1953

The grade placement of the physical science principle "rough dark surfaces absorb and radiate heat energy more readily than smooth light-colored surfaces" in relation to mental ages.

Cormier, Marcellinus, Brother

Boston University

http://hdl.handle.net/2144/11267

Boston University
BOSTON UNIVERSITY
GRADUATE SCHOOL

Thesis

THE GRADE PLACEMENT OF THE PHYSICAL SCIENCE PRINCIPLE "ROUGH DARK SURFACES ABSORB AND RADIATE HEAT ENERGY MORE READILY THAN SMOOTH LIGHT-COLORED SURFACES" IN RELATION TO MENTAL AGES

Submitted by

Brother Marcellinus Cormier, C. F. X.
(B. S., St. John's University, Brooklyn, 1948)

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

1953
First Reader: John G. Read, A.M., Ed.D.
Professor of Education

Second Reader: Roy O. Billett, A.M., Ph.D.
Professor of Education
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th></th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Justification</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Design of the Experiment</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Scope and Limitations</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Discussion of Procedure</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>II. SURVEY OF RELATED LITERATURE</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Teaching By the Use of Principles</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>The inductive method</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>The deductive method</td>
<td></td>
<td>26</td>
</tr>
<tr>
<td>Some results</td>
<td></td>
<td>28</td>
</tr>
<tr>
<td>The Lecture-Demonstration Method of Teaching</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>The Effectiveness of Lecture-Demonstrations</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>The Criteria for a Good Demonstration</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>Statement of the problem</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>Need for research</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>Definition of demonstration</td>
<td></td>
<td>43</td>
</tr>
<tr>
<td>Review of the literature</td>
<td></td>
<td>46</td>
</tr>
<tr>
<td>Use of criteria</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>Selected criteria</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>The Test Technique</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Structure of the Test</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Multiple choice items</td>
<td></td>
<td>56</td>
</tr>
<tr>
<td>Levels of difficulty</td>
<td></td>
<td>57</td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td>59</td>
</tr>
<tr>
<td>The test tryout</td>
<td></td>
<td>59</td>
</tr>
<tr>
<td>The test period</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>Aims and Use of the Test</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>Employing statistics</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>Test-retest method</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>What the test endeavors to determine</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>Characteristics of the Test</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>Validity</td>
<td></td>
<td>64</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>III. EXPERIMENTAL PROCEDURE</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Preparation of the Demonstration</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Designing the apparatus</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Operation of the apparatus</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Preparation of the Test</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Selecting the items</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Testing the items</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Schools Visited</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>The school system used for the survey</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Socio-economic conditions</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>The Testing Program</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Administering the pre-test</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Selecting the control group</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Giving the demonstration</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Administering the post-test</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Administration of mental tests</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>IV. STATISTICAL PROCEDURE AND RESULTS</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Preparing the Data for Statistical Research</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Grading the tests</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Compiling the data</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Performance of Different Mental Age Groups</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Establishing the medians</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>The Item Analysis</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Determination of the mid-80 percent</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Determination of the upper and lower 27 percent</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Determining the difficulty index</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Eliminating the test items</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>The Test of the Principle</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Re-grading the tests</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Pearson coefficient of self-correlation</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>The mastery level</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Proportion passing the test</td>
<td>81</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1. Introduction

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

This year both the background for the study and the design of the experiment are being done under the guidance of Dr. Read and Mr. Herbert Oxendine. Also, the first group will start the experiment with each experimenter taking one principle. The teaching method will be a lecture-demonstration. The amount of learning will be measured by identical tests given before and after the lecture-demonstration. Data to be gathered will consist of the test scores, the pupil's I.Q., M.A., C.A., sex, previous science instruction 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 indication 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 principles 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 \(^1\) states that because scientific knowledge is accumulating at such a rapid rate, there is neither time nor excuse for teaching the elementary scientific concepts in the higher grades. He points out that the scientific background and foundation prerequisite to an understanding of the individual science courses offered in the high schools are lacking in the beginning students. To find a solution to this problem, he suggests that research be started to determine, '...what fundamentals of science can we expect most children of similar ability and cultural background to master at each maturity level'. \(^2\)

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

---

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

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

\(^3\) John Venill, 'Needed Research Studies in the Junior High Schools', Science Education (April, 1948), 32:175-185.
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. Morrison3/ 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/

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-

1/ John Venill, op. cit., p. 175.
3/ Ibid., p. 354.
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 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. 3/ Kingsley defines maturation as "... the

1/ Arno A. Bellack, op. cit., p. 42.
2/ Ibid., p. 623.
3/ Ibid., p. 625.
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.\(^1\)

In order to obtain the maximum efficiency in learning, maturation of the child must be considered carefully. Hilbreth \(^2\) 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. Washburne\(^3\) 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."\(^4\)

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.\(^5\) But a number of studies have


\(^4\) Ibid., p. 3.

been made to determine the role that maturation plays in the development of various concepts, among them that of Pistor \(^1\) 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 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." \(^2\)

Piaget \(^3\) 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.


\(^2\) Frederick Pistor, op. cit. p. 111.

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

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

concluded that "Causal thinking apparently does not develop by stages but by a gradual process." 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. 2/

Haupt 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 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 these factors are extremely difficult to separate for study. In the human body every organ is an integrated

1/ Jean Deusch, op. cit., p. 93.
2/ Ibid., p. 29-42.
part of the whole body. If one organ is malfunctioning, it will affect the normal activity of the whole organism. Similarly, the child is a composite of many factors, each affecting the functioning of the other.

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

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


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.
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 those 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 eights- 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 class room easily.

B. Discussion of Procedure

The first step involved the selection by each investigator of a
principle to be demonstrated and tested. Such accepted lists as that
compiled by Robertson\textsuperscript{1} were consulted.

A review of the literature established that the teaching of prin-ciples
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

\textsuperscript{1} Martin L. Robertson, "Selection of Science Principles Suitable As
Goals of Instruction in the Elementary School", Science Education
Nichols\textsuperscript{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 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 be simple and the speed of action suitable.
6. The demonstration should work; it should be as infallible as possible.
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. Altogether five schools were selected and the pupils of two grade divisions in each school were chosen as subjects for the experiment.

1/ Edward L. Thorndike and Irving Lorge, *op. cit.*
In some cases, investigators chose the elementary grades and in others, the secondary.

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

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

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

The time allotted for the pre-test was approximately fifteen minutes. At the end of this time the answer sheets were collected and half of the pupils in the class were sent to another room, after handling 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 apparatus


\(^2\) Ibid., p. 26.
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 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 a 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 impli-
cation 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.1/

Jones, 2/ Leonelli, 3/ Martin 4/ 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." 5/

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: 1/

"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.  

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 the same schools. The differences however were not statistically significant. Kilgore paired 120 students in high school physics with respect to their previous experience in science courses 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.

The evidence from these studies seems to indicate that the learning of principles of science, and the ability to apply them, may be attainable

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


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, Mack, in describing and defining a demonstration, says in part:

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

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

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

---


Any discussion of the use of the demonstration in science teaching should be related to certain principles of learning. Potthoff 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. They stated that:

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

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

---


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

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

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

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

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

Before presenting a review of the research in which the lecture-demonstration is compared with other methods of science teaching, the evolution of the popularity of the demonstration method should be mentioned.


Henry S. Webb\(^1\) 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"\(^2\) 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\(^3\) 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, eleven 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 studies in seven of the enterprises. The majority of the pupils in three of the enterprises favored the


\(^{3}\) Ibid., p. 76
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 attach on new problems; scientific attitude; ability to 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.

1/ Op. cit., p. 76
2/ Ibid., p. 79
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 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 date.

For comparative purposes the writer has used the studies of Anibel, Knox, and Wiley 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, 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.


5/ Loc. cit.
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 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.

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.

\(^1\) Loc. cit.
Before stating the problem and significant conclusions of a study ranked inferior by Keiser according to the first six of the seven criteria developed by Stuit and Englehart, it should be recognized that the study made by Wiley 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; 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 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 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.

1/ Loc. cit.
3/ Loc. cit.
4/ Note the date of publication of the first group intelligence tests.
5/ Loc. cit.
6. The different methods show decided individual differences both for immediate and delayed reproduction.

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

Stuit and Englehart have also made an excellent critical analysis of the lecture demonstration versus the individual laboratory method of teaching high school chemistry. A summary of their report, which consisted of the combined conclusions of various investigators, is as follows:

Conclusions contending that the laboratory method is superior:

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

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

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

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

5. For providing knowledge and method of attack, the laboratory method is superior for the inferior pupil - Knox.

Conclusions claiming that the demonstration method is superior:

1. Bright pupils are more likely to profit by the lecture-demonstration method than are the others - Anibel.

2. Dull pupils profit more from demonstration than from individual laboratory work - Carpenter.

3. The lecture-demonstration takes less time and costs less - Anibel.

4. The teacher (Demonstration) method is best - Nash and Phillips.

5. Lecture-demonstration method gives better control over the individual since all are under teacher guidance - Pugh.

6. For purpose of providing knowledge for both immediate and permanent retention and for the purpose of providing technique or handling new problems, the demonstration method is much to be preferred to the laboratory method in case of average superior pupil - Knox.

Conclusions contending that the students achieves equally well by either method:

1. Immediate retention is about equal in both lecture-demonstration and individual-laboratory methods - Anibel.

2. There is not as great a difference as is ordinarily supposed in the value of the three methods, lecture, textbook - and laboratory, so far as imparting knowledge is concerned - Wiley.

3. The results of this experiment point to the conclusion that the majority of students in high-school, laboratory-chemistry classes, taught by the demonstration method, succeed as well as when they perform the experiment individually, if success is measured by instruments which measure the same abilities as are measured by these tests, namely, specific information and ability to think in terms of chemistry - Carpenter.

General conclusions based on evaluation of the reported research:

After considering the above conclusions the writers have arrived at a few ideas which seem justifiable in the light of the evidence given by this study.

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

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

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

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

Evidence of the evolution of teaching methods particularly by the visual method is apparent in the next study to be considered, that of
Smith. The visual method has been long recognized by leading educators as one of the most valuable ways of training pupils in all stages of learning. It is difficult, as a result, to find a school that does not, in one way or another, make use of visual aids in teaching. The alert teacher and administrator are constantly seeking suggestions and illustrations by means of which the vague conceptions of the pupils may be made into real facts and parts of their experience. Any method will not be overlooked if it can provide both clarity and simplicity combined.2/

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.

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.


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. His problem was stated thusly:

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

In this study most of the experiments tested the pupils' ability to formulate and apply a principle after eight minutes exposure to the essential experimental basis in the form of a demonstration or directed play. The tentative conclusions made by Croxton 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 satis-

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

2/ Ibid., p. 634
faction 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 little more than might reasonable 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\(^1\) has pointed out the fact that investigators seldom define teaching methods carefully and minutely enough. Various writers have recognized the need in research for accurate definition of terms. Like Noll, Riedel\(^2\) has made a plea for clear definition of teaching methods and experimental procedures. Mack\(^3\) has stated that there are as many definitions of "demonstration" as there are authors treating the subject. Preston\(^4\) also realized this and called for clarification and unification of terminology.

Most of the literature on the demonstration method fails to recognize the difference between the lecture-demonstration, the class experiment, and the illustrated lecture. Preston\(^5\) 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


\(^4\) Carleton E. Preston, "Is the Debate in Common Terms?" *Science Education* (February, 1935), 19:14-16.

\(^5\) Loc. cit.
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, in the lecture-demonstration the flow of information and explanation is from teacher to pupils; in the class experiment it is exactly the opposite."

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

Preston, however, does not distinguish the lecture-demonstration from

---


2/ Ibid., p. 238


the illustrated lecture as does the Encyclopedia of Modern Education: 1/

"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: 2/ "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 3/ to exclude from the demonstration procedure certain standard teaching materials:

"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, 4/ it would seem that this method should also be excluded on the same basis as the other visual aids.

Further, Mack 5/ 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.

**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) audiovisual texts and (5) books on experiments. The following reference sources were consulted: Bibliographic Index, Encyclopedia of Educational Research, Bibliographies and Summaries in Education, Reader's Guide, International Index, Ulvich's Periodical Directory, Vertical File Service, and the Education Index.

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\(^1\) indicates certain steps taken to insure effectiveness of the demonstration and emphasizes only visibility and size of apparatus. Colvin\(^2\) offers three cautions to be observed in class demonstrations. Hoff\(^3\) emphasizes only visibility and planning. Pinkus\(^4\) 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\(^5\) for example, offers several suggestions for performing demonstrations effectively and contributes many excellent ideas. In discussing the art of lecture table demonstration, Davison\(^6\) mentions several rules to follow in demonstrating. Rakestraw\(^7\) 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. Billinger¹ lists five require-
ments for a successful demonstration. Dale² offers fourteen suggestions
for improving demonstrations and eleven questions for evaluating them.
Under "demonstration techniques", Haas³ lists ten steps to be completed
before conducting the experiment and five suggestions for conducting it.
Heiss⁴ elaborates on seven excellent rules for demonstrating. Holley⁵
lists seven things a teacher can do to insure successful demonstrations.
Mack⁶ developed a lengthy checklist of desirable qualities in demonstra-
tion apparatus. In a group thesis edited by Murray⁷, five criteria for
a demonstration were listed which had been developed in a seminar

¹ R. D. Billinger, "Lecture Demonstration Experiments", Journal of
Chemical Education (August, 1937), 14:375-7.

² Edgar Dale, Audio-Visual Methods of Teaching, The Dryden Press,
New York, 1946, p. 125.130.

³ K. B. Haas, "The Demonstration and Field Trip as Training Techniques",
Business Education World (February, 1951), 31:291-293.

⁴ Elwood D. Heiss, Charles W. Hoffman, and Ellsworth S. Obourn, Modern
Methods and Materials for Teaching Science, The Macmillan Company, New

⁵ Charles Elmer Holley, High School Teachers Methods, The Garrard


⁷ Chalmers Murray (Editor), New and Improved Demonstrations, Each
Illustrating a Single Science Principle, Unpublished Master's Thesis,
Boston University, 1950.
discussion. Richardson and Cahoon list five criteria for a good demonstration. Selberg 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 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 stresses the value and importance of vertical mounting of apparatus on special boards. Also, Sutton stresses the need for simplicity and originality. Among his suggestions for improving physics teaching, Weaver 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 specific examples in chemistry.

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.


which he believes to be justified by the experience of a number of
teachers. Gould enumerates on eight to consider in planning and perform-
ing demonstrations. Van Horne offers five suggestions for the preparation
of apparatus and materials and four rules to follow in conducting demonstra-
tions.

Use of Criteria.-- The seminar in 1951 after careful consideration of
criteria of a good demonstration in this field and condensation and tele-
scoping of this material produced the criteria listed below. This group
feels it is of no additional value in going on any further with this
problem. It is an assumption these are good criteria.

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.

1/ Arthur B. Gould, "Demonstration Experiments and Their Place in the
   Teaching of Chemistry", Journal of Chemical Education (February, 1931),
   8:297-302.

2/ Donald Van Horne, "The Lecture Demonstration Method in High School
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.

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

*--These might well be separate criteria.
order on the table so that the action can proceed 
logically.

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 es-
tablishing their validity. The table shows the relative importance of the 
criteria as recognized by these authorities.
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 Ross¹, 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, Odell 3/ 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 adds 4/ "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 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. 5/ In order

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


2/ Ibid., p. 286
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

\(^1\) H. S. Conrad and E. F. Lindquist (Editor), *Educational Measurement*, George Banta Publishing Co., Menasha, Wisconsin, 1951, p. 250
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. Odell¹ has recommended that "on the average elementary-school pupils be expected to respond to three or four such exercises /multiple choice items/ per minute."

By allotting sufficient time for all examinees to attempt all the items, the influential factor of time itself is eliminated. As stated by Lindquist² "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."²

In this respect, Brownell³ 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 [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\(^1\) who states that "statistical reasoning is basic to all predictions".

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

\(^1\) Ibid., p. 176
the test items. The influential time factor has been kept at a minimum. Lindquist\(^1\) concurs in stating that "The procedures [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 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.

\(^1\) E. F. Lindquist (Editor), Educational Measurement, op.cit., p.617.
CHAPTER III
EXPERIMENTAL PROCEDURE

1. Preparation of the Demonstration

Designing the Apparatus.— In order to demonstrate the principle, "rough, dark surfaces absorb and radiate heat energy more readily than smooth, light-colored surfaces", the teaching of which is undertaken in this research, the writer designed a special apparatus which would satisfy all the criteria of a good demonstration. Since the experiment demands the use of a thermometer which can be seen easily by all the pupils in a normal classroom situation, a differential air thermometer is employed. This particular device adds the "dynamic" characteristic which helps focus the pupil's attention on the demonstration.

The display panel consists of two Florence flasks attached to separate differential air thermometers. One of these flasks is painted with an aluminum coating on both inner and outer surfaces. The other is covered with powdered charcoal which is retained on both inner and outer surfaces by a thin coating of shellac. These two bodies now represent the smooth, light-colored surface, and the rough dark surface of the principle, "rough, dark surfaces absorb and radiate heat energy more readily than smooth, light-colored surfaces". The two bodies are mounted on the panel so as to be equidistant from a radiant energy source, in this case a 75-watt lamp. The liquid used for the thermometer columns is ammoniated phenolphthalein solution of a bril-
liant red color. The panel was purposely made attractive to capture interest, and yet, made simple enough in design to avoid overwhelming the pupil with a maze of gadgetry. The mechanism of the differential air thermometer is hidden from sight in order that the pupil's attention may focus on the two flasks along with their rising columns.

**Operation of the Apparatus.**—When the lamp of the apparatus is turned on, the two columns of red liquid begin to rise almost immediately, but the column attached to the black globe ascends much faster and higher by a ratio of approximately four units to one. The principle of absorption by surfaces of varying natures is thus clearly shown. This phase of the operation may be regarded as the first facet of the demonstration. When the lamp is turned off, the columns begin to descend. Again, the column attached to the black, rough-surfaced globe descends faster by nearly the same ratio as that of ascent. This is the second facet of the demonstration illustrating the principle of radiation worded in the statement, "rough, dark surfaces absorb and radiate heat energy more readily than smooth, light-colored surfaces". The entire operation requires less than fifteen minutes to take place.

2. Preparation of the Test

**Selecting the Items.**—The test drafted for the purpose of discovering the index of learning ability of the single scientific principle, "rough, dark surfaces absorb and radiate heat energy more readily than smooth, light-colored surfaces" was designed to contain thirty items. Each of these items is constructed specifically to
determine whether or not the pupil has sufficient understanding of the principle to apply it in a particular situation. That is, whether or not he is able to perceive the cause of a natural phenomenon as a direct application of this principle. Some of the items selected for this purpose were adapted from samples found in standard physics textbooks presently used in secondary schools. The larger number of items were the product of the writer's own experience with the principle. The test was organized according to the following plan: (1) items which are easy, local applications (within the pupil's environment) of both radiation and absorption of heat energy by rough and dark surfaces, (2) items which are easy, local applications of both radiation and absorption of heat energy by smooth and light-colored surfaces, (3) items which are difficult, local applications of radiation and absorption of heat energy by rough and dark surfaces, (4) items which are difficult, local applications of both radiation and absorption of heat energy by smooth and light-colored surfaces, (5) items which are remote applications of the principle, that is, not within the pupil's environmental experience, (6) items which are directly connected with demonstration of the principle from the panel.

In order to eliminate the possibility of penalizing the pupil of the control group who does not see the demonstration before taking the post-test, the items referring directly to the experiment clearly describe the experimental situation. Thus, if the principle is already a part of the pupil's acquired knowledge, he is able to deduce a correct
answer from this description alone.

Testing the Items.--- Once the items were organized and composed with their multiple answers, they were submitted to a group of seminar students engaged in a similar research project. These people reviewed every item of the test critically and suggested necessary changes in both wording and form. Only then were the items chosen at random to form the order in which they appear in the final format of the test. The order of the multiple answers was also determined by random selection, that is, which answer should serve as the first, second, third, fourth, or fifth answer in each item.

In wording the test items it was found necessary to use scientific terms which would be ordinarily unfamiliar to the elementary grade levels chosen for the survey. To overcome this difficulty, the writer prepared a special alphabetized list of the difficult words accompanied with their simple definitions, a sample of which appears in the appendix. This list was printed on a single sheet of paper to facilitate handling and was given out to each pupil along with the test sheet. Oral instructions were given concerning the use of this vocabulary list if the meaning of any word was not clear. During the actual testing program, the writer noted with much satisfaction that the pupils of both the sixth and eighth grades where the survey was conducted made excellent use of the vocabulary list with apparently good results.

3. The Schools Visited

The school system used for the survey.--- The school system
selected and used for this survey was the Boston Archdiocesan Parochial School system. The principals of five different schools of this system were contacted for an interview date. The schools chosen will not be referred to by name at the request of the principals involved, but will be designated as School A, School B, School C, School D, and School E. The areas where these schools are located are Roxbury, South Boston, Jamaica Plain, Somerville, and Brockton. During the interview with the principal a detailed explanation of the survey and its purpose was given, and arrangements were made to have the use of two separate rooms to conduct the experiment. The principal was further requested to have the sixth and eighth grade teachers select at random twenty of their pupils, and to have them report to the room assigned at the proper time. The date and time for the conduct of the experiment were then decided upon, and the teachers of the grades involved were informed about the nature of the program to be carried on that day. In all five schools surveyed, the classes from which the pupils were selected were sufficiently large so that the absence of the testees did not interfere materially with the normal class procedure. For this reason, no teacher was available for prefecting the control group outside the classroom used for the demonstration. The writer provided his own prefect in the person of a responsible senior high-school student who was instructed on how to preside over the silent reading session of the control population.

**Socio-economic conditions.** -- Each of the schools used in this survey represents a typical urban population of school children. The general socio-economic background of the majority of these
children is that of the average working middle class found in any large city of Eastern United States. It is difficult to make a distinct division here since in any individual classroom of any one of these schools the population is heterogenous, socio-economically speaking. Some of the pupils' families are those of the better-than-average income group, while others are from families representing the people whose income is slightly below the average. From school records and the report of the teachers of these boys and girls, the writer found that some are the children of men in the professional ranks; others are the children of those in the skilled and unskilled laboring classes. It is perhaps safe to say that a median could be established at the borderline of the average and slightly below average income group. The writer found the majority of the pupils involved in this survey to be neatly dressed and well-behaved children reflecting the excellent training both at home and in the school they attend. Certainly, they were most anxious to cooperate in carrying out the assignment and in following instructions to the letter.

4. The Testing Program

Administering the pre-test.--- By previous arrangement, five dates were set within a two-week period in May, 1953, for the actual testing program. In each school, the twenty selected pupils from the sixth grade and the twenty from the eighth grade were assembled in one classroom. After a brief informal introduction designed to make the children at ease and cooperative, a copy of the test together with the vocabulary list was handed to each pupil. The
writer himself directed the filling out of the heading of the test to
insure accuracy of the information required. The pupils were then
reassured that the test was not part of their regular school work;
that it would not be rated for grades and submitted to their teacher.
It was also emphasized that the test results were not meant to deter-
mine how much science they already knew, but instead to discover how
much they could learn from a simple demonstration. Further instructions
were given to explain how each item should be answered by placing a
check mark in the bracket before the answer chosen. A twenty-minute
time allotment was set and the test was underway.

Selecting the control group.--- When the pretest time limit
had elapsed, all papers were collected regardless of the fact that
some papers were incompletely answered. The writer then made a rand-
on selection of ten sixth graders and ten eighth graders. These tes-
tees were told to report to the vacant classroom reserved nearby.
In this room a proctor was on hand to distribute non-scientific read-
ing material for silent reading. These pupils, who will be referred
to hereafter as the control group, were asked to refrain from talking
to each other while leaving the room and during the silent reading
session. The pupils remaining after the control group had been selec-
ted were now ready to see the demonstration. These members of the
survey will be referred to from here on as the experimental group.

Giving the demonstration.--- The demonstration apparatus was
now taken from its packing case and set up for all to see clearly.
A brief lecture, a copy of which is contained in the appendix,
explained the apparatus and pointed out the phases of radiation and absorption as they occurred during the operation. The principle was stated at the beginning and again at the end of the lecture to emphasize the points illustrated by the lesson. When the lecture was over, the panel was restored to its packing case and the control group was recalled to take the post-test along with the experimental group.

**Administering the post-test.**—The entire population now reassembled, a new set of test sheets was distributed to each pupil. The heading of the test papers was again filled out completely and the post-test was in progress. The time limit was the same as for the pre-test, namely, twenty minutes. When this time was up, all tests were collected and assorted according to grades and separated into experimental and control sections. The panel demonstration was again set up for operation for the benefit of the control pupils. An informal session followed during which questions were answered concerning the operation of the apparatus. Where time permitted, the test items were reviewed at the request of the pupils and the correct answers revealed. This session proved to be a most interesting phase of the experiment. The children, unhampered by the rigid regulations of the research procedure, were eager to ask all sorts of questions about the test items, the demonstration, and even on scientific problems totally unrelated to the heat phenomena observed. Regretfully, this session had to be cut short in all schools visited, since a time allotment had to be observed. This fact, however, served to convince the writer that youngsters of the age level surveyed are keenly
interested in science education and are definitely capable and ready to receive it advantageously if properly presented.

The administration of mental tests.--- After each school had been visited, the writer left a set of Otis Quick-scoring Mental Ability Tests to each of the teachers of the survey pupils. Sets of both the Beta Test Form A and Beta Test Form B were used for the purpose of determining each child's mental age and IQ. Instructions were left to each teacher to administer these tests at some convenient time within the next two weeks, and to forward the papers to the writer for correction and grading. The teachers concerned were told that duplicate forms of the results of these tests would be sent to them for their own records. This measure of good will insured prompt service since these teachers were anxious to have the records for their personal files.
CHAPTER IV
STATISTICAL PROCEDURE AND RESULTS

1. Preparing the Data for Statistical Research

Grading the tests.— As soon as the testing program was completed, the work of correcting and grading the 200 test papers of the entire population was undertaken. The test contained thirty items; hence each test was scored numerically according to the number of items correctly answered. Items which were not answered were considered wrong. Once these scores were obtained, they were tabulated on data cards especially prepared for this survey. A data card for each pupil was marked to show which item was correctly answered. A red mark indicated an item correctly answered on the pre-test, and a blue mark if it was correct on the post-test.

Compiling the data.— A complete set of pertinent facts was recorded on the data cards mentioned above. This information consisted of the following data: (1) the pupil's name, birthdate, and chronological age, (2) the school, grade, and teacher of the pupil, (3) the Otis form used, the Otis score and I.Q., (4) the date on which the experiment was conducted along with the experimenter's name and the principle taught, (5) the pupil's mental age, (6) the items answered

1\A sample of the pupil's data card appears in the appendix on page 96.

-74-
correctly on pre-test and post-test respectively, (7) the pupil's sex and the group participated in, whether experimental or control. Once this information was completely entered for all pupils, the work of statistical research began.

2. Performance of Different Mental Age Groups

Establishing the medians.— In order to chart the performance of the different mental age levels on the post-test scores for all pupils, the data cards were divided into the two sections, the control and experimental groups. Each set of cards was then stacked in descending order of mental age levels and the post-test scores recorded on a chart. From this chart the medians for each age level were computed. Tables 1 and 2 below contain these medians together with the inter-quartile ranges of each mental age group for both the experimental and control populations. These tables should be examined together for comparison. It is immediately evident that the experimental population gained significantly over the control population after seeing the demonstration, and that the higher the mental age, the better was the performance in either group. The interquartile ranges, or the ranges of the middle 50-percent of the cases, give a measure of variability. In all cases recorded in tables 1 and 2, the range does not extend farther than two points above or below the median score for that particular mental age level.
Table 1. Mental Age Medians and Interquartile Ranges of the Post-test Scores for the Entire Control Population.

<table>
<thead>
<tr>
<th>Mental Age Level in Years</th>
<th>Number of Testees</th>
<th>Median of the Post-test Scores</th>
<th>Interquartile Range of the Post-test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>16*</td>
<td>7</td>
<td>11</td>
<td>12 to 15</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
<td>11</td>
<td>9 to 12</td>
</tr>
<tr>
<td>14</td>
<td>19</td>
<td>10</td>
<td>9 to 11</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>9</td>
<td>8 to 11</td>
</tr>
<tr>
<td>12</td>
<td>18</td>
<td>8</td>
<td>6 to 9</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>8</td>
<td>7 to 9</td>
</tr>
</tbody>
</table>

* The mental age level includes months, from 0 to 11.

It is also significant that the highest score achieved in the interquartile range by the highest mental age group of the control population is lower than the lowest score in the lowest mental age group of the experimental population.

Table 2. Mental Age Medians and Interquartile Ranges of the Post-test Scores for the Entire Experimental Population.

<table>
<thead>
<tr>
<th>Mental Age Level in Years</th>
<th>Number of Testees</th>
<th>Median of the Post-test Scores</th>
<th>Interquartile Range of the Post-test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>16*</td>
<td>13</td>
<td>21</td>
<td>22 to 26</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>23</td>
<td>22 to 21</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>21</td>
<td>20 to 23</td>
</tr>
<tr>
<td>13</td>
<td>16</td>
<td>19</td>
<td>17 to 21</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>19</td>
<td>17 to 21</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
<td>17</td>
<td>16 to 19</td>
</tr>
</tbody>
</table>

* The mental age level includes months, from 0 to 11.
3. The Item Analysis

Determination of the mid-80 percent.— The next step in the statistical procedure consisted of arranging the pupil's data cards according to schools and grades, and stacking each set in descending order of the Otis raw scores. The sequence thus obtained was tabulated for each grade, and the mean and standard deviation were calculated according to the following formulae:

\[ M = M_{as} + i \left( \frac{\Sigma fd}{N} \right) \]  

where \( M_{as} \) is the assumed mean, \( i \) is the interval, \( f \) is the frequency, \( d \) is the deviation from the assumed mean, and \( N \) is the number of cases.

And,

\[ \sigma = \sqrt{i \left[ \frac{\Sigma fd^2}{N} - (\frac{\Sigma fd}{N})^2 \right]} \]  

where \( i \) is the interval, \( f \) is the frequency, \( d \) is the deviation from the assumed mean, and \( N \) is the number of cases. In each case the standard deviation \( \sigma \) times 1.28 yielded the upper and lower limits of the scores which should be the upper and lower 20 percent of the grade. In all grades of all five schools surveyed, the two top scores and the two bottom scores of the Otis test results were thus eliminated from the survey as deviates. This elimination left 160 cases for the statistical research to follow.

Determination of the upper and lower 27 percent.— The mid-80 percent population consisting of 160 cases was now separated into the experimental and control sections. The control group was temporarily
set aside since the item analysis to follow was done on the experimental group only. The experimental population now numbered 45 cases in the sixth grade, and 38 cases in the eighth grade. The upper and lower 27 percent of each of these groups was determined simply by stacking each group of data cards in descending order of the post-test scores and taking 27 percent of the number of cases in each of the two grades. This value, rounded off, gave the number of cases to be taken from the top and from the bottom of each stack. The procedure yielded four separate stacks of data cards which were labeled as follows: (1) sixth grade - upper 27 percent, (2) sixth grade - lower 27 percent, (3) eighth grade - upper 27 percent, (4) eighth grade - lower 27 percent.

Determining the difficulty index.-- On each of the four groups listed in the previous paragraph, every single item of the test was analyzed for its pre-test and post-test performance. Four separate sets of proportions answering the item correctly were tabulated, and each of these proportions was corrected for guessing from Guilford's chart. These corrected proportions were then taken to the Davis table in order to determine the "difficulty index" and the "discrimination index" for each item of the test.

Eliminating test items.-- The items which had to be eliminated from the test as too difficult or too easy were determined according


to the following criteria:

1. The item was eliminated if the difficulty index for the item indicated no gain or a decrease from the lower grade.
2. The item was eliminated if 75 percent of the pupils in the lower grade answered the item correctly on the pre-test.
3. The item was eliminated if 25 percent or fewer of the pupils of the higher grade failed the item on the post-test.
4. The item was eliminated when the difficulty index showed no gain from the pre-test to the post-test at both grade levels.
5. The item was eliminated when the difficulty index showed a gain at the lower grade and a loss at the upper grade from pre-test to post-test.

When the above criteria had been applied to each of the 30 items contained in the test, the writer found that 17 of the items were to be eliminated. These items are the ones which carry an asterisk in the sample copy of the test contained in the appendix.

4. The Test of the Principle

Re-grading the tests.--- The remaining 13 good items are now considered to constitute the test of the principle which will be used to determine the results of the survey. The tests of all 160 cases in the mid-80 percent were then re-graded on the basis of these 13 good items, and the scores tabulated to calculate the means and standard deviations of the test of the principle. The formulae (1) and (2) of this chapter were used for these calculations. Table 3 below records these figures for the pre-test and show remarkably good coincidence between the experimental and control groups for both grades. By inspection, the
division into experimental and control pupils was essentially random.

Table 3. Grades Eight and Six Means and Standard Deviations of the Pre-test for the Experimental and Control Populations on the 13-Item Test of the Principle.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Experimental Population</th>
<th>Control Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (2)</td>
<td>Standard Deviation (3)</td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eight</td>
<td>4.26</td>
<td>2.41</td>
</tr>
<tr>
<td>Six</td>
<td>3.06</td>
<td>1.45</td>
</tr>
</tbody>
</table>

Pearson coefficient of self-correlation.— To find the coefficient of self-correlation of the pre-test for each grade, a correlation table showing the relation between the pre-test and post-test scores of the control population was drawn up. From these tables the coefficient for each grade level was computed according to the method outlined in Lindquist[1] using the formula,

\[
    r_{xy} = \frac{\frac{\sum x'y'}{N} - \frac{\sum x'\sum y'}{N}}{\sqrt{\frac{\sum x'^2}{N} - \left(\frac{\sum x'}{N}\right)^2} \sqrt{\frac{\sum y'^2}{N} - \left(\frac{\sum y'}{N}\right)^2}}
\]

(3)

where \( x' \) represents the deviation of an \( x \) score from the arbitrary

reference point, the assumed mean, in the X distribution, and $y'$ represents the deviation of a Y score from the arbitrary reference point, the assumed mean, in the Y distribution, and $n$ represents the number of cases. For the sixth grade population, the Pearson coefficient of self-correlation cameout as plus 0.71, and for the eighth grade, a plus 0.85. These figures establish a workable reliability of the test of the principle.

The mastery level.— In order to determine the proportion of pupils passing the test of the principle it was necessary to establish a mastery level, or the score which had to be attained for the pupil to pass the test. This level of mastery was established by finding the standard error of the obtained measure from the following formula:

$$\sigma_t = \sigma \sqrt{1 - r_{tt}}$$

where $\sigma$ is the standard deviation of the group, and $r_{tt}$ is the Pearson coefficient of self-correlation computed above. The number of good items, or 13 in this case, minus $2.58 \sigma_t$ determined the level of mastery of the test. For the 13-item test used in this research, the level of mastery was found to be 10 in all cases except the sixth grade pre-test in both the experimental and control sections, and for the eighth grade post-test in the experimental group. In these cases the mastery level was 11.

Proportion passing the test.— The proportion passing the test was found for each section by dividing the number of cases at and above the level of mastery by the number of cases in the section and changing to a percent. Tables 4 and 5 below contain this information.
for comparison. A study of table 4 shows a slight gain from pre-test to post-test in the control population for the eighth grade. Although no gain shows for the sixth grade population, a very slight gain was realized despite the fact that no pupil reached the level of mastery established. The writer maintains that this small gain shows that a learning factor exists in retaking a test although no instruction intervenes. It is very likely that greater familiarity with

Table 4. Grades Eight and Six Pre-test and Post-test Means, Standard Deviations, and Proportion Passing the 13-Item Test in the Control Population.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Proportion Passing in Percent</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Proportion Passing in Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Eight</td>
<td>4.42</td>
<td>2.45</td>
<td>2</td>
<td>4.87</td>
<td>2.64</td>
<td>7</td>
</tr>
<tr>
<td>Six</td>
<td>2.97</td>
<td>1.60</td>
<td>0</td>
<td>3.46</td>
<td>1.86</td>
<td>0</td>
</tr>
</tbody>
</table>

the test when a pupil takes it for a second time enables him or her to achieve a better score in some cases. It must also be considered that if a pupil fails to answer some of the items due to the time limit, on the retest it is very likely that this time limit will not be so penalizing.
Table 5. Grades Eight and Six Pre-test and Post-test Means, Standard Deviations, and Proportion Passing the 13-Item Test in the Experimental Population.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pre-test</th>
<th></th>
<th>Post-test</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Proportion Passing in Percent</td>
<td>Mean</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Eight</td>
<td>4.26</td>
<td>2.41</td>
<td>2</td>
<td>11.00</td>
</tr>
<tr>
<td>Six</td>
<td>3.06</td>
<td>1.45</td>
<td>0</td>
<td>9.10</td>
</tr>
</tbody>
</table>
CHAPTER V
RECAPITULATION

1. Conclusions and Interpretations

What the survey has shown—A careful examination of tables 5 and 6 in the previous chapter reveals some very definite information which indicates that a significant gain in learning has been achieved by the pupils who were exposed to the demonstration of the principle. The proportion passing in both the sixth and eighth grades of the experimental population is sufficiently high to conclude that the principle demonstrated can be taught with effective results in learning as low as the sixth grade. It must be understood, however, that it is not the purpose of the thesis to declare what percent of the grade passing the test is the measure of whether or not a given scientific principle can and should be taught at that grade level. The writer appreciates that this limit is purely arbitrary. Because of this fact, it is assumed that any definite statement as to what constitutes the separating margin is unsafe.

Value of the survey—It remains for those directly connected with the problem of curriculum revision in science education to interpret these facts along with the results of similar surveys now being conducted at Boston University and other universities in Eastern United States under the direction of Doctor John G. Read. Doctor Read
maintains that it will be a matter of several years before enough facts are compiled to make definite statements concerning the grade placement of all the main scientific principles now being taught at both the elementary and secondary levels of education. Only then will these figures become meaningful. This thesis attempts to further this widespread program of research. If nothing else, the writer feels that he has contributed a good demonstration for the principle of heat absorption and radiation by surfaces of varying natures, and a reliable test to measure how much learning has been achieved by this demonstration.

2. Suggestions for Further Research

What the survey fails to show.-- If it can be concluded that the proportion passing the test of the principle is large enough at 44 percent to include the teaching of the principle in the sixth grade, then it follows that the learning ability may be at a still lower grade level. Can this principle be taught effectively in the fourth grade? This survey fails to answer this question. This is the problem which will have to be solved by further research along the same lines as the work done here. For that reason, the writer has submitted both the panel demonstration and a taped recording of his lecture to the Library of the School of Education of Boston University. They are intended to serve as tools for further research at a lower grade level.
APPENDIX
A TEST IN SCIENCE

NAME: ___________________________ GRADE: ______ SCHOOL: ______________________

BOY( ) GIRL( ) AGE: ______ BIRTHDATE: _______________ DATE: __________

INSTRUCTIONS: The teacher will show you how to fill in the blank spaces above. For the test, read each question or statement carefully and place a check mark in the bracket beside the correct answer. Each item has only one correct answer. Be sure to answer all items of the test.

1. Rockets now being designed for space travel will have specially constructed patches on their surfaces to regulate the temperature inside. In order to get heat from the sun's rays, these patches should have ..... 

( ) a. a smooth white surface.  
( ) b. a smooth black surface.  
( ) c. a rough silver surface.  
( ) d. a rough black surface.  
( ) e. a polished silver surface.

2. Gasoline companies have found that painting their large storage tanks with light colors has decreased losses by evaporation because they ..... 

( ) a. do not develop leaks so readily.  
( ) b. keep down the temperature.  
( ) c. radiate heat fast.  
( ) d. do not rust as fast.  
( ) e. absorb heat from the sun.

3. Dirigibles are always painted with a smooth silvery surface to cut down ..... 

( ) a. the expansion of gas from the heat absorbed.  
( ) b. rust forming because of rain and fog.  
( ) c. resistance of the air as it moves ahead.  
( ) d. the glare from the sun.  
( ) e. condensing of water vapor on the surface.

* The asterisk signifies that the item was eliminated in establishing the reliability of the test according to the criteria listed in Chapter IV.
4. A hot piece of charcoal will cool faster than a hot piece of lead of the same size because ......

( ) a. the charcoal is porous.
( ) b. the lead is heavier.
( ) c. the lead is a metal.
( ) d. the charcoal is darker.
( ) e. the charcoal is softer.

5. Coffee will stay hot longer in a polished silver pot rather than in a blackened one because ......

( ) a. it radiates heat slowly.
( ) b. it radiates heat rapidly.
( ) c. it absorbs heat from the air around it.
( ) d. it reflects heat from the air around it.
( ) e. it stops all heat from leaving the inside.

6. A rough black globe and a smooth silver globe of equal size are placed an equal distance from the heat of a lamp. Which of the answers below is true after a while?

( ) a. The smooth silver globe will be hotter.
( ) b. The rough black globe will be hotter.
( ) c. They will both be equally hot.
( ) d. Neither globe will absorb heat.
( ) e. The black globe absorbs heat while the other does not absorb heat at all.

7. In the experiment described in no. 6, each globe has a column of red liquid attached to it which will rise as the globe absorbs heat, and fall as it loses heat. After the light shining on the globes has been turned off, which of the things listed below will happen?

( ) a. Both columns will fall at the same rate.
( ) b. The column of the black globe will fall faster.
( ) c. The column of the silver globe will fall faster.
( ) d. Neither column will fall.
( ) e. Only the column of the black globe will fall.

8. A black frying pan is better to use in frying steak than a shiny one because the pan ......

( ) a. gets hotter faster.
( ) b. holds the heat longer.
( ) c. heats up slowly.
( ) d. reflects the heat of the burners.
( ) e. is not so hot to handle.
9. Two boys are eating ice cream on a hot sunny day. Mike is eating chocolate ice cream, and John is eating some vanilla ice cream. Which of the following does John notice?

( ) a. That neither ice cream is melting very fast.
( ) b. That both ice creams are melting equally fast.
( ) c. That Mike's ice cream is melting faster.
( ) d. That Mike's ice cream is melting more slowly.
( ) e. That neither ice cream is melting at all.

10.* On a sunny winter day snow melts faster where it is covered with soot because the soot ......

( ) a. shields the snow from the sun's rays.
( ) b. reflects the sun's rays.
( ) c. shields the snow from the cold air above it.
( ) d. absorbs the sun's rays well due to its color.
( ) e. is warmer than the sun's rays.

11.* Compared to grassy areas, dark-colored soil absorbs ......

( ) a. less heat from the sun.
( ) b. no heat from the sun.
( ) c. the same amount of heat from the sun.
( ) d. much more heat from the sun.

12. A black umbrella does not make a good parasol to protect yourself from the rays of the sun because of its ......

( ) a. size.
( ) b. weight.
( ) c. color.
( ) d. shape.
( ) e. cost.

13.* The inside of a thermos bottle has a silvery finish ......

( ) a. to make it sanitary.
( ) b. to keep liquids hot.
( ) c. to make it leak-proof.
( ) d. so that hot liquids will cool faster.
( ) e. to make it more attractive.

14. A polished metal surface which is hot will ......

( ) a. give off its heat slowly.
( ) b. give off its heat fast.
( ) c. give off no heat at all.
( ) d. give off its heat all at once.
15. A well polished metal surface will.....

( ) a. reflect very little heat.
( ) b. reflect all heat falling upon it.
( ) c. reflect mostly all the heat falling upon it.
( ) d. reflect no heat at all.
( ) e. reflect half and absorb half of the heat falling upon it.

16. Light-colored clothing is worn in summer because.....

( ) a. it reflects a lot of heat.
( ) b. it lets heat of the body pass through easily.
( ) c. it absorbs the sun's rays readily.
( ) d. it is much lighter in weight.
( ) e. it holds the heat which it absorbs.

17. High tension wires carrying electricity are painted silver to lessen the amount of..... during the summer.

( ) a. rust forming on them
( ) b. sagging due to the heat they absorb
( ) c. glare caused by the sun
( ) d. weight on them

18. The hot water pipes in the cellar should be painted..... to prevent loss of heat by radiation.

( ) a. black
( ) b. silver
( ) c. white
( ) d. gray
( ) e. copper-colored

19. A good way of cutting down on the expansion in steel bridges during the summer months is to.....

( ) a. tighten them to the piers.
( ) b. spray them with oil.
( ) c. tighten the cables of the bridges.
( ) d. paint them black.
( ) e. paint them silver.

20. A piece of charcoal which has been heated in an oven will.....

( ) a. give off its heat fast.
( ) b. give off its heat slowly.
( ) c. give off no heat at all.
( ) d. give off its heat all at once.
21. Thin aluminum foil placed between the walls of a house will ....

( ) a. prevent cold from passing through the walls.
( ) b. hold the heat in the house during the winter.
( ) c. make the walls sound proof.
( ) d. stop drafts between the walls.
( ) e. stop dust from entering the house.

22.* The snow covered regions near the North Pole melt slowly in the warm summer sun because the sun's rays ....

( ) a. are absorbed well by the snow.
( ) b. never reach the North Pole regions.
( ) c. are shining less than one fourth of the day.
( ) d. are scattered by dense clouds all the time.
( ) e. are reflected by the snow.

23. Scientists making high-altitude flights in balloons painted the cabins they rode in with black paint in order to have ....

( ) a. protection against the wind.
( ) b. light inside the cabin.
( ) c. cool air inside the cabin.
( ) d. heat inside the cabin.
( ) e. protection against cosmic rays.

24.* If a copper penny and a nickel are placed on a cake of ice and put in the sun for an hour, which will result?

( ) a. The penny will have sunk deeper into the ice.
( ) b. The nickel will have sunk deeper into the ice.
( ) c. Both will have sunk to equal depths.
( ) d. Neither coin will have sunk into the ice.

25.* The dome of observatories where large telescopes are kept is painted silver to protect the telescopes' delicate instruments from ....

( ) a. extreme heat.
( ) b. dust and dirt.
( ) c. extreme cold.
( ) d. too much light.
( ) e. rain and snow.
26. In order that the radiator in your home will give the greatest amount of heat very fast it should be ..... 

( ) a. painted a dull white.  
( ) b. painted silver.  
( ) c. painted dull black.  
( ) d. painted glossy black.  
( ) e. painted copper-color.

27. A camera will become very hot if left standing out in the sun because ..... 

( ) a. it is sensitive to light.  
( ) b. it reflects heat.  
( ) c. it has a light-proof interior.  
( ) d. it has a black outer surface.  
( ) e. it has metal parts which absorb all the heat.

28. Which would feel the heat most on a hot summer day? 

( ) a. A man wearing a white silk shirt.  
( ) b. A man wearing a black silk shirt.  
( ) c. A man wearing a black cotton shirt.  
( ) d. A man wearing a white cotton shirt.  
( ) e. A man wearing a light blue silk shirt.

29. Which of the following statements is correct? 

( ) a. A good radiator of heat is a good absorber of heat.  
( ) b. A poor radiator of heat is a good absorber of heat.  
( ) c. A good radiator of heat is a poor absorber of heat.  
( ) d. A good radiator of heat absorbs no heat at all.  

30. A movie projector box usually has a dull black surface so that it 

( ) a. won't show the dirt so easily.  
( ) b. will radiate its heat faster.  
( ) c. will hold heat better.  
( ) d. will not rust.  
( ) e. will reflect light better.
LIST OF DIFFICULT WORDS

Absorb: Absorb means to take into itself.

Condensing: Condensing means that water vapor in the air is changing from a gas to water drops on a cold surface.

Cosmic rays: Cosmic rays are rays which pass through space like rays of light or heat.

Designed: Designed means planned on paper.

Dirigible: A dirigible is a lighter than air craft such as a blimp.

Evaporation: Evaporation means a change from water to an invisible gas called water vapor.

Expansion: Expansion means that a gas is enlarging or trying to occupy more space.

Foil: Foil is a very thin sheet of metal.

Globe: A globe is a ball-shaped object.

Glossy: Glossy means having a smooth shiny surface.

Porous: Porous means having tiny holes like a sponge.

Radiate: Radiate means giving off heat from itself.

Radiation: Radiation is the act of giving off heat or light.

Reflects: Reflects means bouncing off an object like rays of light from a mirror.

Thermos bottle: A thermos bottle is a bottle used to keep liquids like coffee in a lunchpail.

Silvery: Silvery means smooth and shiny like silver that is well polished.

Vapor: Vapor is a gas which is usually invisible like the water vapor that is in the air of this room.
LECTURE OF THE DEMONSTRATION

The demonstration which you are about to see will explain to you a scientific principle which has many important applications in the world around you. All things in nature are capable of absorbing, or taking into themselves, heat from other objects around them that contain more heat than themselves. It is also a well-known fact that all things in nature are capable of radiating, or giving off heat to objects around them if they contain more heat than these objects.

Today we are going to learn that the type of surface that objects have tend to make them better radiators and absorbers of heat. Here on this panel, we are going to make two separate globes absorb heat, or take heat into themselves. After they have absorbed heat for a short while, we are going to let them radiate, or give off this heat to the air around them. What we want to see is if the type and color of the surface of the globes have anything to do with how fast these globes are going to absorb and then radiate this heat.

Notice first that the surface of one of the globes is smooth and shiny, while the other is dark and rough. To each of the globes there is a thermometer attached. A thermometer is an instrument which measures how hot an object is. As the globes begin to absorb heat from the lamp above, the red liquid will rise. (The lamp is turned on at this point.) The lamp is giving off what we call "radiant energy" in the form of light and heat, and the globes are absorbing this energy.
That is what causes them to become hot. It does not take long to see what is happening. You can now tell which of the two globes is absorbing heat faster. The big difference in the heights of the columns of red liquid prove that the black globe which has a rough surface seems to absorb much more heat than the smooth, light-colored globe.

Now, let's turn off the lamp (the lamp is turned off at this point) and see what happens. We will adjust the metal clamps at the point where the columns have stopped rising. In just a few moments we will be able to make a comparison between the two thermometers. By counting the lines on the scale of each thermometer you can tell which of the two columns falls faster. Again you can see that the column that is attached to the black globe falls much faster than the other. This shows that the globe which has a dark, rough surface tends to radiate heat much faster than the globe which has a smooth, light-colored surface. That gives you the answer as to what type of surface absorbs and radiates heat more readily.

From this experiment, we can say that bodies which have smooth, light-colored surfaces tend to absorb very little heat; but what heat they do absorb they do not radiate very fast. On the other hand, bodies which are dark in color and have rough surfaces tend to absorb a great deal of heat fast, and to give off or radiate this heat just as fast.

Try to recall this from the experiment which you have just seen:

1. A good absorber of heat is also a good radiator of heat.
2. A poor absorber of heat is also a poor radiator of heat.
You will now have a chance to repeat the test you have taken before seeing this demonstration. If you remember what you have just seen and heard, you will be able to do a much better job on it.
<table>
<thead>
<tr>
<th>Item</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mo.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otis form</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date today</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimenter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kit number</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mental age in mos., 100+
Plate 1. The demonstration panel as it appears when taken from its packing case

Plate 2. The panel with its thermometer mechanism as it appears before being set up

Plate 3. Detail of the thermometer mechanism as it appears when prepared for operation before the demonstration
Plate 6. The demonstration panel in actual operation
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 Test of the Principle
and for the Methods of Statistical Research


