1952

Determination for the modal age level for the grades IV and VI of the difficulty of the principle: a body at rest remains at rest, and a body in motion continues to move at constant speed along a straight line, unless the body is acted upon in either case by an unbalanced force.

Jackman, Robert H
Boston University

http://hdl.handle.net/2144/10665
Boston University
THESIS

DETERMINATION FOR THE MODAL AGE LEVEL FOR THE GRADES IV AND VI OF THE DIFFICULTY OF THE PRINCIPLE: A BODY AT REST REMAINS AT REST, AND A BODY IN MOTION CONTINUES TO MOVE AT CONSTANT SPEED ALONG A STRAIGHT LINE, UNLESS THE BODY IS ACTED UPON IN EITHER CASE BY AN UNBALANCED FORCE.

Submitted by

Robert H. Jackman

(B.S., Northeastern University, 1951)

In partial fulfillment of the requirements for the Degree of Master of Education

1952
First Reader:  John G. Read, A.M., Ed.D.
Professor of Education

Second Reader:  Walter N. Durost, A.M., Ph.D.
Associate 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>
<tr>
<td>III. EXPERIMENTAL PROCEDURE</td>
<td>Page</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Schools Visited</td>
<td>76</td>
</tr>
<tr>
<td>Town A</td>
<td>77</td>
</tr>
<tr>
<td>Geography</td>
<td>77</td>
</tr>
<tr>
<td>Sociological information</td>
<td>77</td>
</tr>
<tr>
<td>Town B</td>
<td>77</td>
</tr>
<tr>
<td>Geography</td>
<td>77</td>
</tr>
<tr>
<td>Sociological information</td>
<td>78</td>
</tr>
<tr>
<td>Testing Procedure</td>
<td>78</td>
</tr>
<tr>
<td>Control groups</td>
<td>81</td>
</tr>
<tr>
<td>Principle used in the experiment</td>
<td>81</td>
</tr>
<tr>
<td>Facets of the demonstration</td>
<td>81</td>
</tr>
<tr>
<td>Interpretation of Comparisons</td>
<td>83</td>
</tr>
<tr>
<td>The Otis Normal Percentile Chart</td>
<td>83</td>
</tr>
<tr>
<td>Advantages of using the Normal Percentile Chart</td>
<td>83</td>
</tr>
<tr>
<td>Modifications in the use of the Otis Normal Percentile Chart</td>
<td>83</td>
</tr>
<tr>
<td>Code used in reading the charts</td>
<td>84</td>
</tr>
<tr>
<td>Steps taken in drawing the percentile curves</td>
<td>84</td>
</tr>
<tr>
<td>Normal Percentile Chart Number 1</td>
<td>87</td>
</tr>
<tr>
<td>Normal Percentile Chart Number 2</td>
<td>89</td>
</tr>
<tr>
<td>Normal Percentile Chart Number 3</td>
<td>91</td>
</tr>
<tr>
<td>Normal Percentile Chart Number 4</td>
<td>93</td>
</tr>
</tbody>
</table>
### III. (continued)

<table>
<thead>
<tr>
<th>Normal Percentile Chart Number</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Percentile Chart Number 5</td>
<td>95</td>
</tr>
<tr>
<td>Normal Percentile Chart Number 6</td>
<td>97</td>
</tr>
<tr>
<td>Normal Percentile Chart Number 7</td>
<td>99</td>
</tr>
<tr>
<td>Normal Percentile Chart Number 8</td>
<td>101</td>
</tr>
</tbody>
</table>

### IV. RECAPITULATION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conclusions</td>
<td>102</td>
</tr>
<tr>
<td>Implications</td>
<td>103</td>
</tr>
<tr>
<td>Future of the Study</td>
<td>103</td>
</tr>
</tbody>
</table>

### APPENDIX

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specimen Test</td>
<td>104</td>
</tr>
<tr>
<td>Specimen Answer Sheet</td>
<td>108</td>
</tr>
<tr>
<td>The Placement of the Demonstration</td>
<td>109</td>
</tr>
<tr>
<td>Script of the Demonstration</td>
<td>110</td>
</tr>
</tbody>
</table>

### BIBLIOGRAPHY

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Books used in the Preparation of the</td>
<td>121</td>
</tr>
<tr>
<td>Demonstration and Test Items</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                                                                 Page
I.  The Frequency of Mention of the Selected Criteria by the Sources Consulted-- 64
II. The Distribution of the Pupils by School and Grade------------------------ 76
III. Standard Testing Procedure--------------------------------------------- 80
IV. Normal Percentile Chart Number 1---------------------------------------- 86
V.  Normal Percentile Chart Number 2---------------------------------------- 88
VI. Normal Percentile Chart Number 3---------------------------------------- 90
VII. Normal Percentile Chart Number 4--------------------------------------- 92
VIII. Normal Percentile Chart Number 5-------------------------------------- 94
IX. Normal Percentile Chart Number 6---------------------------------------- 96
X.  Normal Percentile Chart Number 7---------------------------------------- 98
XI. Normal Percentile Chart Number 8---------------------------------------- 100
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 experimentor
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 Kancevitch
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\textsuperscript{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".\textsuperscript{2}

From a Progress Report of the Committee on Research in Elementary Science for the National Association for Research in Science Teaching, Venill\textsuperscript{3} believes that with

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

\textsuperscript{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.
Washburne¹/ 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 Pistor⁴/ 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.

Deutsche1/ 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.
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. West\(^2\) 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 experiental levels are adequate for learning, lack of interest or proper attitudes, inadequacy of teaching method and materials may account for unprofitable learning.

\(^1\) John Dewey, Experience and Education, The Macmillan Co., New York, 1938, p. 27.

\(^2\) 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 Robertson1/ 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

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. 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, 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
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."


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.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.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


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, Mack1/ 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 Englehart2/ 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."\(^1\) 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\(^2\) 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"\(^3\) 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\(^4\)...

\(^1\) John S. Richardson and G. P. Cahoon, *op. cit.*, p. 67.


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 Anibel\(^1\), Knox\(^2\), and Wiley\(^3\) 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. Anibel\(^4\), 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\(^1\), 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 - Anibet.

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.

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 outcome, 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\(^1\). 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.  

In the study done by Smith 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 Croxton²/ 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 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 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 has made a plea for clear definition of teaching methods and experimental procedures. Mack has stated that there are as many definitions of "demonstration" as there are authors treating the subject. Preston 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.


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\(^1\) 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;\(^2\)

"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:\(^3\) "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\(^4\) to exclude from the demonstration procedure certain standard teaching materials:

\(^1\) Carleton E. Preston, "Is the Debate in Common Terms?" Science Education (February, 1935), 19:14-16.
\(^4\) 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, Mack2/ 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 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 1/ Op. Cit., pp. 19-31.

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\(^1\) offers three cautions to be observed in class demonstrations. Hoff\(^2\) emphasizes only visibility and planning. Pinkus\(^3\) 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,\(^4\) for example, offers several suggestions for performing demonstrations effectively and contributes many excellent ideas. In discussing the art of lecture table demonstration, Davison\(^5\) mentions several rules to follow in demonstrating. Rakestraw\(^6\) touches on six different aspects of the good demonstration in his extensive discussion of lecture-demonstration.

---


Still others in the "general" group have systematically attempted to list criteria in some form. Billinger\(^1\) lists five requirements for a successful demonstration. Dale\(^2\) offers fourteen suggestions for improving demonstrations and eleven questions for evaluating them. Under "demonstration techniques", Haas\(^3\) lists ten steps to be completed before conducting the experiment and five suggestions for conducting it. Heiss\(^4\) elaborates on seven excellent rules for demonstrating. Holley\(^5\) lists seven things a teacher can do to insure successful demonstrations. Mack\(^6\) developed a lengthy checklist of desirable qualities in demonstration apparatus. In a group thesis edited by Murray\(^7\), 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 emphasizes action as the essential quality of good demonstrations and includes, as he elaborates this theme, many other criteria.

Duff 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 presented a lengthy discussion on visibility including many excellent suggestions. Reed discusses in some detail four aspects of good demonstrations and techniques. Wiles also deals only with a few aspects of successful demonstrations.

Dunbar lists eleven desirable characteristics in demonstrations. His list is based on Duff's and includes

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


5/ L. A. Wiles, "The Value of Lecture Table Demonstrations in the Teaching of Chemistry", 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-


bitrary. 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\(^1\); 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.

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

| Frequency... | 11 | 7 | 25 | 19 | 22 | 14 | 4 | 7 | 0 | 6 | 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.\footnote{1} 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."\footnote{2} 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\footnote{3} 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


\footnote{2}{K. W. Vaughn and E. F. Lindquist (Editor), Educational Measurement, George Banta Publishing Co., Menasha, Wisconsin, 1951, p. 174.}

\footnote{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 \(\text{multiple choice items}^3/\) 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,¹ 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."

In this respect, Brownell 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." 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—I the demonstration period. 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." 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 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. Lindquist¹ concurs in stating that "The procedures of testing 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

¹ 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. Schools Visited

Three different schools were visited in two towns to obtain sufficient data for this research. Town A had two grammar schools that were used by the experimenter and Town B had one grammar school with two classrooms for each grade. All the fourth and sixth grade pupils in each of these schools were used in the research. There were 105 fourth grade pupils and 109 sixth grade pupils. Table II illustrates the manner in which the pupils were distributed.

Table II. The Distribution of the Pupils by School and Grade.

<table>
<thead>
<tr>
<th>School</th>
<th>Grade IV</th>
<th>Grade VI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>18</td>
<td>20</td>
<td>38</td>
</tr>
<tr>
<td>A</td>
<td>30</td>
<td>29</td>
<td>59</td>
</tr>
<tr>
<td>B</td>
<td>27</td>
<td>30</td>
<td>57</td>
</tr>
<tr>
<td>B</td>
<td>30</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Totals</td>
<td>105</td>
<td>109</td>
<td>214</td>
</tr>
</tbody>
</table>
Town A

**Geography.** -- This town is located approximately 25 miles west of Boston. It has a population of about 3000. The town has no industries and is classified as rural and residential. The town has few stores, all shopping being done in neighboring towns and cities. A few of the inhabitants are farmers, with the most of the population being employed outside of the town.

**Sociological information.** -- Approximately one third of the population is of Italian extraction with the remaining portion of heterogeneous extraction. Most of the homes are single family dwellings ranging in valuation from $3,000.00 to $40,000.00.

Social life is centered within church clubs, national, and local organizations. Boy Scouts and Girl Scouts are fairly active in the community. The children also have a playground project in the summer which includes swimming, arts and crafts, and competitive sports.

Town B

**Geography.** -- Town B lies on the northwest border of Town A about 30 miles from Boston. The population in Town B is 3500. There are a few industries, however, the majority of the workers are employed in neighboring towns and cities. The shopping facilities are somewhat better than Town A.
Sociological information. -- There is no racial predominance within the town. The social life of Town B is much the same as Town A having national, local, and church organizations as well as a sound Scouting program.

2. Testing Procedure

The administration of the test and the demonstration involving the principle were carried out by the investigator, thus assuring similarity in each situation involving testing and demonstrating. The teachers in both towns were very cooperative and felt the children were benefited by the complete procedure rather than wasting of their classroom time. Upon later visits to these schools for the purpose of giving the Otis results to school administrative officials, many comments were received that indicated the children were anxious to have more science demonstrations and tests.

The first visit to each school by the experimenter was spent in giving the Beta Test: Form CM of the Otis Quick-Scoring Mental Ability Test. At the conclusion of this test the students were informed that at a later date they were to have more tests which would be accompanied by a science demonstration. The pupils accepted the experimenter as a special instructor and throughout the experiment gave their full cooperation and attention to the proceedings. Likewise, the teacher, after having
introduced the experimenter to the class, retired to the back of the room.

On the second visit to each class the teacher again turned the room over to the investigator. The pupils were told that they were going to have a test, followed by a demonstration, which would be followed by a second test, all dealing with science. The students were instructed to perform everything that the experimenter asked and were specifically told to do no talking once the first test had started until the second test was completed. The results were very pleasing, for the pupils in both schools refrained from talking to each other all through the experiment even while they were moving from room to room.

After a brief introduction to the class the answer sheets were passed out and the students were instructed in their use. The experimenter found it best to have the pupils insert only their names on the answer sheet, the remaining information could be more reliably obtained from the teacher which would also mean a saving of time. The test booklets were then distributed and the students began answering the questions immediately. Copies of both the test and answer sheet may be found in the appendix.

Following the completion of the first test each class was divided by random numbers, the answer sheets having been previously numbered. This random selection was made in accordance with the description in the second chapter.
The pupils with the numbers the experimenter called out were moved from the present classroom to another. The experimental group from the sixth grade remained in the sixth grade classroom where they were shortly joined by the experimental group from the fourth grade. The sixth grade control group went to the fourth grade classroom where the fourth grade control group had remained.

A standard method of administering the pre-test, demonstration, and post-test was used which is outlined in Table III. In order to assure more similar conditions in the experimental group, a tape recording of the demonstration script was made and used each time.

Table III. Standard Testing Procedure

<table>
<thead>
<tr>
<th>Pre-test Period</th>
<th>Complete fourth grade in their own classroom</th>
<th>Complete sixth grade in their own classroom</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 minutes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Demonstration Period

| 20 minutes | Control Group of 50% of fourth and 50% of sixth grade in the fourth grade classroom (silent reading period) | Experimental Group of 50% of fourth and 50% of sixth grade in the sixth grade classroom having the demonstration |

Post-test Period

| 20 minutes | Control Group of 50% of each grade in the fourth grade classroom | Experimental Group of 50% of each grade in the sixth grade classroom |
Control Group. -- The Control Group referred to in Table III was set up to measure the effect of incidental learning which might take place between the pre-test and post-test. Those children who represented the control group were not allowed to see the demonstration during the demonstration period. They were allowed to read old editions of the Readers' Digest which were supplied by the experimenter. This 20 minute period of reading represented time filled with unrelated material to the testing situation. This group was not denied the privilege of seeing the demonstration but rather the demonstration was delayed until the post-test had been given.

Principle Used in the Experiment. -- A body at rest remains at rest, and a body in motion continues to move at constant speed along a straight line, unless the body is acted upon in either case by an unbalanced force.

Facets of the Demonstration. -- A copy of the script and a list of materials used in the demonstration are found in the appendix. The following is a listing of the individual parts of the demonstration.

1. An ordinary drinking glass was placed on a table to illustrate that it would not move unless a force caused it to move.

2. A top was set spinning to demonstrate that a body in motion tends to remain in motion.
3. A weight was placed on the end of a piece of paper and then the paper was given a quick pull, which meant the weight would not move. The weight was again placed on top of a piece of paper and the paper was pulled slowly showing that the weight moved with the paper.

4. To explain that the more weight an object has the greater its inertia, the weight was again used in comparison with a desk in the classroom. The weight was more easily pushed than the desk because of its relatively small weight.

5. As a second demonstration of weight being connected with inertia, a light ball and then a heavier ball were rolled down an inclined board (with small guides on the edges). The pupils observed the distance each ball pushed a wooden object at the end of the incline. The heavier ball having more inertia pushed the wooden object further upon striking it.

6. Another sequence used to show that a body at rest tends to remain at rest or if in motion tends to remain in motion was to suspend a weight by a thread with a second thread attached beneath the weight. First the lower thread was given a quick pull, thus showing the students that the lower thread broke. The demonstration was set up again and this time the lower thread was pulled slowly, the upper thread broke with the pull.
3. Interpretations of Comparisons

The Otis Normal Percentile Chart. -- The purposes of the Normal Percentile Chart are twofold: First, to accomplish all the purposes of graphic representation and interpretation of the scores of a group, and second, to do so in the simplest and easiest manner.

Advantages of Using the Normal Percentile Chart. --

First, the percentile scales at the top and bottom of the chart are so constructed that when a normal distribution is plotted the result will be a straight line. If the plotted points appear to lie in an approximately straight line, the distribution is presumably normal and it is merely necessary to fit the best straight line to the plotted points.

A second advantage is the fact that the percentile scores can be read as accurately at the extremes as at the middle of the range.

A third advantage is that the scores of two tests can be compared and read directly from the "percentile curves".

Modifications in the Use of the Otis Normal Percentile Chart. -- The experimenter compared four tests on each chart. This was a result of the nature of the data desired on the fourth and sixth grades in both the experimental and control groups. Information was desired on modal age groups, underage and overage groups.
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
- *4* means fourth grade
- *6* means sixth grade
- *X* means the experimental group
- *C* means the control group
- *Sub* means the underage (mental age) group of deviates
- *Subo* means the overage (mental age) group of deviates
- *Red Lines* report pre-tests
- *Blue Lines* report post-test
- *Solid Lines* denote experimental
- *Dotted Lines* denote control groups

All of the charts deal with the pooled schools, i.e., all the fourth grades are considered as one fourth grade and all the sixth grades are considered as one sixth grade.

Steps taken in Drawing the Percentile Curves. -- Steps were taken to (1) distribute the scores, (2) find the subtotals -- number of cases to and including those in each interval of score, (3) reducing the subtotals to percentiles, (4) locate the points on the chart representing these percentiles, and (5) fit the best straight line to the plotted points.
The following pages contain the comparisons made in this study. The comparisons are shown on Otis Normal Percentile Charts. A work sheet accompanies each chart and following each chart is an explanation and summary of the chart findings.
NORMAL PERCENTILE CHART

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

Grade 4 Experimental and Control Pre-test
Grade 6 Experimental and Control Pre-test

Variable I

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Score intervals</th>
<th>Frequencies</th>
<th>Sub-totals</th>
<th>Sub-totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.1</td>
<td>20</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>.2</td>
<td>19</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.3</td>
<td>18</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>.4</td>
<td>16</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>.5</td>
<td>14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>.6</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>.7</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>.8</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>.9</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

PERCENTILE SCALE

Median

Standard Deviation Scale

The blue point to point line gives a more true representation of the data than the straight line.

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.
Worksheet for Otis Normal Percentile Chart Number 1.

<table>
<thead>
<tr>
<th>Score</th>
<th>4th Grade</th>
<th>6th Grade</th>
<th>4th Grade</th>
<th>6th Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f subtotal</td>
<td>% subtotal</td>
<td>f subtotal</td>
<td>% subtotal</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>30</td>
<td>96.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>96.8</td>
<td>15</td>
<td>100</td>
</tr>
<tr>
<td>17</td>
<td>28</td>
<td>90.3</td>
<td>14</td>
<td>93.3</td>
</tr>
<tr>
<td>13</td>
<td>28</td>
<td>90.3</td>
<td>13</td>
<td>86.6</td>
</tr>
<tr>
<td>12</td>
<td>26</td>
<td>83.9</td>
<td>12</td>
<td>80.0</td>
</tr>
<tr>
<td>11</td>
<td>22</td>
<td>95.6</td>
<td>11</td>
<td>73.3</td>
</tr>
<tr>
<td>10</td>
<td>21</td>
<td>91.3</td>
<td>10</td>
<td>53.3</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>78.2</td>
<td>9</td>
<td>53.3</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>73.9</td>
<td>8</td>
<td>20.0</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>65.2</td>
<td>7</td>
<td>20.0</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>52.1</td>
<td>6</td>
<td>6.66</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>47.8</td>
<td>5</td>
<td>16.0</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>13.0</td>
<td>4</td>
<td>8.00</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td></td>
<td>3</td>
<td>4.00</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Summary of Findings

This chart was constructed to show the effectiveness of the random division within the groups chosen. The control groups of both the fourth grade and sixth grade received higher scores on the pre-test than the experimental groups due to a difference in ability between the two groups. The small sample of pupils lessened the effectiveness of the random division.
<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>School</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Grade 4 Experimental and Control Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>Grade 6 Experimental and Control Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PERCENTILE SCALE**

- The blue point to point line gives a more true representation of the data than the straight line.

<table>
<thead>
<tr>
<th>Variable I</th>
<th>Freq</th>
<th>Subtotals</th>
<th>Per</th>
<th>Score</th>
<th>Subtotals</th>
<th>Per</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</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>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</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>
</tr>
<tr>
<td>7</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>
</tr>
<tr>
<td>5</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>
</tr>
<tr>
<td>3</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>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable II</th>
<th>Freq</th>
<th>Subtotals</th>
<th>Per</th>
<th>Score</th>
<th>Subtotals</th>
<th>Per</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Median</th>
<th>Freq</th>
<th>Subtotals</th>
<th>Per</th>
<th>Score</th>
<th>Subtotals</th>
<th>Per</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Standard Deviation Scale**

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.
Worksheet for Otis Normal Percentile Chart Number 2.

<table>
<thead>
<tr>
<th>Score</th>
<th>4th Grade</th>
<th>6th Grade</th>
<th>Control Group</th>
<th>4th Grade</th>
<th>6th Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f sub-total</td>
<td>%</td>
<td>f sub-total</td>
<td>%</td>
<td>f sub-total</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>31</td>
<td>100.</td>
<td>2</td>
<td>31</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>29</td>
<td>93.5</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>27</td>
<td>87.1</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>26</td>
<td>83.9</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>23</td>
<td>100.</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>22</td>
<td>95.6</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>16</td>
<td>70.0</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>22</td>
<td>95.6</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>16</td>
<td>65.2</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>14</td>
<td>60.9</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>14</td>
<td>30.4</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>10</td>
<td>26.1</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>7</td>
<td>21.7</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>6</td>
<td>13.0</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>5</td>
<td>13.0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>13.0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
<td>8.69</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>8.69</td>
<td>1</td>
<td>1</td>
<td>4.00</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>8.69</td>
<td>2</td>
<td>2</td>
<td>8.69</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8.69</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Summary of Findings

This chart compares the achievement between the experimental and control groups of the two grades. The following facts are significant:

1. The experimental group of grade six showed a definite gain over the control group.

2. The fourth grade experimental group showed a slight increase over the control group.

3. Most of the gain in total score of the experimental groups of both grades is probably attributable to having seen the demonstration. Some of the gain in total score is due to the difference in ability exhibited between the experimental groups in the pre-test.
<table>
<thead>
<tr>
<th>Score</th>
<th>4th Grade</th>
<th>6th Grade</th>
<th>4th Grade</th>
<th>6th Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f</td>
<td>% sub-total</td>
<td>f</td>
<td>% sub-total</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>100</td>
<td>2</td>
<td>100</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>100</td>
<td>2</td>
<td>100</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>96.8</td>
<td>2</td>
<td>96.8</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>90.3</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>93.5</td>
<td>1</td>
<td>93.5</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>87.1</td>
<td>2</td>
<td>87.1</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>83.9</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>87.1</td>
<td>3</td>
<td>87.1</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>70.0</td>
<td>1</td>
<td>70.0</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>87.1</td>
<td>1</td>
<td>87.1</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>83.9</td>
<td>2</td>
<td>83.9</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>96.8</td>
<td>1</td>
<td>96.8</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>90.3</td>
<td>2</td>
<td>90.3</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>84.2</td>
<td>4</td>
<td>84.2</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>78.2</td>
<td>1</td>
<td>78.2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>72.6</td>
<td>7</td>
<td>72.6</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>65.2</td>
<td>4</td>
<td>65.2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>58.1</td>
<td>3</td>
<td>58.1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>52.1</td>
<td>2</td>
<td>52.1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>47.8</td>
<td>1</td>
<td>47.8</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>41.9</td>
<td>1</td>
<td>41.9</td>
</tr>
</tbody>
</table>
NORMAL PERCENTILE CHART

Grade or group | No. of cases | Measure (Examination) | Form | Date | Examiner | Graphs by | School | City
--- | --- | --- | --- | --- | --- | --- | --- | ---
I | | Grade 4 Experimental Group Pre-test and Post-test | | | | | | |
II | | Grade 6 Experimental Group Pre-test and Post-test | | | | | | |

PERCENTILE SCALE

Variable I | Score | Sub-intervals | Per cents | Total cents | Variable II | Score | Sub-intervals | Per cents | Total cents
--- | --- | --- | --- | --- | --- | --- | --- | --- | ---
0 | 20 | 10 | 20 | 0 | 0 | 1 | 20 | 10 | 20 |
1 | 19 | 10 | 19 | 1 | 1 | 1 | 19 | 10 | 19 |
2 | 18 | 10 | 18 | 2 | 2 | 2 | 18 | 10 | 18 |
3 | 17 | 10 | 17 | 3 | 3 | 3 | 17 | 10 | 17 |
4 | 16 | 10 | 16 | 4 | 4 | 4 | 16 | 10 | 16 |
5 | 15 | 10 | 15 | 5 | 5 | 5 | 15 | 10 | 15 |
6 | 14 | 10 | 14 | 6 | 6 | 6 | 14 | 10 | 14 |
7 | 13 | 10 | 13 | 7 | 7 | 7 | 13 | 10 | 13 |
8 | 12 | 10 | 12 | 8 | 8 | 8 | 12 | 10 | 12 |
9 | 11 | 10 | 11 | 9 | 9 | 9 | 11 | 10 | 11 |
10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
11 | 9 | 10 | 9 | 11 | 11 | 11 | 9 | 10 | 9 |
12 | 8 | 10 | 8 | 12 | 12 | 12 | 8 | 10 | 8 |
13 | 7 | 10 | 7 | 13 | 13 | 13 | 7 | 10 | 7 |
14 | 6 | 10 | 6 | 14 | 14 | 14 | 6 | 10 | 6 |
15 | 5 | 10 | 5 | 15 | 15 | 15 | 5 | 10 | 5 |
16 | 4 | 10 | 4 | 16 | 16 | 16 | 4 | 10 | 4 |
17 | 3 | 10 | 3 | 17 | 17 | 17 | 3 | 10 | 3 |
18 | 2 | 10 | 2 | 18 | 18 | 18 | 2 | 10 | 2 |
19 | 1 | 10 | 1 | 19 | 19 | 19 | 1 | 10 | 1 |
20 | 0 | 10 | 0 | 20 | 20 | 20 | 0 | 10 | 0 |

Standard Deviation Scale

-3 \( \sigma \) | -2 \( \sigma \) | -1 \( \sigma \) | Median | +1 \( \sigma \) | +2 \( \sigma \) | +3 \( \sigma \)

* The blue point to point line gives a more true representation of the data than the straight line.
Summary of Findings

The purpose of this chart is to compare the gains between the pre-test and the post-test of the Experimental Groups of both grades. The "learnability" of the demonstration would be the vertical distance between the pre-test and post-test for any grade.

The following fact is significant:

The fourth grade gained as much as the sixth grade, although the total knowledge of the principle tested remained higher in the sixth grade.
Worksheet for Otis Normal Percentile Chart Number 4.

<table>
<thead>
<tr>
<th>Score</th>
<th>4th Grade Pre-Test</th>
<th>6th Grade Pre-Test</th>
<th>4th Grade Post-Test</th>
<th>6th Grade Post-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f sub-total</td>
<td>% sub-total</td>
<td>f sub-total</td>
<td>% sub-total</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>13</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>12</td>
<td>92.3</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>12</td>
<td>92.3</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>11</td>
<td>84.6</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>11</td>
<td>84.6</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>7</td>
<td>53.8</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>6</td>
<td>46.2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>15.4</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>7.69</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The blue point to point line gives a more true representation of the data than the straight line.
OTIS PERCENTILE CHART NUMBER IV

Summary of Findings

The purpose of this chart is to show the results of the underage deviates in the pre-test and post-test of each grade.

The following facts are significant:

1. This chart, with its resulting curves, cannot be considered very accurate for either grade because there were only thirteen pupils in the fourth grade and only five pupils in the sixth grade.

2. There was little actual gain between pre-test and post-test for the fourth grade.
Worksheet for Otis Normal Percentile Chart Number 5.

OVERAGE (MENTAL AGE) DEVIATES-EXPERIMENTAL GROUP

<table>
<thead>
<tr>
<th>Score</th>
<th>4th Grade</th>
<th>6th Grade</th>
<th>Post-Test</th>
<th>4th Grade</th>
<th>6th Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f subtotal</td>
<td>% subtotal</td>
<td>f subtotal</td>
<td>% subtotal</td>
<td>f subtotal</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>14</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>13</td>
<td>92.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>12</td>
<td>85.7</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>14</td>
<td>100</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>6</td>
<td>42.8</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>12</td>
<td>85.7</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>11</td>
<td>78.6</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>10</td>
<td>71.4</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
<tr>
<td>0</td>
<td>8</td>
<td>57.1</td>
<td>2</td>
<td>7</td>
<td>53.8</td>
</tr>
</tbody>
</table>
NORMAL PERCENTILE CHART

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

Grade 1  1  Underage Deviates -- Experimental Pre-test and Post-test
Grade 6  1  Underage Deviates -- Experimental Pre-test and Post-test

PERCENTILE SCALE

Score 1 2 3 4 5 10 20 30 40 50 60 70 80 90 95 96 97 98 99 .5 .6 .7 .8 .9

Percentile 20 19 18 17 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

Median

Standard Deviation Scale

-3σ -2σ -1σ 0σ +1σ +2σ +3σ

Pre-test Underage Experimental Group
Post-test Underage Experimental Group

* The blue point to point line gives a more true representation of the data than the straight line.
The purpose of this chart is to show the results of the overage deviates in the pre-test and the post-test in the experimental group of each grade.

The following facts are significant:

1. "Learnability" of each grade is similar.

2. Overage (mental age) children benefited by the demonstration method of teaching.
### 4TH GRADE EXPERIMENTAL GROUP - POST TEST

<table>
<thead>
<tr>
<th>Score</th>
<th>Modal Age Group</th>
<th>Mental Group</th>
<th>Deviate Under-Age Group</th>
<th>Deviate Over-Age Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>23</td>
<td>100</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>22</td>
<td>95.6</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>16</td>
<td>70.0</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>15</td>
<td>65.2</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>14</td>
<td>60.9</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>10</td>
<td>43.5</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>7</td>
<td>30.4</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>6</td>
<td>26.1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>5</td>
<td>21.7</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>13.0</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>2</td>
<td>8.69</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>2</td>
<td>8.69</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8.69</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
OTIS PERCENTILE CHART NUMBER VI

Summary of Findings

In this chart the fourth grade is broken into three subdivisions; post-tests of the modal age group, post-test of the underage deviates, and post-tests of the overage deviates.

The following facts are significant:

1. The overage deviates exhibit greater "learnability".
2. The underage deviates are significantly lower than the modal age group.
Worksheet for Otis Normal Percentile Chart Number 7.

<table>
<thead>
<tr>
<th>Score</th>
<th>f</th>
<th>subtotal</th>
<th>%</th>
<th>f</th>
<th>subtotal</th>
<th>%</th>
<th>f</th>
<th>subtotal</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1</td>
<td>20</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>19</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>18</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>17</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>16</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>15</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>14</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>13</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>12</td>
<td>100</td>
<td>3</td>
<td>100</td>
<td>100</td>
<td>3</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>11</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>10</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>9</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>8</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
<td>2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>7</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>6</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>5</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
### Normal Percentile Chart

#### Grade or group

<table>
<thead>
<tr>
<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>Grade 6</td>
<td>Experimental Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Variable I

<table>
<thead>
<tr>
<th>Score intervals</th>
<th>Freq. Sub-Percentiles</th>
<th>Score intervals</th>
<th>Freq. Sub-Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.0</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>13</td>
<td>0</td>
</tr>
</tbody>
</table>

#### Percentile Scale

The blue point to point line gives a more true representation of the data than the straight line.

#### Standard Deviation Scale

-3σ | -2σ | -1σ | M | 0 | +1σ | +2σ | +3σ
Summary of Findings

In this chart the sixth grade is broken into three subdivisions: post-tests of the modal age group, post-tests of the underage deviates, and post-tests of the overage deviates.

The following facts are significant:

1. The curves on this chart indicate that there is little gain in total score between the overage deviates and the modal age group.

2. The underage deviates are significantly lower than the modal age group.
Worksheet for Otis Normal Percentile Chart Number 8.

<table>
<thead>
<tr>
<th>Score</th>
<th>Pre-Test</th>
<th>Post-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4th Grade</td>
<td>6th Grade</td>
</tr>
<tr>
<td></td>
<td>f sub-total</td>
<td>%</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
NORMAL PERCENTILE CHART

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

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
<td>Grade 4 Control  Pre-test and Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td></td>
<td>Grade 6 Control  Pre-test and Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Variable I  Variable II
Freq-Subintervals  Percents  Score  Freq-Subintervals  Percents

Pre-test Control Group:
Post-test Control Group:

The blue point to point line gives a more true representation of the data than the straight line.

Standard Deviation Scale

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.
This chart shows the gain or loss made by the control groups between tests.

The following facts are significant:

1. The sixth grade scores were lower on the post-test.

2. The fourth grade did slightly better on the post-test.
CHAPTER IV
RECAPITULATION

1. Conclusions

1. The groups tested were fairly accurate random samples of the group as a whole.

2. On a test-retest basis, the score on the post-test was higher.

3. The post-test clearly established a growth in learning of the experimental group over the control group which is attributable to the demonstration.

4. Learning did take place as a result of the demonstration.

5. The upper grade knew more of the items in both the pre-test and post-test.

6. Both grades gained from the demonstration.

7. Underage (mental age) pupils scored significantly lower than modal age pupils on the pre-test and post-test.

8. Underage, bright children have probably not matured sufficiently to understand the Science Principle demonstrated.

9. Overage (mental age) deviates in each grade exhibited a higher total score than the modal age groups, however, the gain between the pre-test and post-test overage experimental groups is approximately the same as the gain of the modal age experimental groups.

-102-
2. Implications

1. It can not be ascertained at this point if either class had learned the principle because no criteria have been established to determine how much of an increase can be expected to indicate the learning of the principle.

2. Total scores and the increase in scores correlated much closer with mental ages. The fact that total scores and the increase in scores of the average deviate were higher than the scores of the modal age groups leads the experimenter to conclude that with this principle, and this test, the better scores will be achieved by the mentally older pupil, regardless of brightness.

3. Further study with this principle at higher grade levels will indicate the effect of previous science courses or maturation, since none of the classes tested were offered any science courses.

Future of the study -- This experiment will be repeated until 600 individual scores at each grade level are obtained. It is planned to have the demonstrations measured by: (a) a larger sample of fourth and sixth grade pupils, (b) a sample of pupils that have had previous science courses, and (c) a sample of seventh and eighth grade pupils.
APPENDIX
1. What happens to a coin resting on a cardboard on top of an empty drinking glass when the cardboard is quickly pulled sideways from over the glass?

(a) The coin stays on the cardboard.
(b) The coin flies off the cardboard.
(c) The coin drops in the glass.
(d) The coin stays in the same place.

2. If a pound weight is placed on one end of a piece of heavy paper on a table and the other end of the paper is pulled slowly, the pound weight

(a) stays on the piece of paper.
(b) tears the piece of paper.
(c) does not move with the piece of paper.
(d) stays in the same place.

3. When a train stops suddenly, the people who are riding in it

(a) are thrown back in their seats.
(b) are thrown toward the windows of the train.
(c) are thrown forward in their seats.
(d) are not thrown in any direction.

4. What happens if a heavy metal ball is suspended by a thread and a piece of thread attached below the ball is given a quick jerk downward?

(a) Both threads break.
(b) The thread holding the ball breaks.
(c) The thread attached below the ball breaks.
(d) Neither thread breaks.

5. When a train starts suddenly

(a) the passengers reach their destination earlier.
(b) the passengers do not notice that they have started to move.
(c) the passengers are thrown forward.
(d) the passengers are thrown backward.

6. If a pound weight is placed on one end of a piece of heavy paper on a table and the other end of the paper is pulled quickly, the pound weight

(a) stays on the piece of paper.
(b) tears the piece of paper.
(c) does not move.
(d) drops on the floor.
7. A round ball is placed on a card that is over the mouth of a bottle. What happens if the card is suddenly snapped sideways?

(a) The ball drops into the mouth of the bottle.
(b) The ball and card are knocked off the bottle.
(c) The bottle tips over.
(d) The ball drops beside the bottle.

8. What happens if a heavy metal ball is suspended by a thread and a piece of thread attached below the ball is pulled downward steadily?

(a) Both threads break.
(b) The thread holding the ball breaks.
(c) The thread attached below the ball breaks.
(d) Neither thread breaks.

9. The passengers in a train that is speeding around a sharp curve are thrown

(a) forward.
(b) backward.
(c) toward the inside of the curve.
(d) toward the outside of the curve.

10. A 3000 lb. truck is moving at the rate of 25 miles an hour; a 1000 lb. automobile is moving at the rate of 25 miles an hour.

(a) The truck will be harder to stop.
(b) The automobile will be harder to stop.
(c) It all depends on the driver.
(d) There is no difference in trying to stop either one.

11. An ocean liner cannot stop suddenly even when moving slowly because

(a) the water carries it forward.
(b) it's motors are not powerful enough.
(c) it has a large weight.
(d) it has great length.

12. If a piece of heavy paper is pulled quickly from beneath a glass half filled with water,

(a) the glass remains standing and no water is spilled.
(b) the glass tips over spilling the water.
(c) the glass will break, spilling the water.
(d) the paper tears leaving a piece under the cylinder.
13. A baseball pitcher's "wind-up" helps him to throw the ball
(a) with greater speed.
(b) with better accuracy.
(c) without looking at the batter.
(d) with more of a curve on it.

14. It is harder to stop a train going 50 miles an hour than it is to stop an automobile moving at the same speed, because
(a) the train is traveling on rails.
(b) the automobile has four wheel brakes.
(c) the train has greater weight than the automobile.
(d) the automobile has rubber tires.

15. A large and heavy ball possesses _______ inertia than a light ball
(a) less
(b) more
(c) the same

16. A clown, the fat man and the midget, jump into a deep snow bank
(a) The clown goes in the deepest.
(b) The fat man goes in the deepest.
(c) The midget goes in the deepest.
(d) They all go in the same depth.

17. To tighten the hammer head on the handle of a hammer, one would
(a) strike the hammer head a few sharp blows.
(b) strike the hand end of the hammer handle a few sharp blows.
(c) turn the hammer head a few times.
(d) strike the head end of the hammer handle a few sharp blows.

18. What happens to a boy on a swing when another boy stops the swing from going forward by pulling down on a rope that is tied to the seat?
(a) He falls backwards.
(b) He falls forward.
(c) He falls sideways.
(d) He does not fall backward, forward or sideways.
19. If a horizontal pull is exerted on a spring balance attached to a flat object, the force required to start the object sliding is \underline{\hspace{10cm}} the force needed to keep the object sliding.

(a) less than
(b) more than
(c) the same as

20. Three boys are coasting on a hill. Two boys go on one sled, one boy on another sled. Both sleds start from the same place on the hill, when the sleds stop

(a) both sleds have coasted the same distance.
(b) the sled with one boy on it has coasted farther than the sled with two boys.
(c) the sled with two boys on it has coasted farther than the sled with one boy.
ANSWER SHEET

Name________________________________________ Date of Birth__________ Age______

Sex (M or F) ________________________________

Name of School _____________________________ Town ____________________________

Have you had any previous courses in science? Yes No (Please circle)

List the science courses you have had below:

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

Instructions:

Please place the letter of the answer which you think is most nearly correct for each question in the test booklet opposite its number below. That is, if you think answer "a" is correct for question No. 1, put after 1. in the space provided (a). Do not mark the test booklet.

ANSWER EVERY QUESTION.

1. ( ) 11. ( ) 21. ( )
2. ( ) 12. ( ) 22. ( )
3. ( ) 13. ( ) 23. ( )
4. ( ) 14. ( ) 24. ( )
5. ( ) 15. ( ) 25. ( )
6. ( ) 16. ( ) 26. ( )
7. ( ) 17. ( ) 27. ( )
8. ( ) 18. ( ) 28. ( )
9. ( ) 19. ( )
10. ( ) 20. ( )
THE DEMONSTRATION EQUIPMENT

1. Ringstand with weight suspended by a thread with another thread attached beneath the weight.
2. Inclined plane with guides
3. A tennis ball (light ball), a baseball (heavy ball) and a light wooden object at the bottom of the inclined plane.
4. Ordinary drinking glass
5. Small Weight
6. Top
7. Tape recorder
Inertia Script

This is probably the first time many of you have heard a tape-recorder. Instead of talking today, I prepared this many days ago and for the remainder of the demonstration you will hear my voice coming from this machine.

For the next ten minutes I would like to have you watch me very closely and listen to what I have to say. Please do not raise your hand to ask questions and do not talk to anyone, just watch the demonstrations. After the demonstrations are over and you have answered the questions I will give you, I will answer your questions. Now we are ready to begin.

I am going to try and show you that all matter has no ability in itself to move or to stop moving once it has started to move. Matter is inert, which means that it is inactive or lazy. Matter we can say is anything that takes up room or occupies space. Books, baseballs, carts, water, milk, and gas that is burned in stoves are all forms of matter. Since all matter has the property of inertia, bodies in motion tend to remain in motion and bodies at rest tend to remain at rest. Matter requires a push or a pull to overcome the inertia, or laziness, of a body to set it in motion or stop it when it is in motion.

A body that is at rest cannot of itself start to move. For example, we have this glass here on the table, we all know it will not move unless something causes it
to move. On the other hand, a body that is in motion cannot stop moving, but tends to continue its motion until disturbed by some outside force. This top is set to spinning, it will continue to spin unless I touch it, or it bumps into something. If the top does not bump into anything it finally stops spinning because of a force we call friction. If we could do away with friction the top would spin forever. This tendency of all bodies that are at rest to remain at rest and of all bodies that are in motion to continue their motion is called inertia.

Other demonstrations you will see in the next few minutes are all going to be concerned with bodies having no ability in themselves to move, or if moving to stop moving.

We can show this tendency of all bodies that are at rest to stay at rest in many different ways. First of all we place this weight on the end of this piece of paper. Now, if the paper is pulled quickly, we notice that the weight remained stationary. A body at rest tends to remain at rest. The pull on the paper was too fast to overcome the inertia or laziness of the weight. If the paper is pulled slowly, the weight will move because the steady pull travels through the paper to the weight and the inertia of the weight has been overcome.
The more weight an object has, the greater is its inertia. We can easily push this weight with our hand, but we find it is much harder to move this desk, in fact the desk will not move. If we roll a small or light ball against this block, we notice it only moves a little. Now we will try this heavy ball, we see that the block has moved much farther. Thus we can say that the heavy ball has more inertia than the lighter ball. We learn from this demonstration that the more weight an object has, the more inertia it has.

Another way to show that a body at rest tends to remain at rest or if in motion tends to remain in motion is by hanging a weight by a thread and having another piece of thread hanging from the weight. First, the thread hanging from the weight is given a quick pull. The thread beneath the weight breaks because the force of the pull did not pass through into the weight, all the pull was taken by the lower thread, the weight remained at rest. Now after attaching another thread, I pull slowly but strongly, you see that the thread holding the weight breaks while the lower thread does not. With a strong and steady pull the inertia or laziness of the weight was overcome, and there was now the force of the pull on the upper thread plus the movement of the weight. A body in motion tends to remain in motion, the weight was put in motion by the steady pull, thus the upper thread broke because of the combined weight and strength of the pull.
That demonstration ends the talk on bodies in motion tending to remain in motion and bodies at rest tending to remain at rest.

Now please do not talk with anyone as I will now give you another test.
BIBLIOGRAPHY


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.


Books Used in the Preparation of the Demonstration and the Test Items.