>
<td>15-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Underage</td>
<td>14-10</td>
<td>12-7</td>
</tr>
</tbody>
</table>

(concluded on next page)
Table 9. (concluded)

<table>
<thead>
<tr>
<th>Grade</th>
<th>Group</th>
<th>Sub-group</th>
<th>Modal Chronological Age</th>
<th>Modal Mental Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Nine</td>
<td>Experimental</td>
<td>Overage</td>
<td>15-7</td>
<td>15-7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modal Mental</td>
<td>14-9</td>
<td>14-10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Underage</td>
<td>14-7</td>
<td>13-4</td>
</tr>
</tbody>
</table>

Recording of Data.— The Otis Normal Percentile Chart was used to record the scores of the various groups and to compare the scores of one group with another.

There are many advantages in using this chart. The chart is so constructed that a normal distribution curve becomes a straight line. In this type of chart the scores at both extreme ends of the distribution are read with just as much accuracy as those in the middle.

Code used in reading the charts.— The following is the code used in constructing or in reading the percentile charts which follow:

- pre means pre-test
- post means post-test
- 7 means seventh grade
- 9 means ninth grade
- x means the experimental group
- c means the control group
Sub$_u$ means the underage (mental age) group of deviates.

Sub$_o$ means the overage (mental age) group of deviates.

Red lines report pre-tests
Blue lines report post-tests
Solid lines denote experimental groups
Dotted lines denote control groups
CHAPTER IV
SUMMARY AND FINDINGS

1. Otis Normal Percentile Charts

Chart I.-- The purpose of Chart I is to establish if the random division of pupils was effective. In the case of the seventh grade this division was effective. However, the ninth grade control group proved superior to the ninth grade experimental group.

Chart II.-- This chart shows the achievement of the seventh and ninth grade modal mental age control and experimental groups. The two experimental groups improved greatly over the control groups. The seventh grade experimental was greater than the ninth grade experimental as shown by the larger area between the lines representing the control and experimental groups.

Chart III.-- In this chart the gains between the pre-tests and the post-tests of the modal mental age experimental groups of both grades are compared. As the chart shows, both grades gained by approximately the same amount. The ninth grade scored higher than the seventh grade in both tests.

Chart IV.-- Chart IV compares the pre-tests and post-tests of the underage control and experimental groups. It shows that both groups benefited from the demonstration and
by the same amount. The ninth grade scored higher than the seventh grade in both tests.

Chart V.-- The purpose of Chart V is to compare the achievement on the pre-tests and post-tests of the average experimental groups. In the pre-test and the post-test the two groups scored closely together with the ninth grade group scoring slightly higher.

Chart VI.-- The achievement of the three seventh grade experimental groups are compared in Chart VI. The underage deviate group scored lower than the modal mental age group. The overage deviate group scored the highest.

Chart VII.-- The achievement of the three ninth grade experimental groups are compared in Chart VII. The underage deviate group scored much lower than the other two groups. The modal mental age group and overage deviate group scored very close together. The overage deviate group was slightly higher than the modal mental age group. This graph indicates that the saturation point in the understanding of the principle is being reached.

2. Recapitulation

The random division of the pupils was effective in the seventh grade control and experimental groups. The ninth grade was superior in the control group.

All experimental groups showed a gain in the post-tests. Therefore, the underage deviate groups, modal mental age groups
and overage groups benefited from the demonstration.

In the seventh grade there was a difference in the scores in the three experimental groups in the post-tests. However, in the ninth grade only the underage deviate group scored low. The modal mental age group and the overage deviate group scores were very close together in the post-tests. This would seem to indicate that the saturation point in the understanding of the principle is being reached.

At this time, it is impossible to draw any definite conclusions. In time when each grade level has been tested and a substantial number of pupils tested, a definite statement can be made.
APPENDIX A

Otis Normal Percentile Charts and Worksheets
## Worksheet for Chart No. 1
(Mental Mental Age Group)

<table>
<thead>
<tr>
<th>INTERVAL</th>
<th>EXPERIMENTAL PRE-TEST</th>
<th>CONTROL PRE-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>F  S  F</td>
<td>F  S  F</td>
</tr>
<tr>
<td>(1)</td>
<td>(2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13)</td>
<td></td>
</tr>
<tr>
<td>28-29</td>
<td>1 1 16 100.0</td>
<td></td>
</tr>
<tr>
<td>26-27</td>
<td>0 15 0.0</td>
<td></td>
</tr>
<tr>
<td>24-25</td>
<td>0 15 0.0</td>
<td></td>
</tr>
<tr>
<td>22-23</td>
<td>5 15 95.6</td>
<td></td>
</tr>
<tr>
<td>20-21</td>
<td>1 1 16 100.0 2 61 100.0</td>
<td></td>
</tr>
<tr>
<td>18-19</td>
<td>0 17 0.0 2 19 90.0 2 20 100.0 2 9 56.5</td>
<td></td>
</tr>
<tr>
<td>16-17</td>
<td>1 17 93.5 2 17 90.0 2 5 90.0 1 7 43.6</td>
<td></td>
</tr>
<tr>
<td>14-15</td>
<td>1 16 92.0 2 15 71.0 1 16 80.0 1 7 45.5</td>
<td></td>
</tr>
<tr>
<td>12-13</td>
<td>2 15 82.5 1 7 80.0 3 10 80.0 3 4 25.0</td>
<td></td>
</tr>
<tr>
<td>10-11</td>
<td>1 1 71.5 2 6 39.0 3 13 0.0 1 1 6.5</td>
<td></td>
</tr>
<tr>
<td>8-9</td>
<td>7 12 66.0 2 4 29.0 6 18 65.0</td>
<td></td>
</tr>
<tr>
<td>6-7</td>
<td>4 5 27.5 2 2 10.0 2 7 39.0</td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td>1 1 5.6</td>
<td>2 5 25.0</td>
</tr>
<tr>
<td>2-3</td>
<td>2 3 15.0</td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>1 1</td>
<td>5.0</td>
</tr>
</tbody>
</table>

**Codes:**
- F = Frequency
- S = Sub-total
- P = Percent
### Worksheet for Chart No. 2

(Modal Mental Age Group)

<table>
<thead>
<tr>
<th>INTERVAL</th>
<th>EXPERIMENTAL POST-TEST</th>
<th>CONTROL POST-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>26-29</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>26-27</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>24-25</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>22-23</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>20-21</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>18-19</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>16-17</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>14-15</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12-13</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10-11</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>8-9</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>6-7</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>4-5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2-3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Code:**
- **F** = Frequency
- **S** = Sub-total
- **P** = Percent
### Worksheet for Chart No. 5

**Modal Mental Age Group**

<table>
<thead>
<tr>
<th>INTERVAL</th>
<th>EXPERIMENTAL: PRE-TEST</th>
<th>EXPERIMENTAL: POST-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>26-27</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>24-25</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>22-23</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>20-21</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>18-19</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>16-17</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>14-15</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>12-13</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>10-11</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>8-9</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>6-7</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4-5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2-3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0-1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Code:**
- **F** = Frequency
- **S** = Sub-total
- **P** = Percent
## Normal Percentile Chart

<table>
<thead>
<tr>
<th>Grade or group</th>
<th>No. of cases</th>
<th>Measure (Examination)</th>
<th>Form</th>
<th>Date</th>
<th>Examiner</th>
<th>Graphs by</th>
<th>School</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Pre-X7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-X7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pre-X9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post-YQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table I

<table>
<thead>
<tr>
<th>Variable II</th>
<th>Subscores</th>
<th>Score intervals</th>
<th>Frequencies</th>
<th>Subscores</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Percentile Scale

- Pre-test Experimental
- Post-test Experimental
- Pre-test Experimental Grade 7
- Post-test Experimental Grade 7
- Pre-test Experimental Grade 3
- Post-test Experimental Grade 3

### Standard Deviation Scale

-3σ  -2σ  -σ  0σ  +σ  +2σ  +3σ

---

Published by World Book Company, Yonkers-on-Hudson, New York, and Chicago, Illinois. Copyright 1938 by World Book Company. Copyright in Great Britain. All rights reserved.
<table>
<thead>
<tr>
<th>INTERVAL</th>
<th>EXPERIMENTAL: PRE-TEST</th>
<th>EXPERIMENTAL: POST-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>F  S  F</td>
<td>F  S  F</td>
</tr>
<tr>
<td>(1)</td>
<td>(2) (3) (4) (5) (6) (7)</td>
<td>(8) (9) (10) (11) (12) (13)</td>
</tr>
<tr>
<td>26-29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26-27</td>
<td>1 25 100.0</td>
<td>1 15 100.0</td>
</tr>
<tr>
<td>24-25</td>
<td>2 24 96.0</td>
<td>2 14 93.2</td>
</tr>
<tr>
<td>22-23</td>
<td>1 16 100.0</td>
<td>3 22 88.0</td>
</tr>
<tr>
<td>20-21</td>
<td>1 14 93.2</td>
<td>1 19 76.0</td>
</tr>
<tr>
<td>18-19</td>
<td>2 25 100.0</td>
<td>0 13 0.0</td>
</tr>
<tr>
<td>16-17</td>
<td>1 23 92.0</td>
<td>3 13 86.6</td>
</tr>
<tr>
<td>14-15</td>
<td>0 22 0.0</td>
<td>1 10 66.7</td>
</tr>
<tr>
<td>12-13</td>
<td>5 22 88.0</td>
<td>3 9 59.9</td>
</tr>
<tr>
<td>10-11</td>
<td>2 17 69.0</td>
<td>3 6 39.9</td>
</tr>
<tr>
<td>8-9</td>
<td>1 15 60.0</td>
<td>0 3 0.0</td>
</tr>
<tr>
<td>6-7</td>
<td>7 14 56.0</td>
<td>2 3 20.0</td>
</tr>
<tr>
<td>4-5</td>
<td>4 7 29.0</td>
<td>1 1 6.7</td>
</tr>
<tr>
<td>2-3</td>
<td>3 3 12.0</td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**WORKSHEET FOR CHART NO. 5**

**(Average Deviate Group)**

<table>
<thead>
<tr>
<th>INTERVAL</th>
<th>PRE-TEST</th>
<th>POST-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F S F</td>
<td>F S F</td>
</tr>
<tr>
<td>28-29</td>
<td>1 22 100.0</td>
<td></td>
</tr>
<tr>
<td>26-27</td>
<td>1 11 100.0</td>
<td>4 21 90.0</td>
</tr>
<tr>
<td>24-25</td>
<td>2 22 100.0</td>
<td>1 10 99.1</td>
</tr>
<tr>
<td>22-23</td>
<td>2 20 90.0</td>
<td>1 9 99.1</td>
</tr>
<tr>
<td>20-21</td>
<td>1 18 81.0</td>
<td>0 8 0.0</td>
</tr>
<tr>
<td>18-19</td>
<td>1 17 76.5</td>
<td>0 8 0.0</td>
</tr>
<tr>
<td>16-17</td>
<td>3 16 73.0</td>
<td>2 8 79.2</td>
</tr>
<tr>
<td>14-15</td>
<td>1 15 58.5</td>
<td>4 6 59.4</td>
</tr>
<tr>
<td>12-13</td>
<td>3 12 54.0</td>
<td>0 2 0.0</td>
</tr>
<tr>
<td>10-11</td>
<td>1 9 40.0</td>
<td>1 2 19.6</td>
</tr>
<tr>
<td>8-9</td>
<td>5 0 33.0</td>
<td>0 1 0.0</td>
</tr>
<tr>
<td>6-7</td>
<td>0 3 0.0</td>
<td>1 1 9.9</td>
</tr>
<tr>
<td>4-5</td>
<td>2 5 15.5</td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>1 1 4.5</td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Code:**  
F = Frequency  
S = Sub-total  
P = Percent
### WORKSHEET FOR CHART NO. 6

#### EXPERIMENTAL: POST-TEST 1 SEVENTH GRADE

<table>
<thead>
<tr>
<th>INTERVAL</th>
<th>MODAL MENTAL AGE</th>
<th>UNDERAGE GROUP</th>
<th>OVERAGE GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>S</td>
<td>F</td>
<td>S</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>22-29</td>
<td>1 16</td>
<td>100.0</td>
<td>1 25</td>
</tr>
<tr>
<td>26-27</td>
<td>2 18</td>
<td>95.5</td>
<td>2 24</td>
</tr>
<tr>
<td>24-25</td>
<td>2 16</td>
<td>92.5</td>
<td>3 23</td>
</tr>
<tr>
<td>22-23</td>
<td>2 18</td>
<td>91.0</td>
<td>1 19</td>
</tr>
<tr>
<td>20-21</td>
<td>2 15</td>
<td>81.5</td>
<td>1 19</td>
</tr>
<tr>
<td>18-19</td>
<td>2 11</td>
<td>60.5</td>
<td>1 18</td>
</tr>
<tr>
<td>16-17</td>
<td>5 9</td>
<td>49.5</td>
<td>5 17</td>
</tr>
<tr>
<td>14-15</td>
<td>3 4</td>
<td>22.0</td>
<td>4 12</td>
</tr>
<tr>
<td>12-13</td>
<td>1 1</td>
<td>5.5</td>
<td>2 8</td>
</tr>
<tr>
<td>10-11</td>
<td></td>
<td></td>
<td>3 6</td>
</tr>
<tr>
<td>8-9</td>
<td></td>
<td></td>
<td>3 3</td>
</tr>
<tr>
<td>6-7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Code:**
- **F** = Frequency
- **S** = Sub-total
- **P** = Percent
### NORMAL PERCENTILE CHART

Grade or group | No. of cases | Measure (Examination) | Form | Date | Examiner | Graphs by | School | City
---|---|---|---|---|---|---|---|---

<p>| Variable I | Subs- | Per | Score | Subs- | Per | 1 | 2 | 3 | 4 | 5 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 95 | 96 | 97 | 98 | 99 | .5 | .6 | .7 | .8 | .9 | .99 | .999 |</p>
<table>
<thead>
<tr>
<th>Subtotal</th>
<th>Per cents</th>
<th>Score intervals</th>
<th>Freq.</th>
<th>Subtotals</th>
<th>Per cents</th>
<th>Experimental Grade 7</th>
<th>Underage Deviates</th>
<th>Normal Mental Age</th>
<th>Average Deviates</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>28-29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26-27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24-25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16-17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12-13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Standard Deviation Scale

-3σ -2σ -σ 0 +σ +2σ +3σ

---

*Published by World Book Company, Yonkers-on-Hudson, New York, and Chicago, Illinois. Copyright 1938 by World Book Company. Copyright in Great Britain. All rights reserved.*

*PRINTED IN U.S.A. ONPC-9*
## Worksheet For Chart No. 7

<table>
<thead>
<tr>
<th>INTERVAL</th>
<th>EXPERIMENTAL</th>
<th>POST-TEST</th>
<th>SEVENTH GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MODAL</td>
<td>MENTAL AGE</td>
<td>UNDERAGE GROUP</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>S</td>
<td>F</td>
</tr>
<tr>
<td>28-29</td>
<td>1</td>
<td>16</td>
<td>100.0</td>
</tr>
<tr>
<td>26-27</td>
<td>0</td>
<td>15</td>
<td>0.0</td>
</tr>
<tr>
<td>24-25</td>
<td>1</td>
<td>15</td>
<td>95.8</td>
</tr>
<tr>
<td>22-23</td>
<td>2</td>
<td>14</td>
<td>87.5</td>
</tr>
<tr>
<td>20-21</td>
<td>4</td>
<td>12</td>
<td>75.0</td>
</tr>
<tr>
<td>18-19</td>
<td>2</td>
<td>8</td>
<td>50.0</td>
</tr>
<tr>
<td>16-17</td>
<td>2</td>
<td>6</td>
<td>37.5</td>
</tr>
<tr>
<td>14-15</td>
<td>1</td>
<td>4</td>
<td>25.0</td>
</tr>
<tr>
<td>12-13</td>
<td>3</td>
<td>3</td>
<td>18.0</td>
</tr>
<tr>
<td>10-11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Code:**
- F = Frequency
- S = Sub-total
- F = Percent
# NORMAL PERCENTILE CHART

<table>
<thead>
<tr>
<th>Grade or group</th>
<th>No. of cases</th>
<th>Measure (Examination)</th>
<th>Form</th>
<th>Date</th>
<th>Examiner</th>
<th>Graphs by</th>
<th>School</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Post-X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Post-X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Post-X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Table 1

<table>
<thead>
<tr>
<th>Sub-</th>
<th>Score</th>
<th>Sub-</th>
<th>Per</th>
</tr>
</thead>
<tbody>
<tr>
<td>tota</td>
<td>(cents)</td>
<td>totals</td>
<td>(cents)</td>
</tr>
<tr>
<td>1</td>
<td>1.2</td>
<td>3.4</td>
<td>5</td>
</tr>
<tr>
<td>1.1</td>
<td>1.2</td>
<td>3.4</td>
<td>5</td>
</tr>
<tr>
<td>1.1</td>
<td>1.2</td>
<td>3.4</td>
<td>5</td>
</tr>
<tr>
<td>1.1</td>
<td>1.2</td>
<td>3.4</td>
<td>5</td>
</tr>
</tbody>
</table>

### Percentile Scale

- **Experimental Grade 9**
- **1 = Underage Deviates**
- **2 = Modal Mental Age**
- **3 = Overage Deviates**

## Standard Deviation Scale

- $-3\sigma$ to $+3\sigma$

---

Published by World Book Company, Yonkers-on-Hudson, New York, and Chicago, Illinois. Copyright 1938 by World Book Company. Copyright in Great Britain. All rights reserved.
APPENDIX B

Testing Materials
TEST

Directions: Fill in the blanks at the top of the answer sheet. Each question in the test has four choices lettered A, B, C and D. On your answer sheet place a cross (x) under the letter of the correct answer. Answer every question. Do not write on the test paper. Use the answer sheet.

Sample question: What color is snow?
A. Black
B. White
C. Red
D. Green

Answer on the Answer Sheet:

A  B  C  D
( ) (x) ( ) ( )

The correct letter was B. Therefore, a cross has been placed under the letter B. Do not do this.

/  B  C  D
( ) (x) ( ) ( )
**SCIENCE TEST**

**Directions:** Answer every question. Do not write on this paper. Use the answer sheet.

1. What happens when air is heated?
   A. Air increases in size.
   B. Air decreases in size.
   C. Air will weigh more.
   D. Nothing happens.

2. When a metal is cold, it is --
   A. very cold.
   B. painful to touch.
   C. larger when it was hot.
   D. smaller than when hot.

3. What happens to a liquid if it contracts?
   A. It will become colder.
   B. Nothing will happen.
   C. It will get hot.
   D. It will expand.

4. Why are sidewalks made with many cement blocks instead of one long piece?
   A. It will prevent the sidewalk from cracking.
   B. The cracks allow the rain to wash off.
   C. It makes the sidewalk look pretty.
   D. Cement comes only in blocks.

5. During the winter a water-tank is filled to the very top. What will happen in the summertime?
   A. Nothing will happen.
   B. The water will contract.
   C. More water can be placed in the tank.
   D. It will overflow.

6. What will happen if a metal ball is heated?
   A. It will contract.
   B. It will expand.
   C. It will evaporate.
   D. It will explode.
7. A frying pan will become a little bigger when it is —
   A. placed in a snowbank.
   B. not being used.
   C. placed on a hot stove.
   D. placed on a cold stove.

8. When a copper wire is heated, what will happen to it?
   A. It will become cold.
   B. The wire will become shorter.
   C. Nothing will happen.
   D. It will become longer.

9. When a liquid expands, it must become —
   A. colder.
   B. smaller.
   C. hotter.
   D. contracted.

10. When air is cooled, it becomes —
    A. bigger.
    B. expanded.
    C. smaller.
    D. cloudy.

11. Two nails are the same length. One nail is heated. The other is placed in ice-water. They are now measured. What has happened?
    A. They are still the same length.
    B. The cold nail is longer.
    C. Both are now smaller.
    D. The hot nail is longer.

12. Telephone wire are placed tightly between two poles on a hot summer day. What will happen to the wire during the winter?
    A. Nothing.
    B. Break.
    C. Expand.
    D. Carry a lot of calls.

13. A large saucepan was filled to the very top with boiling water. The saucepan was allowed to cool. What happened to the water?
A. The water overflowed.
B. The water contracted.
C. Nothing.
D. The water kept boiling.

14. A balloon was filled with air. It was then held over a hot stove. What happened to the balloon?
   A. The air in the balloon became cold.
   B. It became smaller.
   C. Nothing.
   D. It became larger.

15. An expansion tank is used in a hot-water heating system. Why is this tank necessary?
   A. Hot water occupies more space.
   B. Hot water occupies less space.
   C. It is used to store water.
   D. People bathe in it.

16. A radiator in the living-room will be larger --
   A. when it is cold.
   B. when it is painted silver.
   C. when it is painted black.
   D. when it is hot.

17. During a cold day a man filled his car tires with a lot of air. The weather the next day was very warm. What happened to the tires?
   A. They became very flat.
   B. They remained the same.
   C. They exploded.
   D. The air contracted.

The next question is number 18. Check your answer sheet. Are you on number 18?

18. A car radiator is used to cool the engine. What would happen to the radiator if it did not have an overflow pipe?
   A. The water would expand and break the radiator.
   B. The engine would not run.
   C. The water would remain cold.
   D. Nothing.
19. Glass is held in a window by strips of metal. In the summer, the glass is held tight. What will happen in the winter?
   A. The glass will be just as tight.
   B. The glass will be loose.
   C. The glass will be held very tight.
   D. The glass will be hot.

20. What will happen if a balloon filled with air is placed in a pan of ice water?
   A. The balloon will become larger.
   B. The air will contract and the balloon will expand.
   C. The air will expand and the balloon will contract.
   D. The balloon will become smaller.

21. A glass stopper on a bottle will not come out. What would you do?
   A. Place the stopper and bottle in cold water.
   B. Place only the bottle in hot water.
   C. Place the stopper and bottle in hot water.
   D. Place only the bottle in cold water.

22. The shorter a pendulum on a clock, the faster the clock will run. What will happen if the clock is placed over a hot radiator.
   A. The clock will be a little fast.
   B. The clock will keep perfect time.
   C. The clock will be a little slow.
   D. The clock will stop.

23. Many times when a building is on fire, the doors and windows are blown out. Why does this happen?
   A. Fire is very explosive.
   B. Fire causes the air to expand.
   C. The windows and doors were not built correctly.
   D. Fire causes the doors and the windows to expand.

24. A gasoline station received 2500 gallons early one cool summer morning. The meter read "Full". Only five gallons were sold. At noontime, the meter again read "Full". What do you think?
   A. It is impossible.
   B. It is possible.
   C. The meter did not work.
   D. No gasoline was sold.
25. During one season of the year, a steel bridge is 900 feet long. What season of the year will the bridge be 900 feet 5\(\frac{1}{2}\) inches?
   A. Spring.
   B. Summer.
   C. Fall.
   D. Winter.

26. A large tin can is heated. The cover is then placed on it very tightly. This cover is air tight. The can is then placed in ice water. What will happen?
   A. The air pressure of the room will compress and break the can.
   B. The air in the can will expand and break the can.
   C. Nothing will happen except the can will be cooled off.
   D. The heat of the can will heat the water.

27. A metal flag pole is 200 feet high. The sun rises in the East and sets in the West. At noontime, the flag pole bends a slight amount. In what direction does it bend?
   A. North.
   B. East.
   C. South.
   D. West.

28. What is the best way to remove the metal covers of fruit or jam jars?
   A. Place only the covers in cold water.
   B. Saw of the covers.
   C. Break the jar.
   D. Pour hot water over the covers.

29. On what kind of a day would you let air out of a car tire to keep it at the right pressure?
   A. A hot day.
   B. A mild Spring day.
   C. A cold day.
   D. A mild Fall day.

30. If we applied the same amount of heat to the following items, which would expand the most?
   A. Water.
   B. A metal.
   C. Air
   D. A frying pan.
This is the end of the test. There are 50 questions. Have you answered them all?

Go back and check all your answers.
ANSWER SHEET

Name: ____________________________

School: __________________________

Date of Birth: ____________________

Age: ____________________

How many years have you studied science?

Do you plan to go to college?

Directions: Answer every question.

Sample question: What color is snow?

A. Black
B. White
C. Red
D. Green

Answer on answer sheet:

A B C D
(x) ( ) ( )

Do not do this:

A B C D
(x) ( ) (x)

1. A B C D
( ) ( ) ( )

2. A B C D
( ) ( ) ( ) ( )

3. A B C D
( ) ( ) ( ) ( )

4. A B C D
( ) ( ) ( ) ( )

5. A B C D
( ) ( ) ( ) ( )

6. A B C D
( ) ( ) ( ) ( )

7. A B C D
( ) ( ) ( ) ( )

8. A B C D
( ) ( ) ( ) ( )

9. A B C D
( ) ( ) ( ) ( )

10. A B C D
( ) ( ) ( ) ( )

11. A B C D
( ) ( ) ( ) ( )

12. A B C D
( ) ( ) ( ) ( )
The next answer space is for question 13. Are you reading question 13?

13. A B C D
   ( ) ( ) ( ) ( )

14. A B C D
   ( ) ( ) ( ) ( )

15. A B C D
   ( ) ( ) ( ) ( )

16. A B C D
   ( ) ( ) ( ) ( )

17. A B C D
   ( ) ( ) ( ) ( )

18. A B C D
   ( ) ( ) ( ) ( )

19. A B C D
   ( ) ( ) ( ) ( )

20. A B C D
   ( ) ( ) ( ) ( )

21. A B C D
   ( ) ( ) ( ) ( )

22. A B C D
   ( ) ( ) ( ) ( )

23. A B C D
   ( ) ( ) ( ) ( )

24. A B C D
   ( ) ( ) ( ) ( )

25. A B C D
   ( ) ( ) ( ) ( )

26. A B C D
   ( ) ( ) ( ) ( )

27. A B C D
   ( ) ( ) ( ) ( )

28. A B C D
   ( ) ( ) ( ) ( )

29. A B C D
   ( ) ( ) ( ) ( )

30. A B C D
   ( ) ( ) ( ) ( )
# THE RECORD CARD

<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.</td>
<td>E</td>
<td>C</td>
<td>B</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td>Name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td>Yr. Mo. Day.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td>Birthdate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td>School</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td>Teacher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td>Otis form</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td>Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td>IQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Date today</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Experimenter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Principle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Item</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Kit number</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mental age in mos., 100+</th>
</tr>
</thead>
<tbody>
<tr>
<td>8-10</td>
</tr>
<tr>
<td>11-13</td>
</tr>
<tr>
<td>14-16</td>
</tr>
<tr>
<td>17-19</td>
</tr>
<tr>
<td>20-22</td>
</tr>
<tr>
<td>23-25</td>
</tr>
<tr>
<td>26-28</td>
</tr>
<tr>
<td>29-31</td>
</tr>
<tr>
<td>32-34</td>
</tr>
<tr>
<td>35-37</td>
</tr>
<tr>
<td>38-40</td>
</tr>
<tr>
<td>41-43</td>
</tr>
<tr>
<td>44-46</td>
</tr>
<tr>
<td>47-49</td>
</tr>
<tr>
<td>50-52</td>
</tr>
<tr>
<td>53-55</td>
</tr>
<tr>
<td>56-58</td>
</tr>
<tr>
<td>59-61</td>
</tr>
<tr>
<td>62-64</td>
</tr>
<tr>
<td>65-67</td>
</tr>
<tr>
<td>68-70</td>
</tr>
<tr>
<td>71-73</td>
</tr>
<tr>
<td>74-76</td>
</tr>
<tr>
<td>77-79</td>
</tr>
<tr>
<td>80-82</td>
</tr>
<tr>
<td>83-85</td>
</tr>
<tr>
<td>86-88</td>
</tr>
<tr>
<td>89-91</td>
</tr>
<tr>
<td>92-94</td>
</tr>
<tr>
<td>95-97</td>
</tr>
<tr>
<td>98-00</td>
</tr>
<tr>
<td>1-03</td>
</tr>
<tr>
<td>4-06</td>
</tr>
<tr>
<td>7-09</td>
</tr>
<tr>
<td>10-12</td>
</tr>
<tr>
<td>13-15</td>
</tr>
<tr>
<td>16-18</td>
</tr>
<tr>
<td>19-21</td>
</tr>
<tr>
<td>22-24</td>
</tr>
</tbody>
</table>
Today we are going to learn an important scientific principle. However, before I tell you what this principle is, it is necessary that you understand the meaning of the two words "expand" and "contract". (Write words on blackboard.)

Let us take the word "expand" first. Expand means to grow larger or become bigger. When an athlete talks about his expansion, he is talking about how large his chest will become when he breathes in a large amount of air. If a storekeeper talks about expanding his business, he will either make the store bigger or make more money. In either case, the store or the profits will become larger.

Now let us take the word "contract". Contract means just the opposite of expand. Contract means to become smaller. Suppose you take a bath sponge that most people keep in their bathrooms and squeeze it between your hands. You have made the sponge smaller. The sponge has been contracted. Remember expand means to become larger and contract means to become smaller.

In order to make the meaning of these two words clearer, I have here a ball and ring. This ball and ring are made out of metal. Notice that the ball will just pass through the ring. I will now light this lamp. The lamp produces a very
hot flame. I will now place the ball in the flame. We will have to wait for a few moments for the metal to become hot.

Now that the ball is hot, I will try to put it through the ring. As you can see the metal ball has become larger and will not pass through the ring. In other words, the metal ball has expanded.

You will notice that in this pan there is some ice. The water is very cold. The ball is now placed in the ice water. The ball is cooled. Now the ball will easily pass through the ring. The cold has made the ball smaller or contracted it.

Let us see what will happen if we heat the ring. The ball goes through the ring without any trouble. Now we will cool the ring. The ball will just fit. Again the heat caused the metal to expand and the cold water caused it to contract.

You have probably noticed the other equipment here. Before I can use them, there is something else that you should know. Everything in the world -- solids, liquids and gases -- is made up of very tiny particles. You all know the A-Bomb is made up of atoms. Atoms are also very tiny. Two, three or more atoms make up a molecule. It is these molecules that make up everything in the world.

These molecules act like people. They are always moving. If the molecules become hot, they move very fast. If they become cold, they will move very slow.

We will now continue with this comparison of molecules and
people. Let us compare this flask with a bus. The flask is full of molecules of air. There is plenty of room for them to move about. On the bus, when everybody is seated, the people have plenty of room to move around without getting in each other's way. Now, when I heat this flask, the molecules of air will begin to move around and bump each other. There is not enough room for all the molecules. On the bus if all the people began to move around at the same time there would not be enough room for everybody. Some of the people would be pushed out. In the flask, some of the molecules of air will pushed out.

In order to prove to you that air will expand when heated, I have set up this experiment. When the flask is heated, the molecules of air begin to move faster. The air will expand. The molecules push down on this water in the "U" tube. As they molecules move faster and faster, air bubbles are pushed through the water into the air in the room. Just as some of the people were pushed out of the bus.

Let us now go back to the example of the bus. Suppose we wished to fill the bus with as many people as possible. In order to do this the people would have to sit and stand very still. If they moved around, they would begin to take up more room than they should. We can slow down molecules by making them cold.

I will take this pan of ice water and place the flask in it. The flask becomes cold and also the molecules of air inside
of it. The molecules do not move as quickly; hence, will not take up as much room. The air will contract. Notice that the water in the "U" tube moves up towards the flask, replacing the space that the air once occupied. The air has now contracted or has taken up less space.

Let us now turn our attention to this flask of colored water. Notice that when I place this rubber stopper with the glass tube into the flask some water goes up the tube. This is because the air between the stopper and the water has pushed down on the water and forced some water into the tube. We have already learned that when a solid or a gas is heated, they will expand and when they are cooled, they will contract. Now let us find out if the same thing will happen to a liquid when heated and cooled. In this case, we will use water.

I will now heat the water. Watch to see what happens. The molecules of water are now beginning to move very fast. They are taking up more space. The water is rising in the tube. The water is expanding.

Let us wait a moment until the flask cools. I am going to place the flask into the ice water. If I did this right after I heated the flask, it would crack. The flask and the water inside of it are now becoming cold. The molecules are slowing down. The water is contracting.

You have now learned a very important scientific principle. You have learned that when you heat any solid, liquid or gas they will expand and when you cool them, they will contract.
Perhaps you can think of many ways in which man has made use of this principle. It has been used to make life more comfortable and enjoyable for all of us.
APPENDIX C

Bibliography
BIBLIOGRAPHY

1. Anibel, Fred G., "Comparative Effectiveness of Lecture-Demonstration and Individual-Laboratory Method", 

2. Arthur, Paul, Lecture Demonstration in General Chemistry, 

3. Babitz and Keyes, "An Experiment in Teaching Pupils to 
   Apply Scientific Principles", Science Education 
   (December, 1939), 23:367-70.

   to the Organization of a Twelve Year Science Sequence", Science Education 
   (April, 1948), 34:176-177.

5. Bellack, Arno A., "Sequence and Grade Placement", 
   Journal of Educational Research (April, 1948), 
   41:610-623.

   Journal of Chemical Education (August, 1937).
   14:375-7.

7. Breckenridge, Marian E. and Elee Vincent, Child 
   Development, W. B. Saunders Company, Philadelphia, 
   1949.

8. Cahoon, G. P., "Using Demonstrations for Providing Pupil 
   Experiences in Thinking", Science Education 
   (October, 1946), 30:198-201.

9. Colvin, Stephen Sheldon, An Introduction to High 
   School Teaching, Ch. 12, The Macmillan Company, New York, 
   1924.

10. Coyle, J. P., E. C. Hansen, and R. B. Coe, "Demonstra- 
    tions Made More Visible", Chicago Schools Journal 
    (November, 1941), 23:64-8.

11. Croxton, W. C., "Pupils Ability to Generalize", School 
    Science and Mathematics (June, 1936), 36:627-634.

12. Cunningham, Harry A., "Lecture Demonstration Versus 
    Individual Laboratory Method in Science Teaching - 
    A Summary", Science Education (March, 1946), 
    30:70-82.


20. Freud, Henrietta Z. and N. D. Cheronis, "Retention in the Physical Science Survey Course", *Chemical Education* (June, 1940), 18:288-293.


77. West, Joe Young, "Do We Expect Too Much or Too Little of Children from their Experiences in Science?", Science Education (October, 1949), 33:296-298.
