AN ABSTRACT OF THE THESIS OF

William Herbert Fryback for the Ph. D. in General Science
(Name) (Degree) (Major)

Date thesis is presented May 10, 1965

Title EVALUATION OF MULTI-LEVEL READING MATERIALS,
INTRA-CLASS DISCUSSION TECHNIQUES AND STUDENT EXPERI-
MENTATIONS ON ACHIEVEMENT IN FIFTH GRADE ELEMENTARY
SCIENCE

Abstract approved Redacted for Privacy
(Major professor)

The purpose of this study was to statistically evaluate one kind
of new elementary science curriculum materials. The materials;
(1) are written at five reading levels to partially meet the reading
ability spread found in the intermediate grades, (2) provide ways to
stimulate class discussion, and (3) provide three levels of sophisti-
cation for student experimentation on both an individual and a class
basis. The significance of each of these parts was evaluated by an
analysis of variance.

Population

The materials were used with 29 fifth grade classes in three
school districts within a radius of 60 miles of Portland. The classes
were divided into six different treatment groups which were
distributed among the schools.

Treatment Groups

The six treatment groups were composed as follows:

Treatment 1. Students read only the middle (C) level of the five reading levels provided in a full set of the materials. They discussed the materials and also carried out the experiments provided with the (C) level materials.

Treatment 2. Students read, according to ability, at three levels; A (the lowest), C (average) and E (the highest). Students discussed the materials and worked the experiments provided with these three reading levels.

Treatment 3. Students read, according to ability, the materials at five reading levels; A, B, C, D, E. They discussed the materials and worked the experiments provided in the full set of materials.

Treatment 4. Students used the materials as in Treatment 3 except they did not discuss any of the materials with their teacher or classmates.

Treatment 5. Students used the materials as in Treatment 3 except they did not work any of the experiments.

Treatment 6. Students did not use any of the materials.
Null Hypotheses Tested

There are no greater differences in achievement on criterion test items about atmosphere, among classes of fifth grade students who: (1) study materials written at one, three or five reading levels, (2) do, or do not interact in classroom discussion, or (3) do, or do not perform both individual and group science experiments, than would be expected to result from the vagaries of random sampling from a single, normally distributed, infinite population.

Results

1. There were no significant differences on achievement gain scores among the first three treatment groups.

2. There were no significant differences on achievement gain scores among any of the first three treatment groups and treatment group four.

3. There was a significant difference on achievement gain scores between treatment group five and any of the other treatment groups.

Conclusions

Based on the results of this investigation it can be stated that:

1. There are no advantages in achievement for students who
respectively read these materials written at their reading ability level.

2. There are no achievement advantages to students who discuss these materials with other members of their class.

3. There are significant achievement advantages to students who work the science experiments provided in these materials.

The results indicate to the researcher that science experiments which involve active participation on the part of the learner provide significant advantage to him and that they might well be made the central part of elementary science curriculum materials. Although the provisions for reading ability levels and class discussion did not show significant influence on achievement, further investigation needs to be carried out to evaluate their effectiveness on other objectives; (1) arousing curiosity in science, (2) stimulating creativeness, and (3) developing skills of inquiry.
EVALUATION OF MULTI-LEVEL READING MATERIALS, INTRA-CLASS DISCUSSION TECHNIQUES AND STUDENT EXPERIMENTATIONS ON ACHIEVEMENT IN FIFTH GRADE ELEMENTARY SCIENCE

by

WILLIAM HERBERT FRYBACK

A THESIS submitted to OREGON STATE UNIVERSITY

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

June 1965
APPROVED:

Redacted for Privacy

Chairman of Department of Science Education

In Charge of Major

Redacted for Privacy

Chairman of Department of General Science

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 10, 1965

Typed by Marion F. Palmateer
ACKNOWLEDGMENTS

The world moves through the association of people with people. It has been this writer's good fortune to have associated with the finest examples of people who could be found any place. Their example, help and inspiration have been his wellspring. Their only reward is limited to a simple "thank you" by the confines of a sheet of paper, but perhaps their "torch" can be passed on as this one tries to serve others as he has been served. Heartfelt thanks are extended to the following:

To my wife Marilyn who has given beyond self.

To Dr. Stanley Williamson whose example is an inspiration to all who know him.

To Dr. Donald Stotler and Dr. Donald Parker, authors and both of whom are fondly regarded as friends.

To Dr. Jack Edling and Dr. James Beaird of Teaching Research who have given counsel and encouragement unselfishly.

To Dr. Fred Fox for friendship and constructive criticism.

To Mr. Ben Purvis for aid in computer programming.

To Science Research Associates who loaned their materials and who gave other tangible assistance.

To the students, teachers and administrators of the schools who used the materials.
To a very helpful graduate committee and a legion of friends and associates.

Finally, to my children who have quietly sensed the pressures and who helped immeasurably.

W. H. F.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td></td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Assumptions</td>
<td>7</td>
</tr>
<tr>
<td>The Following Null Hypotheses Were Identified and Tested</td>
<td>8</td>
</tr>
<tr>
<td>Organization of the Materials</td>
<td>8</td>
</tr>
<tr>
<td><strong>REVIEW OF LITERATURE</strong></td>
<td>9</td>
</tr>
<tr>
<td>Need for Graded Reading Materials</td>
<td>16</td>
</tr>
<tr>
<td>Importance of Class Discussion</td>
<td>16</td>
</tr>
<tr>
<td>Need for Student Experimentation</td>
<td>25</td>
</tr>
<tr>
<td>Summary</td>
<td>26</td>
</tr>
<tr>
<td><strong>THE STUDY</strong></td>
<td>29</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>31</td>
</tr>
<tr>
<td>School Contacts</td>
<td>34</td>
</tr>
<tr>
<td>Student Population</td>
<td>36</td>
</tr>
<tr>
<td>Criterion Test</td>
<td>39</td>
</tr>
<tr>
<td>Teacher Questionnaire</td>
<td>47</td>
</tr>
<tr>
<td><strong>PRESENTATION AND INTERPRETATION OF DATA</strong></td>
<td>48</td>
</tr>
<tr>
<td>Mean Gain Score Data by Class, Treatment and Reading Level</td>
<td>48</td>
</tr>
<tr>
<td>Data Concerning the Influence of Materials Written at Five Reading Levels</td>
<td>53</td>
</tr>
<tr>
<td>Data Concerning the Influence of Whether or Not Students Discuss What They Read</td>
<td>57</td>
</tr>
<tr>
<td>Data Concerning Whether or Not Students Performed Experiments</td>
<td>59</td>
</tr>
<tr>
<td>Data Concerning Influences of Classes within Treatments</td>
<td>61</td>
</tr>
<tr>
<td>Data Derived to Aid in the Interpretation of Results</td>
<td>61</td>
</tr>
<tr>
<td>Data Concerning the Teacher Questionnaire</td>
<td>65</td>
</tr>
<tr>
<td><strong>SUMMARY, CONCLUSIONS AND RESULTS</strong></td>
<td>69</td>
</tr>
<tr>
<td>Summary of Results</td>
<td>69</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Conclusions</td>
<td>70</td>
</tr>
<tr>
<td>Recommendations</td>
<td>76</td>
</tr>
<tr>
<td>Summary</td>
<td>78</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>79</td>
</tr>
<tr>
<td>APPENDIX I</td>
<td>81</td>
</tr>
<tr>
<td>APPENDIX II</td>
<td>94</td>
</tr>
<tr>
<td>APPENDIX III</td>
<td>106</td>
</tr>
<tr>
<td>APPENDIX IV</td>
<td>113</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Schematic Organization of &quot;Learnings in Science&quot; Laboratory</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Mean Gain Scores for Six Treatments and for Five Reading Levels</td>
<td>51</td>
</tr>
<tr>
<td>3</td>
<td>Mean Gain Scores for First Three Treatments and for Five Reading Levels</td>
<td>52</td>
</tr>
<tr>
<td>4</td>
<td>Mean Gain Scores for Treatments 3 and 4 and for Five Reading Levels</td>
<td>52</td>
</tr>
<tr>
<td>5</td>
<td>Mean Gain Scores for Treatments 3 and 5 and for Five Reading Levels</td>
<td>53</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reading and Work Level Organization for Learning in Science Research Booklets</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>Description of Treatment Groups</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>Number of Students Assigned in Different Reading Groups</td>
<td>37</td>
</tr>
<tr>
<td>4</td>
<td>Mean Intelligence Quotient, Reading Level and Number of Students at Each Reading Level for Treatments and Classes</td>
<td>38</td>
</tr>
<tr>
<td>5</td>
<td>Mean Intelligence Quotients and Grade-Equivalent Reading Scores for the Six Treatments</td>
<td>39</td>
</tr>
<tr>
<td>6</td>
<td>Criterion Test Question Validation Data</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>Number of Questions for Each Big Idea</td>
<td>42</td>
</tr>
<tr>
<td>8</td>
<td>Number of Questions at Each Reading Level by Categories of Questions</td>
<td>42</td>
</tr>
<tr>
<td>9</td>
<td>Schematic of Analyses</td>
<td>46</td>
</tr>
<tr>
<td>10</td>
<td>Class and Treatment Mean Gain Scores by Treatments</td>
<td>49</td>
</tr>
<tr>
<td>11</td>
<td>Mean Gain Scores for Groups of Students Reading at Different Reading Ability Levels</td>
<td>50</td>
</tr>
<tr>
<td>12</td>
<td>Significance of Reading Level (RL) and Treatment (T) When Different Combinations of Treatments are Compared</td>
<td>54</td>
</tr>
<tr>
<td>13</td>
<td>Significance of Treatments on Questions Unique to Reading Levels</td>
<td>55</td>
</tr>
<tr>
<td>14</td>
<td>Significances of Treatments on Questions in Special Categories</td>
<td>56</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>15</td>
<td>Significances of Treatment 4 Involving Class Discussion</td>
<td>57</td>
</tr>
<tr>
<td>16</td>
<td>Significances of Treatment 4 on Questions Unique to Different Reading Levels and to Questions in Six Special Categories</td>
<td>58</td>
</tr>
<tr>
<td>17</td>
<td>Significances of Treatment 5 Involving Nonperformance of Experiments</td>
<td>59</td>
</tr>
<tr>
<td>18</td>
<td>Significances of Treatment 5 Involving Nonperformance of Experiments on Questions Unique to Reading Levels and on Questions in Six Different Categories</td>
<td>60</td>
</tr>
<tr>
<td>19</td>
<td>Significances of Class Mean Scores within Each Treatment Group</td>
<td>62</td>
</tr>
<tr>
<td>20</td>
<td>Significances of Class Mean Scores within Each Treatment Group on Questions Unique to Reading Levels and to Questions in Six Categories</td>
<td>63</td>
</tr>
<tr>
<td>21</td>
<td>Some Mean Values by Treatments</td>
<td>64</td>
</tr>
<tr>
<td>22</td>
<td>Coefficients of Correlation</td>
<td>65</td>
</tr>
<tr>
<td>23</td>
<td>Rank Order from Highest to Lowest Agreement for the Questions on the Teacher Questionnaire</td>
<td>66</td>
</tr>
</tbody>
</table>
EVALUATION OF MULTI-LEVEL READING MATERIALS, INTRA-CLASS DISCUSSION TECHNIQUES AND STUDENT EXPERIMENTATIONS ON ACHIEVEMENT IN FIFTH GRADE ELEMENTARY SCIENCE

INTRODUCTION

The trend in science education for the public schools today is away from subject matter as the main goal and toward the development of curiosity, positive attitudes, creativeness and skills of inquiry which contribute to the understandings of conceptual schemes in science. This should not be interpreted to mean that science teachers are no longer interested in their students learning subject matter. Quite to the contrary, it means that science educators believe that students whose curiosity for science has been stimulated and whose skills of inquiry have been sharpened by active participation in investigations will gain equally as much in subject matter achievement and immeasurably more in the areas of positive attitudes and creativeness.

The major efforts in science curriculum preparation today are directed toward giving students experiences in real investigations through which they learn relationships, patterns and conceptual schemes which bind the disciplines of science into one. This is far different from simply memorizing and verifying the work of others.

Science educators have realized for some time that the major objectives of science education are two-fold: (1) to produce a small
population of men and women with enthusiasm, skills and ideas who will spend their lives in the field of science; and (2) to produce a large population of men and women who are scientifically literate and who exhibit the behaviors of sympathy and empathy for the scientific enterprise. The former become our professional scientists, the latter become the laymen (and everyone is a layman outside his own specialty) who must use and help support the scientific enterprise by their monies and their legislative directives via the ballot.

Both of these populations can be developed by similar methods, but not "out of the same book." This simply means that nearly all people are subject to the same laws of learning, but that every individual reacts differently to them, and we must accommodate for these differences within the framework of the laws of learning.

No one has really been able to conceive a better ratio than that of one teacher to one pupil in meeting individual differences, but even that does not guarantee pupil needs will be met. Realistically, in these days of teacher shortage and more and more young people to be educated, it is folly to yearn for the teacher and the student on opposite ends of the log. Neither does it mean that students must be lumped into groups of hundreds where the pace and limits and procedures are set at the same level for all. Individual differences can be met by the production of materials that are motivational, variable, flexible and incomplete.
The term motivational probably is well understood; materials must stimulate the student to make good use of them. Variable means that the materials do not have to be organized in such fashion that every student has access to identical materials. Some should be at different reading and concept levels. Flexible means that they do not have to be used in the same manner by each individual. Some students may finish with a topic in one day whereas another student may work on the topic for a week. The last, incomplete, may be the most nebulous and yet the most important. It simply means materials cannot be a recording of information. It is far better that they have only a few answers for reinforcement, but many more unanswered challenging situations, therefore, incomplete. One attribute of a good teacher may be that he can point out what is not known in a subject matter area, rather than to be able to recite what is known. Good curriculum materials should have the same potential.

Learning to use instructional materials like these is difficult but teachers can learn the techniques necessary to work with students as individuals or in small groups, each learning in his own way if the materials suitable to the student are available. Teachers know about individual differences in students and they can usually identify their needs, but they do not have access to the materials necessary to meet these needs. Consequently, teachers must form large, economical groups of students and set the limits of student behavior compatible
with the conformity necessary to present a lesson "out of the same book." This technique has become so commonplace that most people simply accept it.

In order for teachers to do anything different the necessary curriculum materials must be produced. As our school populations increase a greater variability within the population accompanies it. Still our emphasis seems to be greater production of the same kinds of inadequate materials to be used in the same inadequate administrative units. The solution does not lie only in providing still more school rooms and teachers, but in the development of the concept that students can largely be self-learners when they are provided with the proper kinds of materials, and are reinforced by success, the contributions of their peers and indirect consultation by teachers.

We do not know just what it is that sparks creativity but in consideration of the remarks of many of our creative people toward their school experience the factor may not have been the kind of school experiences they received. Of course, there may be a negative kind of correlation, but the point is, we must think more broadly concerning curriculum materials production. It is not enough to design one kind of material and consider it good for everybody. It may be good for some, but this does not guarantee that it will be at all good for others.

One of the greatest stumbling blocks in the preparation of materials which try to meet the criteria listed above is that of proper
evaluation. We have not developed reliable instruments by which to measure adequately the objectives of curiosity, positive attitudes, creativity and skills of investigation. We have not been able to show evidence. The reliability of consistently producing these behaviors in students is still a "seat of the pants" operation. In spite of this, many educators are taking the lead in the production of curriculum materials which try to accomplish these larger objectives.

The efforts began at the secondary level through efforts of the Physical Science Study Committee, the Biological Science Study Committee, the Chem Study Committee and the Chemical Bond Approach Committee. Very recently there have been a couple of efforts to integrate physics and chemistry. At the elementary level there have been the Princeton Studies, the Illinois Program, the American Association for the Advancement of Science Program and many others. The above mentioned have all been subsidized by Federal funds. All are oriented away from subject matter "coverage" and toward the objectives already listed. They attempt to help students discover the structure in science through active participation in investigations, many of which do not have known answers. Most of their laboratory investigations are designed for small groups. However, the instructional techniques in nearly all of them still cling somewhat to the group presentation method.

Evaluation of these programs has been difficult because the only
kind of reliable evaluation that has been developed is for subject matter achievement and this is not the principal objective of these new materials. Nevertheless, subject matter achievement evaluation has been carried out on them and in most cases they have been found to equal the more conventional subject matter oriented materials.

Similar limitations present themselves in carrying out this evaluation of a new kind of material (page 9) developed for science in the elementary school. The authors of the materials under study have attempted the complex task of supplying what they term "multi-level" materials. Basically this means (1) to provide the learner with materials prepared at his reading level; and more generally (2) to involve him in class discussion about both the things he has read and the things which his classmates have read; (3) to provide him with experiences in both individual and group experimentation; and (4) to encourage him to evaluate his own progress. More broadly it is defined in these terms: to provide materials for a learning situation where each student may start where he is and move ahead as fast and as far as his learning rate and capacity will let him. The objectives of "multi-level" materials are to arouse curiosity in students, to stimulate creativeness, and to provide for the acquisition of the skills of inquiry and the knowledge of science. Since it is not possible to measure all these objectives reliably, this researcher has evaluated the materials from the standpoint of achievement or knowledge of science.
The rationale was that if it can be shown that the design of the materials effect a significant difference in achievement gain, then producers of these and other materials like them can be encouraged to continue trying to perfect their productions.

Other objectives of the authors may be better measured after they have produced enough materials to influence student behavior over a longer period of time.

**Statement of the Problem**

The purpose of this study is to determine the effectiveness, in terms of achievement, of three basic and unique aspects in preparation of one kind of science materials for the elementary school: (1) presentation of materials written at five reading levels; (2) stimulation of class discussion through lead questions in a unique Teacher's Instructional Aid Booklet; and (3) the contribution of active student involvement in both individual and group science experimentation. Effectiveness of these three was determined through a statistical comparison of student pretest and post test gain scores on a criterion test designed specifically for these materials (Appendix I).

It is important to note that the study is designed to get results which reflect the manner in which most teachers will use the materials, i.e., with only a minimum of in-service instruction. Theoretically it would have been possible to control the experience level and the
science training background of the teachers, the amount of time spent and the ability levels of the students. However, it was the objective of the research to get results which show the more typical situation represented when a school district purchases materials for its teachers and expects them to use the materials without special training and with heterogeneous groups of students. The researcher chose the real situation in order to better evaluate the materials as they will actually be used.

**Assumptions**

The following assumptions were made in carrying out the study:

1. Materials are written at reading grade levels described.
2. Elementary teachers, with limited science backgrounds, can handle the complex materials in the manner recommended by the authors.
3. Classroom teachers followed prescribed instructions.

**The Following Null Hypotheses Were Identified and Tested**

1. There are no greater differences in achievement on criterion test items about atmosphere, among classes of fifth grade students who study materials written at one, three or five reading levels, than would be expected to result from the vagaries of random sampling from a single, normally distributed, infinite population.
2. There are no greater differences in achievement on criterion test items about atmosphere, among classes of fifth grade students who do, or do not interact in classroom discussion stimulated by leading questions taken from materials written at different reading levels and provided for the teacher in a "discussion aid" booklet, than would be expected to result from the vagaries of random sampling from a single, normally distributed, infinite population.

3. There are no greater differences in achievement on criterion test items about atmosphere, among classes of fifth grade students who do, or do not perform both individual and group science experiments, than would be expected to result from the vagaries of random sampling from a single, normally distributed, infinite population.

Organization of the Materials

The materials, Earth's Atmosphere Laboratory, published by Science Research Associates are quite unique from the "one textbook" concept of instructional materials.

They are designed primarily to meet individual reading needs so that a well organized teacher can allow students to move through the materials at their own rates. The materials are arranged around five central "big ideas" about atmosphere. Each "big idea" is discussed at five different reading levels. Even the activities are at three to five different levels of sophistication. There are activities
"Five Big Ideas" or "Concepts"

I  Air takes up space
II Air has weight
III Air is made up of different gases
IV Air changes with altitude
V Air is a fluid

Class motivational chart with drawings depicting something about each "Big Idea." A curiosity arouser.

Research booklets
Research booklets
Research booklets
Research booklets
Research booklets

Class Discussion
Class Discussion
Class Discussion
Class Discussion
Class Discussion

Experiments
Experiments
Experiments
Experiments
Experiments

Home activity
Home activity
Home activity
Home activity
Home activity

Classroom experiment
Classroom experiment
Classroom experiment
Classroom experiment
Classroom experiment

*3-4 1 3-4 1 3-4 1 3-4 1 3-4 1
*4-5 2 4-5 2 4-5 2 4-5 2 4-5 2
*5-6 3 5-6 2 5-6 3 5-6 2 5-6 3 5-6 2
*6-7 4 6-7 4 6-7 4 6-7 4 6-7 4
*7-8 5 7-8 5 7-8 5 7-8 5 7-8 5

Class Discussion
Class Discussion
Class Discussion
Class Discussion
Class Discussion

* Grade equivalent reading level

STUDENT RECORD BOOK

Record answers to questions
Score own questions and graph progress
Record predictions for experiments
Record results of experiments
Record conclusions about experiments

Figure 1. Schematic Organization of "Learnings in Science" Laboratory
which call for interaction with classmates, but a student could use the materials without teacher assistance.

The following description of materials used in the study can best be understood by referring to Figure 1 on the previous page.

Laboratory Contents

**One-level Picture Chart.** Teachers are encouraged to use this 24 inch x 30 inch chart to promote thinking about this area of atmosphere. It has five large drawings, one for each of the five basic concepts (called "big ideas") about atmosphere. It is designed to motivate students to want to learn about the concepts.

**Student Record Book.** Each student received a record book which has three principal parts: (1) A one-reading-level article which is read by the teacher as a model of the articles which the students will read. Its purpose is to instruct about how to use the information. (2) Several pages on which students record their answers to questions, write out their experiment predictions, list materials for the experiments and then their responses to the results of the experiments. (3) A chart on an instructional page shows them how to figure the percentage right which they make on questions which test comprehension on the reading selections. There is also a graph on to which they plot their percentage scores and can keep a complete record of their
progress while working with the laboratory materials.

Research Booklets. These are reading selections. Five concepts about atmosphere were identified and ten writers contributed selections about the concepts. Rather than write a single selection (or chapter) about each "big idea" designed for children of average reading ability in a class, five research booklets for each "big idea" were written. Each research booklet was prepared for a different reading level which ranged from grade three to grade eight. Each of the five research booklets, for each "big idea," was not the same article with the reading level simply adjusted to a different reading difficulty. Instead, each of the five research booklets for a "big idea" was written by a different author who presented the "big idea" from a unique point of view. This is important because it means that the different research booklets have some content which is common, or similar to that of the other four (for any "big idea") and some content which is unique, or different from the others. The authors did not try to cover the whole "big idea" in each research booklet. Class discussion and experimentation was intended to supplement individual reading selections in such ways that the total experience for the class would contain most elements of any "big idea." The authors designate this organization as the "multi-level" approach and through its use any student can start where he is and move as quickly and as far
as he is able. The following shows the five reading grade-level organizations of the research booklets and the number of words used.

Table 1. Reading and Work Level Organization for Learnings in Science Research Booklets

<table>
<thead>
<tr>
<th>Research booklet reading level</th>
<th>Reading ability grade-equivalent</th>
<th>Number of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3-4</td>
<td>325-425</td>
</tr>
<tr>
<td>B</td>
<td>4-5</td>
<td>475-525</td>
</tr>
<tr>
<td>C</td>
<td>5-6</td>
<td>650-750</td>
</tr>
<tr>
<td>D</td>
<td>6-7</td>
<td>850-950</td>
</tr>
<tr>
<td>E</td>
<td>7-8</td>
<td>1050-1150</td>
</tr>
</tbody>
</table>

The reading difficulty of the lowest level (A) was figured according to the Spache formula (17, p. 116). The other four levels were determined by counting the number of sentences and the number of syllables in each one-hundred words. To determine the reading level, these counts were compared on a "reading-ease calculator" published by Science Research Associates.

At the conclusion of each reading selection in a research booklet, the student is asked to respond to two sets of multiple choice questions. The first set asks him to recall accurately what he has read and the second asks him to use what he has learned in new situations.

The last part of the research booklet calls for two kinds of experiments. (1) A home activity designed to be an individual experience but which may require the assistance of a helper who is
recommended to be a member of the student's family. (2) A class experiment which is designed to be a small group experience and which calls for mutual cooperation. Experiences in both situations are described by the authors as "open-ended" experiments, i.e., an experiment whose answer may not be known and one which the student may enlarge upon as much as he wishes.

**Self-evaluation.** One of the uniquenesses of the materials is the provision for self-evaluation. The laboratory contains 15 "key cards" which have the correct answers to the objective questions which follow each reading selection. There are also 20 "key model booklets" which contain discussions about the predictions and the results of the experimentation. Students are instructed to use these to score and evaluate their own work on each of the "big ideas." A record of this evaluation is kept in the student record book described above.

**Teacher's Handbook.** This contains instructions to the teacher about the organization of the laboratory and how to use the materials. It also has a discussion about the general organizational philosophy of the materials and what they will do for students if used properly.

**Teacher's Instructional Aid Booklet.** Its primary purpose is to help teachers direct class discussion. It contains several questions which the teacher can ask to stimulate class discussion. Some of
them are aimed at the uniquenesses of specific reading levels. A question may be asked about which only students who read at level B had an opportunity to read. Other questions may ask about the uniquenesses of other reading levels. It is the intention of the authors that this technique will bring about two desirable ends. (1) Slow readers, who usually forfeit participation in class discussion to members of the class who are better readers, will be brought into the discussion sessions because only they may be able to answer questions about the uniquenesses of their own reading selections. (2) Gaps in materials at any one level will be filled in and ideas stimulated by the discussion of the uniquenesses will bring about a common understanding of principles or concepts.
REVIEW OF THE LITERATURE

Experimental evidence similar to this study is virtually non-existent--but there are many authoritative statements which support the theory of the development of the materials.

Need for Graded Reading Materials

There are many references expressing a need for the recognition of the wide reading abilities among students in any one classroom. There is need for a variety of materials to meet the reading differences.

Since an individual's achievement may be affected by a wide variety of conditions, one may expect variability in reading scores on any level of intelligence. The scatter in reading scores of a given chronological and mental age is wide, often extending over the whole range of the reading test (18, p. 27).

Recognition that there are individual differences in reading abilities and interests in every grade has given impetus to the individualization of reading instruction. The third grade, in many schools, contains nonreaders as well as pupils who can read books of seventh grade difficulty. This range of differences widens as the children advance through the grades (18, p. 40).

An experiment in the Chicago schools for homogeneous grouping for reading instruction found that they had to make nine different reading groups with grade equivalent reading abilities from 1.7 to 7.0 in grades six, seven and eight (2, p. 135-139).
Sixth grade classes in a school in Bellingham, Washington were given the Gates Reading Survey. It was found that only six of a particular group of 27 students were reading at the sixth grade level. The other scores were spread out fairly evenly from 2.8 to 8.8 (9, p. 113).

This kind of evidence would make it seem desirable to have an abundance of grade-level reading materials in all areas of the curriculum to meet the needs of the individual reading differences in students.

With such variations as these to be found in the average classroom, it is of course unwise and unfair to demand the same results within a given length of time from all the pupils of the group. The relative simplicity, however, of giving a single assignment to an entire class, the apparent economic and administrative necessity of maintaining fairly large classes for each teacher, the reluctance of many teachers to experiment with innovations which seem to require more extended planning, and the seductively business-like precision which seems to pervade a sharply graded school system, have all combined to preserve the delusion that all pupils in a given grade are capable of uniform achievement, with the more vicious fallacy that the teacher is impartially "tREATING THEM ALL ALIKE" by demanding the same daily tasks of all. As a matter of fact, a uniform requirement means only half-work for part of the class, and impossible effort for another part. Thus the teacher who tries to strike a fair average in the assignment is really unfair to the brighter pupils by depriving them of any challenge to capacity effort, and equally unfair to the slow pupils by depriving them of any opportunity for genuine mastery, and corresponding educational growth. Some teachers, realizing the plight of the latter group, direct their main teaching efforts at the slow pupils, and justify their course by declaring that the "BRIGHT ONES WILL LEARN ANYHOW." Others, with equal logic but less compassion, allow the stronger pupils to absorb the chief attention and set the pace, on the theory that those who cannot meet such a standard should drop back into a grade where they can do the work. In reality none of
these three plans is fair or democratic, in the sense of granting to all pupils equal opportunities for the best educational progress of which they are individually capable (20, p. 351-352).

A number of authorities have pointed out the desirability of having a diversity of materials in the teaching of science which meet both interest and reading ability levels. It is much more common for teachers in the elementary school to have a diversity of materials for instruction in reading than to have them in other areas of the curriculum.

Not enough has yet been done to provide in our classrooms the variety of science reading materials that we need in order to meet the wide range of science interests among children and also to stretch the abilities of the better readers while meeting the needs of those children with less reading skill. If we recognize that very often our groups have a spread of as much as four grades in reading achievement, and that the better teaching of reading will increase rather than decrease this spread, we should see that it is necessary to have a variety of reading materials in science available to every classroom (12, p. 33).

One of the major obstacles to the development of science curriculum materials at different reading levels is the financial obligation involved. It is not just a little financial venture to launch a program of providing materials for individual differences. In some respects it may even be considered a missionary effort because it actually cuts down the volume of mass production sales on the part of textbook publishers.

Many persons have suggested that three levels of the same textbook should be prepared for every course.
(1. The slow and retarded learner, 2. the average student, and 3. the superior student.) Thus the same basic areas would be covered, but the level of presentation and understanding would be differentiated for the students. In theory, this is a fine plan. But publication of a single textbook now involves an investment of $75,000 or more (10, p. 151).

The lack of curriculum materials at reading level has complicated the teacher's task in classroom instruction. Teachers develop files of materials on their own, but this requires several years and it is difficult to keep the selections up-to-date.

Teachers of remedial reading, teachers of retarded students and, in general, classroom teachers confronted with the problem of teaching a group of students having widely divergent reading levels from a single text which proves "too hard" for those at the lower levels often ask the question, "But where can I find the materials which will aid me in teaching each student at his own level?" The short annotated list of books, periodicals and bulletins containing lists of controlled vocabularies—the highest interest reading materials is offered as a partial answer to this oft-repeated question (15, p. 400).

Upon closer examination we find Schutte suggests eight books, three bulletins and four periodicals. Out of these 15, 14 of them are directly related to the slow or handicapped reader. What is being done for the good reader who is bored with listening to others? The one who wants to go ahead? Of course, there are materials lists for him too, but still the problem of the teacher having to seek them out is present.

One of the most thought-provoking statements concerning the diversity of materials needed in the schools is the following:
Grading Children. The graded system was advocated because pupils in the same grade could have the same class books. It is the writer's observation that many learning disabilities in the schools of today are caused by this idea of providing every child—regardless of his capacity or achievement—in a given grade with the same basal textbook. This type of regimentation undoubtedly contributes substantially to retardation in reading and to the perpetuation of individual reading difficulties. Furthermore, it is apparent that the correction of this particular situation on the basis of preventive measures is not likely to take place until every teacher becomes a student of individual differences in relation to reading activities (3, p. 17).

It will certainly be stimulating to teachers to have available materials whose philosophy it is to meet individual differences. It should cause them to be even more cognizant of the facts of individual differences since the materials will demand that they be used properly. The practice of using the same materials in quantities large enough that each child has the same materials has caused teachers to become somewhat calloused toward individual differences—at least in an administrative sense.

In the final analysis, differentiation of instruction must be made in the classroom by a well-prepared teacher concerned primarily with the developmental needs of individuals. Many signs point to the beginning of a qualitative era, which undoubtedly will result in a reorganization of the elementary school to meet pupil needs (3, p. 21).

The fact of individual differences argues for pluralism and for an enlightened opportunism in the materials and methods of instruction. Early in this paper it was attested, rather off-handedly, that no single ideal sequence exists for any group of children. The conclusion is to be drawn from that assertion is not that it is impossible to
put together a curriculum that would satisfy a group of children or a cross-section of children. Rather, it is that if a curriculum is to be effective in the classroom it must contain different ways of activating children, different ways of presenting sequences, different opportunities for some children to "skip" parts while others work their way through, different ways of putting things. A curriculum, in short, must contain many tracks leading to the same general goal (4, p. 334).

The reading problem is generally recognized and teachers of reading are attacking on all fronts to raise the general level of reading ability. This is most helpful—but few have recognized that much help and encouragement could be given slow readers by presenting reading materials at reading level in all areas of the curriculum—not just when reading is being taught.

Evidence of the effect of reading programs upon scholastic improvement is scanty. There have been relatively few research reports on this question. Of these, only one study using control groups reported significant gains in academic grades for students in the reading classes. Perhaps this lack of evidence of the effectiveness of reading courses may be attributed to (1) the fact that some of the students who are enrolled in such classes do not take them very seriously, and (2) the fact that many courses of this type do not offer sufficient instruction and practice in reading methods suitable to each subject (18, p. 28).

The Philadelphia Reading Program emphasis on helping slow readers is only a single example of the many such programs which are being carried out over the country.

As the program developed, the reading teachers spent more and more of their time helping the teachers of regular classes to improve the reading instruction in all subjects. In accord with the developmental concept
of reading, several senior high schools have introduced courses in advanced reading and study skills for college preparatory students and courses in developmental reading for all students, in addition to the already existing courses in remedial work for retarded readers. At least three senior high schools have organized in-service training programs involving all the members of the professional staff, including counselors, librarians, and department heads. These programs have the triple purpose of (1) improving reading skills by means of instruction in the various content areas, (2) improving achievement in the content areas by means of instruction in appropriate reading and study skills, and (3) improving general school adjustment by helping students experience the satisfactions that accompany greater scholastic success. Bibliographies and study guides for the professional staff have been prepared and distributed, special faculty meetings have been planned, and expert consultant service has been secured (18, p. 36).

This brief description of an evolving reading program illustrates three important trends: toward continuity in reading instruction from kindergarten to college; toward integration of reading instruction with the teaching of every subject; and toward development of a comprehensive, whole-school program which serves the needs of students representing a wide range of reading potential and proficiency. This program also accents the importance of providing for the continuous growth of teachers, administrators, counselors, librarians, and the reading specialists themselves (18, p. 37).

Many evolving reading programs consist of grading pupils into more-or-less homogeneous reading groups and challenging each group according to its ability. This is fine—that there is emphasis on reading ability during a 20 or so minute reading period. However, what about the many more minutes in the school day when the learner must face constantly reinforced nonachievement in his other class subjects because his reading ability has not been taken into account? It is
recognized that the fundamental problem of different reading ability levels should be constantly attacked by special reading sessions. But aren't these special sessions usually supplied with materials that students can handle at their particular reading level? Why then, should we not demand that other curriculum materials be written at reading level? Shall we let children live with the vague notion that they learn to read only in reading class?

Most of the reports on this method (Joplin Plan) have consisted, to date, of unevaluated descriptions. Convincing experimental evidence of the effectiveness of individualized reading is lacking. Jackson compared two programs of individualized reading with a conventional reading program that used basic and study-type readers in grades 3, 4, 5 and 6. One experimental group used the individualized procedure for 40 percent of the total three hundred minutes of reading time, another, for 70 percent. The results were inconclusive. Kaar, working with third grade children, found that the individualized approach did not produce better results on standardized reading tests than did the more usual combination of group instruction plus some individual help. However, the teachers were enthusiastic about the individualized procedure and believed that the children read more books and wasted less time than with previous group methods. Jenkins reported more favorable results for the individualized reading group—an average growth in vocabulary of 1.96 years as compared with 1.09 years for the control group. The averages in total reading gains were 1.14 years for the control group and 1.41 for the individualized reading group (18, p. 42).

Other studies favorable to this type of program are those by Bernard and by Hart. Bohnhorst concluded that group instruction with basal readers and individualized instruction with self-selection may vary in effectiveness with pupils of different reading abilities. Teachers were also found to differ in their judgments regarding the relative effectiveness of the two kinds of programs. Enthusiastic
support of "self-selection" in reading is given in Bulletin No. 29 of the Association for Childhood Education. Articles are written from the standpoint of the reading consultant, the principal and the teacher. Parents' responses to this approach are reported as very favorable. Many values are suggested. The individualized procedure described here should not be confused with free recreational reading. Skills are taught as needed and group experiences are included (18, p. 42).

Since reading is involved in every subject, it is inevitably part of every teacher's work. Yet many teachers ask questions such as: How can I incorporate the teaching of reading into my biology (or other subject) classes? Should I neglect the teaching of biology for the teaching of reading? What can a teacher do in a classroom where students of widely varied ability have to use the same textbook, which is too difficult for many of them? Unless teachers make reading instruction an intrinsic part of their teaching, students are likely to waste their study time. When this happens, the teachers become frustrated; they feel students cannot read (18, p. 79).

Another facet, not under study in this evaluation, but certainly worthy of mention, is a need which is at least partially met by the use of diverse science materials. That is the need of "pure enjoyment" of reading. "But science makes another contribution to reading that too few teachers have recognized, and that is 'pure enjoyment!'" (5, p. 128). This remark should be coupled with the reports by Kingsley (9, p. 113-118) concerning an experiment in individualized reading.

It was carried out in Bellingham, Washington in a sixth grade where it was discovered that only six of a particular group of 27 were reading at sixth grade level. The other scores were spread out fairly evenly from 2.8 to 8.8. The reading emphasis was based on the philosophy
that if children enjoy reading, they will read. These students were allowed to choose their own books rather than being assigned texts. Each child had an individual conference with the teacher about his book and teachers were very happy with the results.

**Importance of Class Discussion**

Again, the experimental evidence investigating the value of class interactions in science classes is lacking. Just the same it is recommended by authorities.

For a thorough understanding of bodies of knowledge discussion is an important tool. A greater depth and degree of understanding characterizes knowledge gained in this way (16, p. 430).

There is another reason why discussion of a certain kind can be a valuable means to developing an understanding of the processes of biological science. It is that discussion can utilize the "energy of wanting" in the pursuit of educational goals (16, p. 430).

Discussion is especially appropriate as a means of utilizing the energy of wanting because it can tap motivating forces. It provides something in its initial phases to which the energy of the student can attach. This is true because the students' desire to like and be liked can be directed toward the teacher if the initial phase of discussion is conducted properly. In later phases of discussion this energy originally attached to the teacher as a person, can be transferred, by appropriate discussion techniques, to the desire to want and to practice the qualities which the teacher as an educated person exemplifies (16, p. 431).
Need for Student Experimentation

The third feature of the materials to be tested is the result of the approach to student involvement in scientific experimentation. It is believed that the best reinforcement of basic learnings and concept formation is to be achieved by involving the student in both individual and group science experimentation experiences. The philosophy of the materials under study here seem to reflect the trends in science teaching as reported in the National Society for the Study of Education, 59th Yearbook.

1. A movement away from teacher demonstrations and toward pupil experimentation and problem-solving.

2. A trend away from the requirement that all pupils in a science class perform the same experiments in the same period and toward the practice of using a variety of experiments and projects performed at the same time by individuals and small groups (11, p. 230).

The following statement sums the trends of thinking among many educators today toward the value of the laboratory experiences of students.

If a teacher agrees with these statements, he will plan for the laboratory phase to occur early in the teaching-learning sequence. It will generally precede the teacher telling, describing, or the textbook reading phase (13, p. 31).

The philosophy of laboratory experimentation is strengthened by the report of a study on retention of science subject matter with college students. Students were drilled on the concept of density by
numerical problems and laboratory demonstrations and were given further problem solving situations. After three weeks the same students were given numerical problems on density and they got them 100 percent right. After a six weeks time elapse, 4200 students were still able to get 92 percent of the numerical problems correct, but only 17 percent of these could prove the solution in the laboratory without help. It was discovered that 85 percent of this 17 percent had had high school laboratory experience (6, p. 103). It is easy to grasp the significance of this even though the percentage which had had high school laboratory experience and still were unable to prove the numerical problems in the laboratory was not reported.

Betts (3, p. 46) recommends that the teacher has a greater part to play than simply making science materials available to students. Based on results for 20 fifth and sixth grade students of high ability who (1) observed and predicted, (2) asked questions which could be answered yes or no and (3) manipulated science concept demonstration materials, the author found significant concept development only in one of four concepts tested, action-reaction. There was none for concepts of displacement, inertia or depth-pressure relationships. He states, "there needs to be a specific focusing of attention on the relationships involved."

Atkin (1, p. 46) hypothesizes similar action. "Indeed, it does not seem crucial to teach the children to invent concepts, because
they can and do invent concepts readily. The educational problem, rather, is to teach children to carry out their creative thinking with some intellectual discipline.

Another study used four fifth and four sixth grades in each of four school districts. Half of the classes in each district were taught by one teacher in self-contained classrooms except for music, art and physical education. The other half of the classes in each district were taught by special teachers in science and mathematics. The object was to determine if gain scores between the two groups were significant. It was found there were no significant differences for mathematics. There were significant differences in most situations for science, but there were no significant preferences among children for several teachers as opposed to only one. There were no significant differences among children in ability to solve new problems at the end of their experiences in the classrooms. The special teachers had more than twice the amount of college course preparation in sciences as teachers in self-contained classrooms. The significant results could have been due to this fact rather than to the room organization (8, p. 582-583).

In a review of science textbooks currently used in elementary schools Piltz (14, p. 277) states, "Another major lack in presently available text series is the scarcity of suggestions for individual investigations and discovery."
The commonness of this situation is being overcome by the emphasis on "open-ended" type experiments in many new curriculum materials. Taba (19, p. 19) defines an open-ended question as, "One which sets a focus that is perceptible to students and which permits spontaneity and participation in the search by students of different abilities by allowing individuals to respond in terms of their own perceptions instead of pursuing the line of thought and doing it on a level established by the teacher."

In paraphrase, an open-ended experiment could be: One which sets a direction that is perceptible to students and which permits spontaneity and participation in its pursuit by students of different abilities by allowing individuals to respond in terms of their own perceptions instead of pursuing the line of thought and verbal or printed directions on a level established by a teacher or a manual.

**Summary**

In general, writers emphasize the importance of students being provided with materials at their own reading ability levels. However, they also point out that this may do things for the student other than just increase his achievement. One writer indicates that the "joy of reading" may be increased when students read at their appropriate ability levels. Both discussion and experimentation are emphasized because of the student participation in these activities. These same
reasons are tied into the philosophy of the development of the materials under study here and the results ascertained from the investigation are important to those who would produce science materials for the elementary schools.
THE STUDY

This investigation was designed to determine the significant relationship between mean class gain scores on an achievement test, developed by the researcher, for three basic features of the Earth's Atmosphere Laboratory published by Science Research Associates. The three evaluated features were: (1) the multi-level reading approach; (2) the unique technique for class discussion; and (3) the individual and group experimentation experiences.

It was carried out over a period of approximately six weeks, beginning the last week in January 1964.

Experimental Design

The experimental design was set up to use five fifth grade classes in each of six different treatment groups. The conditions which set the boundaries of the treatment groups are described in Table 2.

No attempt was made to match any particular room with any particular treatment. There was, however, an attempt to scatter the treatments among schools and school districts. This was done without assigning the same treatment in any two rooms in the same building, with three exceptions. Rooms 221 and 223 were in the same school and both used treatment two. Rooms 251 and 253 were in the same school and both used treatment five. (These exceptions occurred
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Number of Classes</th>
<th>Symbol</th>
<th>Description of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>RDX (C)</td>
<td>All students read only the middle (C) level of the five reading levels supplied in a normal laboratory. Research booklets at reading levels A, B, D and E were removed from the laboratories. The students discussed the materials and carried out the experiments.</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>RDX (A-C-E)</td>
<td>Research booklets for reading levels B and D were removed from these laboratories. Students at grade-equivalent 0.00-4.9 read level A; students at grade-equivalent 5.0-5.9 read level C; and students at grade-equivalent 6.0-10.0 read level E research booklets. The students discussed the materials and carried out the experiments.</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>RDX (A-E)</td>
<td>Students at each of the five designated grade equivalents read research booklets at their appropriate level. The students discussed the materials and carried out the experiments.</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>RX (A-E)</td>
<td>Students at each of the five designated grade equivalents read research booklets at their appropriate level. The teachers were not given the Teacher's Instructional Aid Booklet (for discussion) and teachers were instructed to have no class discussion of the materials. The students did carry out the experiments.</td>
</tr>
<tr>
<td>Treatment</td>
<td>Number of Classes</td>
<td>Symbol</td>
<td>Description of treatment</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------</td>
<td>-------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>RD(A-E)</td>
<td>Students at each of the five designated grade equivalents read research booklets at their appropriate level. The students discussed the materials but they did none of the recommended experiments. The part of the materials with drawings and suggestions for experiments were masked so teachers and students were not able to see them.</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>Test-re-test</td>
<td>These classes never did see any of the materials.</td>
</tr>
</tbody>
</table>

R = student reads research booklets.

D = student discusses, with his class, the "big idea" according to the Teacher's Instructional Aid Booklet.

X = student carries out both the home activity and the classroom experiment.

A, B, C, D, E, = designated reading level.
because it was necessary for them to share materials.)

Within each treatment group all students were identified into one of five reading ability levels: A (lowest), B (next lowest), C (average), D (next highest) and E (highest), as discussed on page 13 and shown on Table 1.

An analysis of variance technique was applied to the gain scores between a criterion pretest and post test designed to measure achievement. The influence of treatments and reading levels on achievement gain scores was determined for significance at the .05 (or better) level.

Significance of treatment and reading level on achievement gain scores was also determined for small groups of questions which were unique to each of the five different reading levels and on small groups of questions which fell into one of six categories described on page 42.

All these determinations were grouped under one of three headings:

1. Influence of reading levels.
2. Influence of discussion.
3. Influence of experiments.

School Contacts

First contacts with schools were through the administrative officers in October 1963. A meeting with all of their fifth grade teachers...
was arranged and the philosophy of the Learnings in Science Laboratories was discussed and teachers were shown samples of the laboratories. The design of the experiment was explained and their cooperation was asked. Teachers were given two or three days to consider whether or not they wished to participate. At the end of this consideration period, they unanimously agreed to help carry out the study.

They were cautioned not to use any science materials which concerned atmosphere either before or during the time the study was made ready.

Another meeting was held with the teachers in late January 1964. At this time teachers were given written instructions (Appendix II) about the way they should conduct their classes according to the treatment group to which they were assigned. Questions were raised and answered. During the same meeting the following materials were distributed: (1) Learnings in Science Laboratories to individual teachers or to designated pairs of teachers for sharing, where feasible; (2) a student record book for each student; (3) accessory kits of laboratory experiment materials to be shared between two rooms, and (4) the criterion test along with standard answer sheets.

The teachers were not given an instructional workshop on how to use the unique materials. The research was intended to get a measurement of the results as they would probably appear in a typical field condition since most teachers will not have the benefit of an instructional workshop.
The study used 29 heterogeneous, fifth-grade classrooms distributed among three mid-Willamette Valley, Oregon school districts. All the heterogeneous fifth grades in the three districts were used. The study was designed to use 30 classes, but quite late in the planning it was discovered that three classes in one school were homogeneously grouped. Consequently, the experiment was moved to another school in the same district which had only two heterogeneous classes of fifth graders; hence, the study involved 29 fifth-grade classrooms rather than the intended 30.

There was a total of 677 students involved in the study. Information on the mental ability and reading levels for students in each class was obtained from school records. Twenty-seven classes were given the Metropolitan Reading Achievement Tests in September 1963. The other two classes used the California Reading Achievement Tests given in October 1963. Since there was an approximate delay of nearly five months between the time of the testing and the use of the study materials, students were assigned to reading groups one-half year above their grade-equivalent scores as indicated on Table 3.
Table 3. Number of Students Assigned in Different Reading Groups

<table>
<thead>
<tr>
<th>Reading level</th>
<th>Grade-equivalent reading level of the materials</th>
<th>Grade-equivalent reading level of the student</th>
<th>Total students</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3-4</td>
<td>0.0 - 3.9</td>
<td>139</td>
</tr>
<tr>
<td>B</td>
<td>4-5</td>
<td>4.0 - 4.9</td>
<td>136</td>
</tr>
<tr>
<td>C</td>
<td>5-6</td>
<td>5.0 - 5.9</td>
<td>142</td>
</tr>
<tr>
<td>D</td>
<td>6-7</td>
<td>6.0 - 6.9</td>
<td>118</td>
</tr>
<tr>
<td>E</td>
<td>7-8</td>
<td>7.0 -10.0*</td>
<td>142</td>
</tr>
</tbody>
</table>

*The highest range of the test.

Table 4 shows the distribution of the student population. Each section shows statistics for the classes within a treatment. Included are the classes by code number, the mean intelligence quotient for each class and for each treatment, the total number of students in each of the five reading levels in each class and each treatment, and the mean reading level for each class and each treatment.

Intelligence quotient scores were taken from three sources. (1) Twelve classes used the Otis Quick Scoring Mental Ability Test, form B, given in April 1963; (2) 15 classes used the Science Research Associates Primary Mental Abilities Test, Grades 4-6, given in January 1964; and (3) two classes used the California Mental Maturity Test given in March, 1964. The standard deviation range of these three tests is similar.

In almost every class the range of reading ability was from
Table 4. Mean Intelligence Quotient, Reading Level and Number of Students at Each Reading Level for Treatments and Classes

<table>
<thead>
<tr>
<th>Treatment Class</th>
<th>Mean L.Q.</th>
<th>Number of students at each reading level*</th>
<th>Total equivalent reading level</th>
<th>Average grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1</td>
<td>100</td>
<td>7 8 2 1 2 20 4.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 2</td>
<td>111</td>
<td>4 3 5 2 8 22 5.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 3</td>
<td>102</td>
<td>5 6 8 3 5 27 5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 4</td>
<td>110</td>
<td>5 6 5 1 8 25 4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 5</td>
<td>109</td>
<td>3 2 5 8 5 23 6.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>107</td>
<td>24 25 25 15 28 117 Mean 5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 1</td>
<td>102</td>
<td>8 4 2 5 1 20 4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 2</td>
<td>103</td>
<td>7 5 5 4 5 26 4.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 3</td>
<td>107</td>
<td>4 3 7 3 4 21 5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 4</td>
<td>106</td>
<td>1 7 6 5 1 20 5.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 5</td>
<td>112</td>
<td>2 6 4 2 8 22 5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>106</td>
<td>22 25 25 19 19 109 Mean 5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 1</td>
<td>106</td>
<td>2 3 7 3 4 19 5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 2</td>
<td>103</td>
<td>9 5 5 5 3 27 5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 3</td>
<td>101</td>
<td>12 4 2 2 3 23 4.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 4</td>
<td>108</td>
<td>2 4 3 6 5 20 6.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 5</td>
<td>108</td>
<td>4 6 4 3 5 22 4.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>105</td>
<td>29 22 21 19 20 111 Mean 5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 1</td>
<td>105</td>
<td>5 4 7 4 6 26 5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 2</td>
<td>109</td>
<td>3 2 5 6 9 25 6.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 3</td>
<td>102</td>
<td>11 10 2 3 2 28 4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 4</td>
<td>106</td>
<td>4 8 5 9 1 27 5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 5</td>
<td>112</td>
<td>3 3 7 1 9 23 5.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>105</td>
<td>26 27 30 19 27 129 Mean 5.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 1</td>
<td>108</td>
<td>3 2 7 4 7 23 5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 2</td>
<td>105</td>
<td>6 3 5 6 3 23 5.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 3</td>
<td>112</td>
<td>4 5 4 8 7 28 5.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 4</td>
<td>107</td>
<td>5 3 5 9 5 27 6.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>109</td>
<td>18 13 21 27 22 101 Mean 5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 1</td>
<td>107</td>
<td>3 6 4 4 3 20 5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 2</td>
<td>100</td>
<td>8 5 4 1 1 19 4.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 3</td>
<td>108</td>
<td>4 4 5 5 6 24 5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 4</td>
<td>106</td>
<td>4 4 4 3 7 22 5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 5</td>
<td>117</td>
<td>1 5 4 6 9 25 5.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>108</td>
<td>20 24 21 19 26 110 Mean 5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>139</td>
<td>136 142 118 142 677</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*A - Lowest reading ability level. Grade-equivalent level 3-4.

B - Highest reading ability level. Grade-equivalent level 7-8.
nonreaders to those reading at the tenth grade level or above. Table 5 is a summary of these same mean scores by treatment groups and it is to be noted how similar the mean Intelligence Quotient scores are and also the mean reading ability levels for each treatment group.

Table 5. Mean Intelligence Quotients and Grade-Equivalent Reading Scores for the Six Treatments

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Mean I. Q.</th>
<th>Mean grade-equivalent reading score</th>
<th>Total students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>107</td>
<td>5.2</td>
<td>117</td>
</tr>
<tr>
<td>2</td>
<td>106</td>
<td>5.1</td>
<td>109</td>
</tr>
<tr>
<td>3</td>
<td>105</td>
<td>5.1</td>
<td>111</td>
</tr>
<tr>
<td>4</td>
<td>105</td>
<td>5.3</td>
<td>129</td>
</tr>
<tr>
<td>5</td>
<td>109</td>
<td>5.5</td>
<td>101</td>
</tr>
<tr>
<td>6</td>
<td>108</td>
<td>5.2</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>677</td>
</tr>
</tbody>
</table>

Criterion Test

Purpose

The criterion test (Appendix I) was given first as a pretest and later as a post test in order to obtain a gain score indicative of achievement.

Questions were devised from one of three sources: (1) directly from the sets of questions which were printed at the end of the reading selections; (2) directly from the reading selection materials; and (3) others were developed in such a manner that the answers had to be
inferred from the information or situations presented in the reading selections.

**Question Selection**

The questions were developed over a period of time, tried out on a few fifth grade classes, then were accepted, reworded, discarded or substitutions made until 118 questions had been tried with students in 15 classes not used in the study.

Taking into account the above described tryouts, 60 questions (Table 6) were selected for the final criterion test. An attempt was made to select questions for difficulty which would show neither a "ceiling" nor "flooring" effect in the total score.

The phi coefficient score ($\phi$) estimates a Pearson product-moment coefficient of correlation where both variables are expressed dichotomously. A value of .20 is considered satisfactory. The mean phi coefficient for this test was .32. The values were read off a table reported by Flanagan (7, p. 298).

Concern for validity required that questions be well distributed among the five big ideas, the five reading levels and the six special categories of questions used in the analysis. Tables 7 and 8 show these distributions.

In Table 8, the total figures on the right indicate the number of questions, from a total of 60, which a student reading at any one
<table>
<thead>
<tr>
<th>Q</th>
<th>U</th>
<th>L</th>
<th>( \phi )</th>
<th>G</th>
<th>Q</th>
<th>U</th>
<th>L</th>
<th>( \phi )</th>
<th>G</th>
<th>Q</th>
<th>U</th>
<th>L</th>
<th>( \phi )</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>47</td>
<td>40</td>
<td>.07</td>
<td>50</td>
<td>21</td>
<td>47</td>
<td>41</td>
<td>.06</td>
<td>30</td>
<td>41</td>
<td>100</td>
<td>56</td>
<td>.70</td>
<td>61</td>
</tr>
<tr>
<td>2</td>
<td>73</td>
<td>57</td>
<td>.16</td>
<td>46</td>
<td>22</td>
<td>75</td>
<td>37</td>
<td>.39</td>
<td>34</td>
<td>42</td>
<td>87</td>
<td>25</td>
<td>.62</td>
<td>72</td>
</tr>
<tr>
<td>3</td>
<td>87</td>
<td>37</td>
<td>.54</td>
<td>57</td>
<td>23</td>
<td>79</td>
<td>30</td>
<td>.50</td>
<td>15</td>
<td>43</td>
<td>50</td>
<td>19</td>
<td>.24</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>25</td>
<td>.83</td>
<td>81</td>
<td>24</td>
<td>50</td>
<td>19</td>
<td>.34</td>
<td>42</td>
<td>44</td>
<td>53</td>
<td>12</td>
<td>.47</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>53</td>
<td>.13</td>
<td>100</td>
<td>25</td>
<td>69</td>
<td>50</td>
<td>.20</td>
<td>12</td>
<td>45</td>
<td>88</td>
<td>79</td>
<td>.13</td>
<td>62</td>
</tr>
<tr>
<td>6</td>
<td>84</td>
<td>40</td>
<td>.42</td>
<td>-5</td>
<td>26</td>
<td>75</td>
<td>37</td>
<td>.38</td>
<td>315</td>
<td>46</td>
<td>30</td>
<td>5</td>
<td>.43</td>
<td>52</td>
</tr>
<tr>
<td>7</td>
<td>57</td>
<td>31</td>
<td>.27</td>
<td>57</td>
<td>27</td>
<td>32</td>
<td>14</td>
<td>.25</td>
<td>0</td>
<td>47</td>
<td>20</td>
<td>3</td>
<td>.41</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>37</td>
<td>0</td>
<td>.66</td>
<td>184</td>
<td>28</td>
<td>100</td>
<td>25</td>
<td>.83</td>
<td>27</td>
<td>48</td>
<td>25</td>
<td>0</td>
<td>.58</td>
<td>38</td>
</tr>
<tr>
<td>9</td>
<td>98</td>
<td>83</td>
<td>.46</td>
<td>5</td>
<td>29</td>
<td>50</td>
<td>19</td>
<td>.34</td>
<td>100</td>
<td>49</td>
<td>47</td>
<td>38</td>
<td>.10</td>
<td>74</td>
</tr>
<tr>
<td>10</td>
<td>87</td>
<td>25</td>
<td>.62</td>
<td>49</td>
<td>30</td>
<td>65</td>
<td>29</td>
<td>.37</td>
<td>33</td>
<td>50</td>
<td>87</td>
<td>62</td>
<td>.32</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>16</td>
<td>12</td>
<td>.07</td>
<td>88</td>
<td>31</td>
<td>37</td>
<td>0</td>
<td>.66</td>
<td>100</td>
<td>51</td>
<td>63</td>
<td>52</td>
<td>.12</td>
<td>67</td>
</tr>
<tr>
<td>12</td>
<td>80</td>
<td>50</td>
<td>.33</td>
<td>6</td>
<td>32</td>
<td>75</td>
<td>47</td>
<td>.30</td>
<td>60</td>
<td>52</td>
<td>83</td>
<td>36</td>
<td>.49</td>
<td>21</td>
</tr>
<tr>
<td>13</td>
<td>100</td>
<td>37</td>
<td>.78</td>
<td>91</td>
<td>33</td>
<td>45</td>
<td>33</td>
<td>.13</td>
<td>0</td>
<td>53</td>
<td>58</td>
<td>73</td>
<td>- .17</td>
<td>66</td>
</tr>
<tr>
<td>14</td>
<td>75</td>
<td>25</td>
<td>.50</td>
<td>22</td>
<td>34</td>
<td>43</td>
<td>15</td>
<td>.33</td>
<td>100</td>
<td>54</td>
<td>75</td>
<td>45</td>
<td>.31</td>
<td>53</td>
</tr>
<tr>
<td>15</td>
<td>65</td>
<td>29</td>
<td>.36</td>
<td>0</td>
<td>35</td>
<td>53</td>
<td>33</td>
<td>.21</td>
<td>0</td>
<td>55</td>
<td>75</td>
<td>45</td>
<td>.31</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>44</td>
<td>19</td>
<td>.28</td>
<td>45</td>
<td>36</td>
<td>25</td>
<td>12</td>
<td>.19</td>
<td>200</td>
<td>56</td>
<td>80</td>
<td>64</td>
<td>.19</td>
<td>11</td>
</tr>
<tr>
<td>17</td>
<td>40</td>
<td>7</td>
<td>.45</td>
<td>0</td>
<td>37</td>
<td>34</td>
<td>10</td>
<td>.34</td>
<td>9</td>
<td>57</td>
<td>83</td>
<td>18</td>
<td>.64</td>
<td>58</td>
</tr>
<tr>
<td>18</td>
<td>45</td>
<td>33</td>
<td>.13</td>
<td>27</td>
<td>38</td>
<td>45</td>
<td>19</td>
<td>.29</td>
<td>118</td>
<td>58</td>
<td>25</td>
<td>9</td>
<td>.27</td>
<td>50</td>
</tr>
<tr>
<td>19</td>
<td>25</td>
<td>12</td>
<td>.20</td>
<td>30</td>
<td>39</td>
<td>22</td>
<td>12</td>
<td>.16</td>
<td>20</td>
<td>59</td>
<td>33</td>
<td>21</td>
<td>.15</td>
<td>43</td>
</tr>
<tr>
<td>20</td>
<td>88</td>
<td>50</td>
<td>.42</td>
<td>0</td>
<td>40</td>
<td>50</td>
<td>0</td>
<td>.72</td>
<td>61</td>
<td>60</td>
<td>57</td>
<td>15</td>
<td>.46</td>
<td>39</td>
</tr>
</tbody>
</table>

Q = question number

U = percentage of students in the upper half who answered correctly on post test

L = percentage of students in the lower half who answered correctly on post test

\( \phi \) = phi coefficient score

G = percentage gain score from pretest to post test
Table 7. Number of Questions for Each Big Idea

<table>
<thead>
<tr>
<th>Big idea</th>
<th>Number of questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>12</td>
</tr>
<tr>
<td>II</td>
<td>12</td>
</tr>
<tr>
<td>III</td>
<td>12</td>
</tr>
<tr>
<td>IV</td>
<td>14</td>
</tr>
<tr>
<td>V</td>
<td>10</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>60</strong></td>
</tr>
</tbody>
</table>

Table 8. Number of Questions at Each Reading Level by Categories of Questions

<table>
<thead>
<tr>
<th>Reading level</th>
<th>Factual-unique</th>
<th>Conceptual-unique</th>
<th>Factual-common</th>
<th>Conceptual-common</th>
<th>Conceptual-Conceptual-common</th>
<th>Conceptual-Conceptual-none</th>
<th>Discussion</th>
<th>Experiments</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>3</td>
<td>12</td>
<td>3</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>3</td>
<td>12</td>
<td>3</td>
<td>18</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>3</td>
<td>12</td>
<td>3</td>
<td>18</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>39</td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>3</td>
<td>12</td>
<td>3</td>
<td>18</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>41</td>
</tr>
<tr>
<td>E</td>
<td>3</td>
<td>2</td>
<td>[12*]</td>
<td>3*</td>
<td>[18*]</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>[40]</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11</strong></td>
<td><strong>14</strong></td>
<td><strong>12</strong></td>
<td><strong>3</strong></td>
<td><strong>18</strong></td>
<td><strong>2</strong></td>
<td><strong>2</strong></td>
<td><strong>2</strong></td>
<td><strong>60</strong></td>
</tr>
</tbody>
</table>

* Questions common to all reading levels, therefore each is a total.

** All questions in this category also appeared under one of the other six categories. They were used for special analysis.

1. **Factual-unique.** Those which call for recall of a fact and are unique to one of the five reading levels.

2. **Conceptual-unique.** Those which require a response which shows understanding or application of a concept and are unique to one reading level.

3. **Factual-common.** Factual-type which appear in all five reading levels.

4. **Conceptual-common.** Concept-type which appear in all five reading levels.

5. **Conceptual-none.** Conceptual-type, the answer for which does not appear in any of the reading levels and which depends upon an understanding of, or application to, related facts or concepts.

6. **Discussion.** Those which were taken from the teacher's instructional aid booklet and which were specifically covered in class by discussion.
reading level had an opportunity to answer. The balance of the questions for each reading level is represented by those unique to the other reading levels. Answers to the latter could have been learned through class discussion.

Test Reliability

Test reliability coefficient score was .88. It was determined by the split-half method on the post test scores of the 111 students in treatment group three which used the materials as recommended by the authors. The reliability of the extended test was determined by the Spearman-Brown formula (21, p. 15). This correction is represented by the formula,

\[ r_{xy} = \frac{2r}{1 + r} \]

where:  
- \( r_{xy} \) = reliability coefficient of the extended test  
- \( r \) = obtained coefficient of the split-half test.

Pretest

Teachers gave the criterion test in two 20-minute sessions on Thursday and Friday, January 23 or 24, 1964. All tests and answer sheets were collected on Monday, January 27, except for absentees for whom teachers held the test materials. These were completed and mailed to the researcher as soon as they were completed.
Teachers began using the materials in their classrooms during the week of January 27, 1964. They were asked to try to time their use of the materials so they could finish during the first week of March. This afforded them about six weeks in which to complete their particular treatment of the materials.

**Classroom Visits**

Each classroom was visited at least three times during the six-week period. Because of schedule conflicts, it wasn't always possible to see the classes using the materials but short visits with the teachers were made during which time they had an opportunity to discuss the use of the materials. In addition, they were contacted as a group by letters (Appendix III) giving encouragement, additional information and answers to questions about problems of administration of the materials.

**Post Test**

Post tests of the criterion test, along with answer sheets, were distributed to teachers during the week of March 2 as they finished with the instructional materials. Testing was completed in all classes by March 11, except for a few absenteeees who were given the post test during the week of March 23 following spring vacation. Results on all students who missed more than five days of class time during the
six-week period were discarded. The results for students who transferred in or out of the district during the time of the experiment were also discarded.

Scoring

The answer sheets for both the pretest and post test were sent to Science Research Associates of Chicago, Illinois to be machine scored. A right or wrong score for each of the 60 items on the criterion test was punched on to IBM cards. There was a card for the pretest and one for the post test for each of the 677 students. The cards also had student identifying data on them.

A short Fortran program was written for the IBM 1620 computer which computed the gain score between the pretest and post test for each student. The gain score was punched and assembled into a data deck. Similar programs were written to assemble data decks for the other special analyses.

Analyses

The gain score data were submitted to an analysis of variance program written for disproportionate subclass frequencies. It was composed at the University of California, Riverside, California November 10, 1963. Analyses were computed for significant differences at the .05 (or better) level. Table 9 shows a schematic of all
the possible comparisons of gain score data.

Table 9. Schematic of Analyses.

<table>
<thead>
<tr>
<th>Source</th>
<th>Treatments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions for reading levels A-E</td>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Question unique to reading level:</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
</tr>
<tr>
<td>Questions in category:</td>
<td></td>
</tr>
<tr>
<td>Factual-unique</td>
<td></td>
</tr>
<tr>
<td>Conceptual-unique</td>
<td></td>
</tr>
<tr>
<td>Factual-common</td>
<td></td>
</tr>
<tr>
<td>Conceptual-common</td>
<td></td>
</tr>
<tr>
<td>Conceptual-none</td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td></td>
</tr>
</tbody>
</table>

Selected analyses were made and other results were interpolated or extrapolated from them. These are marked with an asterisk on the results tables. Where reliable interpolations or extrapolations could not be made, further analyses were run. These are shown in the next section.
Teacher Questionnaire

A short questionnaire (Appendix IV) to survey the teachers' opinions concerning the materials was designed to be completed by the teachers at the conclusion of the investigation. Teachers were not forewarned that they would be expected to do this. Results of the survey are reported in the next section.
PRESENTATION AND INTERPRETATION OF DATA

As described on page 31 the study was designed to determine if student achievement gain scores between pretest and post test were significantly affected when students used: (1) science materials which were written at several different reading ability levels; (2) science materials written at several different reading ability levels but did not discuss them with their classmates; and (3) science materials written at several different reading ability levels but did not perform the experiments.

This was accomplished by forming six different treatment groups and using an analysis of variance technique on mean gain scores as tabulated in Table 10.

Mean Gain Score Data by Class, Treatment and Reading Level

Table 10 shows the mean gain score for each class within a treatment and also for each treatment. These scores represent the mean number of points gained between the pretest and post test. It is noted that they vary considerably by classes. In Treatment 1 the mean gain scores range from 5.75 to 12.35. Similar ranges are seen in the other treatment groups. There is no consistent correlation between the mean gain score and the intelligence quotient as shown on Table 4. The least variance among scores is among the classes in
### Table 10. Class and Treatment Mean Gain Scores by Treatments

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Class</th>
<th>Class mean gain score</th>
<th>Mean treatment gain score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students read only at C (average) reading level. Discussed materials and performed experiments.</td>
<td>1</td>
<td>6.03</td>
<td>8.67</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>12.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5.75</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students read respectively at A (lowest), C (average) and E (highest) reading levels. Discussed materials and performed experiments.</td>
<td>1</td>
<td>5.37</td>
<td>7.93</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>11.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>12.15</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students read respectively at A, B, C, D or E reading levels. Discussed materials and performed experiments.</td>
<td>1</td>
<td>6.48</td>
<td>8.14</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>13.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>10.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4.79</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students read respectively at A, B, C, D or E reading levels. Did not discuss materials but did perform experiments.</td>
<td>1</td>
<td>9.92</td>
<td>6.87</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5.81</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students read respectively at A, B, C, D or E reading levels. Discussed materials, but did not perform experiments.</td>
<td>1</td>
<td>1.94</td>
<td>3.32</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3.36</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students did not use any of the materials.</td>
<td>1</td>
<td>1.75</td>
<td>2.45</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.93</td>
<td></td>
</tr>
</tbody>
</table>
Treatment 4. In Treatment 5 there was one very low score of 0.49.

In Treatment 6 class number four outscored two of the classes in Treatment 5, one of the classes in Treatment 3 and one of the classes in Treatment 2. Several factors may be responsible for these extremes, one being the influence of the teacher. Undoubtedly some teachers were more enthused in the area of science than others.

Table 11 shows that except for comparisons of Treatments 3 and 4, the mean scores for the reading levels order themselves from lowest, for the lowest reading level (A), to highest, for the highest reading level (E).

Table 11. Mean Gain Scores for Groups of Students Reading at Different Reading Ability Levels.

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean gain scores for reading level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>All six treatments</td>
<td>3.32</td>
</tr>
<tr>
<td>Treatments 1, 2 and 3</td>
<td>4.79</td>
</tr>
<tr>
<td>Treatments 3 and 4</td>
<td>5.11</td>
</tr>
<tr>
<td>Treatments 3 and 5</td>
<td>2.69</td>
</tr>
</tbody>
</table>

This is the typical achievement gain score pattern we would expect to see based upon reading ability. Later tables show that the reading level mean gain scores are variable enough that they are significant at the .05 level or better. This is true in nearly all comparisons.
Figures 2, 3, 4 and 5 present a visual representation of the gain scores by treatment and reading ability level. In Figure 2 it can be seen that the mean gain scores for the first four treatments range close together, but they drop off sharply for Treatments 5 and 6. The reading levels show a steady rise from the lowest level A to the highest level E.

Figure 3 shows a similar pattern when only the first three treatment groups are compared with the reading levels. Figure 4 shows reading level D group as the highest gain score achievers. It wasn't possible to determine if there were anything peculiar to Treatment 4 which would account for this one deviation. Figure 5 shows a return to the same pattern.

![Figure 2](image)

**Figure 2.** Mean Gain Scores for Six Treatments and for Five Reading Levels.
Figure 3. Mean Gain Scores for First Three Treatments and for Five Reading Levels.

Figure 4. Mean Gain Scores for Treatments 3 and 4 and for Five Reading Levels.
Figure 5. Mean Gain Scores for Treatments 3 and 5 and for Five Reading Levels.

Data Concerning the Influence of Materials Written at Five Reading Levels

Table 12, column one, shows the gain score differences between reading ability levels (RL) to be significant when all six treatments are considered together. The same is true for the influence of the treatment (T). If this data were considered alone, one would assume that both influences were significant and the study would be finished. However, it was carried further and column two shows that when only Treatments 1, 2 and 3 are compared (their variable involved different reading levels) the effect of the treatment is non-significant. It makes no difference on student achievement whether
students have an opportunity to read respectively, science materials written at five different reading ability levels, three different reading ability levels, or whether they all have access to materials written at only one (the average) reading ability level. The same pattern of non-significance for treatment influence shows in comparisons of Treatments 1 and 2, 1 and 3, and 2 and 3, all of which are indicated in columns three, four and five.

Table 12. Significance of Reading Level (RL) and Treatment (T) When Different Combinations of Treatments are Compared.

<table>
<thead>
<tr>
<th>Source</th>
<th>1 Treatments (1-6)</th>
<th>2 Treatments (1x2x3)</th>
<th>3 Treatments (1x2)</th>
<th>4 Treatments (1x3)</th>
<th>5 Treatments (2x3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RL T</td>
<td>RL T</td>
<td>RL T</td>
<td>RL T</td>
<td>RL T</td>
</tr>
<tr>
<td>Total</td>
<td>.01 .01</td>
<td>.01 ns</td>
<td>* .01 * ns</td>
<td>* .01 * ns</td>
<td>* .01 * ns</td>
</tr>
<tr>
<td>criterion test</td>
<td>.01 .01</td>
<td>.01 ns</td>
<td>* .01 * ns</td>
<td>* .01 * ns</td>
<td>* .01 * ns</td>
</tr>
</tbody>
</table>

* Interpolated or extrapolated value.

Column one on Table 13 shows the results of comparisons of gain scores on specific questions selected from the criterion test. These questions represent only those unique (discussed in only one of the reading level materials) to specific reading levels. Column one shows that performance on questions unique to reading levels A and B were nonsignificant for both reading level (RL) and treatment (T). The treatment influence was significant for questions unique to level
C, and both treatment and reading level influences were significant for questions unique to levels D and E. Again, however, as shown in columns two, three, four and five, when selected treatment groups are compared, the significances are very spotty and as a whole, considered nonsignificant.

Table 13. Significance of Treatments on Questions Unique to Reading Levels.

<table>
<thead>
<tr>
<th>Source</th>
<th>1 Treatments (1-5)</th>
<th>2 Treatments (1x2x3)</th>
<th>3 Treatments (1x2)</th>
<th>4 Treatments (1x3)</th>
<th>5 Treatments (2x3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions unique to reading level</td>
<td>RL T</td>
<td>RL T</td>
<td>RL T</td>
<td>RL T</td>
<td>RL T</td>
</tr>
<tr>
<td>A(5)**</td>
<td>ns ns *ns *ns *ns *ns *ns *ns *ns *ns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B(5)**</td>
<td>ns ns *ns *ns *ns *ns *ns *ns *ns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(4)**</td>
<td>ns .05 *ns .05 ns .05 *ns *ns ns ns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D(6)**</td>
<td>.05 .01 *ns *ns *ns *ns *ns *ns *ns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(5)**</td>
<td>.01 .01 *ns *ns *ns *ns .05 ns *ns *ns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Interpolated or extrapolated value.

** Number of questions as reported on Table 8.

The comparisons on Table 14 involve gain score differences on specific questions selected from the total criterion test which represent one of six groups. (See Table 12 for a presentation of the categories of questions.) When all six treatment groups were compared
together (column 1) it was found that the influence of both the reading level and the treatment were significant. However, when individual treatments were compared (columns 2 to 5) neither treatments nor reading levels were significant.

Table 14. Significances of Treatments on Questions in Special Categories.

<table>
<thead>
<tr>
<th>Source</th>
<th>1 Treatments (1-6)</th>
<th>2 Treatments (1x2x3)</th>
<th>3 Treatments (1x2)</th>
<th>4 Treatments (1x3)</th>
<th>5 Treatments (2x3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RL</td>
<td>T</td>
<td>RL</td>
<td>T</td>
<td>RL</td>
</tr>
<tr>
<td>Questions which are:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factual-unique</td>
<td>0.01</td>
<td>0.01</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Conceptual-unique</td>
<td>0.01</td>
<td>0.01</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Factual-common</td>
<td>0.01</td>
<td>0.01</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Conceptual-common</td>
<td>0.01</td>
<td>0.01</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Conceptual-none</td>
<td>0.01</td>
<td>0.01</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Discussion</td>
<td>0.01</td>
<td>0.01</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
</tbody>
</table>

* Interpolated or extrapolated value.

** Number of questions as reported on Table 8.
Data Concerning the Influence of Whether or Not Students Discuss What They Read

Table 15 shows the results of comparisons of Treatment 4, (in which the variable was nondiscussion of the materials) to Treatments 1, 2 and 3, (in which the variable involved was different reading levels). Column one shows again that the influence of the reading levels is significant, but that the treatment influence is nonsignificant. Whether or not students discuss with their classmates what they have read was found to have no significance. The reader should also refer to Figure 4 for a visual representation of the data.

Table 15. Significances of Treatment 4 Involving Class Discussion.

<table>
<thead>
<tr>
<th>Source</th>
<th>1 Treatments (4x1)</th>
<th>2 Treatments (4x2)</th>
<th>3 Treatments (4x3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RL T</td>
<td>RL T</td>
<td>RL T</td>
</tr>
<tr>
<td>Total criterion test</td>
<td>.01 ns</td>
<td>* .01 * ns</td>
<td>.01 ns</td>
</tr>
</tbody>
</table>

* Interpolated or extrapolated value.

Table 16 shows that the influence of treatments or reading levels when Treatment 4 is compared to Treatments 1, 2 or 3 have only spotty significances on questions unique to certain reading levels or on questions which fall into one of the special categories as listed. The significances are not consistent enough to make positive conclusions.
Table 16. Significances of Treatment 4 on Questions Unique to Different Reading Levels and to Questions in Six Special Categories.

<table>
<thead>
<tr>
<th>Source</th>
<th>1 Treatments (4x1)</th>
<th>2 Treatments (4x2)</th>
<th>3 Treatments (4x3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RL</td>
<td>T</td>
<td>RL</td>
</tr>
<tr>
<td>Questions unique to reading levels:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (5)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>B (5)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>C (4)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>D (6)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>E (5)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Questions which are:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factual-unique (11)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Conceptual-unique (14)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Factual-common (12)**</td>
<td>*ns</td>
<td>*ns</td>
<td>.05</td>
</tr>
<tr>
<td>Conceptual-common (3)**</td>
<td>ns</td>
<td>.01</td>
<td>*ns</td>
</tr>
<tr>
<td>Conceptual-none (18)**</td>
<td>ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>Discussion (8)**</td>
<td>.05</td>
<td>.05</td>
<td>*ns</td>
</tr>
</tbody>
</table>

* Interpolated or extrapolated value.
** Number of questions as reported on Table 8.
Data Concerning Whether or Not Students Performed Experiments

Table 17 shows the results of comparisons of the first three treatments (in which the variable involved was reading levels) with Treatment 5 (in which the variable was that students did not do any of the experiments suggested by the authors). Columns one, two and three show the adverse influence on achievement gain scores in Treatment 5 to be significant. Column four shows the same evidence when Treatment 5 is compared to Treatment 4. This is strong evidence that doing experiments is important to student achievement and that it evidently reinforced learning because in all treatments in which students performed experiments, the gain scores were significantly higher.

Table 17. Significances of Treatment 5 Involving Nonperformance of Experiments.

<table>
<thead>
<tr>
<th>Source</th>
<th>1 Treatments (5x1)</th>
<th>2 Treatments (5x2)</th>
<th>3 Treatments (5x3)</th>
<th>4 Treatments (5x4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RL</td>
<td>T</td>
<td>RL</td>
<td>T</td>
</tr>
<tr>
<td>Total criterion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>test</td>
<td>*.05</td>
<td>*.05</td>
<td>.01</td>
<td>.01</td>
</tr>
</tbody>
</table>

* Interpolated or extrapolated value.

The reader should also see Figure 5 for a visual representation of the above data.
Table 18 shows that, except for questions unique to reading levels A and B, a similar significant influence of lower achievement gain scores for Treatment 5 is generally indicated for reading levels C, D and E and for the six categories of questions.

Table 18. Significances of Treatment 5 Involving Nonperformance of Experiments on Questions Unique to Reading Levels and on Questions in Six Different Categories

<table>
<thead>
<tr>
<th>Source</th>
<th>1 Treatments (5x1)</th>
<th>2 Treatments (5x2)</th>
<th>3 Treatments (5x3)</th>
<th>4 Treatments (5x4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions unique to reading levels:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (5)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>B (5)**</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>C (4)**</td>
<td>*05</td>
<td>*05</td>
<td>*ns</td>
<td>*ns</td>
</tr>
<tr>
<td>D (6)**</td>
<td>*05</td>
<td>*05</td>
<td>01</td>
<td>05</td>
</tr>
<tr>
<td>E (5)**</td>
<td>05</td>
<td>05</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Questions which are:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factual-unique (11)**</td>
<td>05</td>
<td>01</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Conceptual-unique (14)**</td>
<td>05</td>
<td>05</td>
<td>01</td>
<td>01</td>
</tr>
<tr>
<td>Factual-common (12)**</td>
<td>05</td>
<td>05</td>
<td>05</td>
<td>01</td>
</tr>
<tr>
<td>Conceptual-common (3)**</td>
<td>ns</td>
<td>01</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Conceptual-none (18)</td>
<td>05</td>
<td>01</td>
<td>*05</td>
<td>*05</td>
</tr>
<tr>
<td>Discussion (8)**</td>
<td>05</td>
<td>05</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

* Interpolated or extrapolated value.

** Number of questions as reported on Table 8.
Data Concerning Influences of Classes within Treatments

The results in Table 19 were determined to check the influence of class means within the different treatment groups. There were significant differences for the influence of the reading levels as expected; however, there were also significant differences for the influence of class means. This was true for each treatment except for Treatment 6 in which students did not use the materials. This gives evidence that there were one or more factors operating within the different classrooms which influenced class mean gain scores. This opens up an avenue for another study of these influences.

Tables 20 and 21 show that the influences of the class mean gain scores are generally nonsignificant for questions unique to the different reading levels and on questions which fall into one of the six special categories. The most consistent of the latter group is the category, factual-unique, which shows significance for the influence of the class in Treatments 2, 3 and 5. However, none of the results are consistent enough that positive conclusions can be reached.

Data Derived to Aid in the Interpretation of Results

Data on Table 21 were derived through records kept by the teachers (as instructed) while their classes used the materials, or
Table 19. Significances of Class Mean Scores Within Each Treatment Group.

<table>
<thead>
<tr>
<th>Source</th>
<th>Treatment-1 RDX (C)</th>
<th>Treatment-2 RDX (A, C, E)</th>
<th>Treatment-3 RDX (A-E)</th>
<th>Treatment-4 RX (No Discussion)</th>
<th>Treatment-5 RD (No Experiment)</th>
<th>Treatment-6 (No Material)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RL</td>
<td>Class</td>
<td>RL</td>
<td>Class</td>
<td>RL</td>
<td>Class</td>
</tr>
<tr>
<td>Total criterion</td>
<td>.05</td>
<td>.01</td>
<td>.05</td>
<td>.01</td>
<td>.01</td>
<td>.05</td>
</tr>
<tr>
<td>test</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>.01</td>
<td>ns</td>
<td>.01</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>
Table 20. Significances of Class Mean Scores within Each Treatment Group on Questions Unique to Reading Levels and to Questions in Six Categories.

<table>
<thead>
<tr>
<th>Source</th>
<th>Treatment-1 RDX (C)</th>
<th>Treatment-2 RDX (A, C, E)</th>
<th>Treatment-3 RDX (A-E)</th>
<th>Treatment-4 RX (No Discussion)</th>
<th>Treatment-5 RD (No Experiment)</th>
<th>Treatment-6 (No Material)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions unique to reading levels:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A(5)**</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>.05</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>B(5)**</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>C(4)**</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>.05</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>D(6)**</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>.01</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>E(5)**</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Questions which are:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factual-unique (11)**</td>
<td>.01</td>
<td>ns</td>
<td>.05</td>
<td>ns</td>
<td>.05</td>
<td>ns</td>
</tr>
<tr>
<td>Conceptual-unique (14)**</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>.01</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Factual-common (12)**</td>
<td>.05</td>
<td>.01</td>
<td>ns</td>
<td>.01</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Conceptual-common (3)**</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>.01</td>
<td>ns</td>
</tr>
<tr>
<td>Conceptual-none (18)**</td>
<td>ns</td>
<td>.01</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Discussion (8)**</td>
<td>ns</td>
<td>.05</td>
<td>.05</td>
<td>.01</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

* Number of questions as reported on Table 8.
from previously reported data supplied here for the sake of comparisons. Data on this table should be surveyed at the same time as that on Table 22.

Table 21. Some Mean Values by Treatments.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean gain score</th>
<th>Mean minutes using materials</th>
<th>Mean years teachers' experience</th>
<th>Mean quarter hours teachers' preparation</th>
<th>Mean student intelligence quotient</th>
<th>Mean student reading level grade-equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.67</td>
<td>667</td>
<td>20.8</td>
<td>27.4</td>
<td>107</td>
<td>5.2</td>
</tr>
<tr>
<td>2</td>
<td>7.93</td>
<td>818</td>
<td>7.4</td>
<td>15.4</td>
<td>106</td>
<td>5.1</td>
</tr>
<tr>
<td>3</td>
<td>8.14</td>
<td>729</td>
<td>10.4</td>
<td>15.4</td>
<td>105</td>
<td>5.1</td>
</tr>
<tr>
<td>4</td>
<td>6.87</td>
<td>674</td>
<td>15.0</td>
<td>27.4</td>
<td>105</td>
<td>5.3</td>
</tr>
<tr>
<td>5</td>
<td>3.32</td>
<td>261</td>
<td>1.3</td>
<td>15.0</td>
<td>109</td>
<td>5.5</td>
</tr>
<tr>
<td>6</td>
<td>2.45</td>
<td>none</td>
<td>9.8</td>
<td>18.7</td>
<td>108</td>
<td>5.2</td>
</tr>
</tbody>
</table>

Data on Table 22 was derived by computing some correlations between factors. It shows a .69 correlation (a reliable correlation should be above .80) between the mean gain score of the treatments and the mean number of minutes spent using the materials. Table 21 shows that students in Treatment 5 spent only 261 minutes with the materials. The fewer number of minutes obviously represents the fact that other treatment groups required quite a lot of extra time to complete the experiments. Students in Treatment 5 did not do any of the experiments.
None of the other correlations on Table 22 show high enough values to be considered positive. However, an interesting one is the -.29 correlation between the treatment gain scores and the mean intelligence quotient. This adds evidence to the opinion which some people hold that the intelligence quotient score is not always a reliable indication of achievement ability.

Table 22. Coefficients of Correlation

<table>
<thead>
<tr>
<th>Source</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean gain score vs. minutes spent on the materials</td>
<td>.69</td>
</tr>
<tr>
<td>Mean gain score vs. years teacher's experience</td>
<td>.45</td>
</tr>
<tr>
<td>Mean gain score vs. quarter hours teacher's preparation in science</td>
<td>.14</td>
</tr>
<tr>
<td>Mean gain score vs. mean intelligence quotient</td>
<td>-.29</td>
</tr>
<tr>
<td>Mean gain score vs. reading level grade-equivalent</td>
<td>.20</td>
</tr>
</tbody>
</table>

Data Concerning the Teacher Questionnaire

A valuable assessment of the materials may be determined from the results on a questionnaire (Appendix IV) which was given to the teachers after they had completed use of the materials (Table 23). They were asked to rate each statement with AA (Agree), AD (Agree somewhat), N (Neutral), DA (Disagree somewhat) or DD (Disagree). These were assigned the following values respectively: 5, 4, 3, 2, and 1.
All statements were positive in character and were therefore ranked according to the highest score.

Table 23. Rank Order from Highest to Lowest Agreement for the Questions on The Teacher Questionnaire

<table>
<thead>
<tr>
<th>Rank</th>
<th>Question number</th>
<th>Rank</th>
<th>Question number</th>
<th>Rank</th>
<th>Question number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>10</td>
<td>25</td>
<td>19</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>11</td>
<td>18</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>12</td>
<td>20</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>14</td>
<td>7</td>
<td>23</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>15</td>
<td>9</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>16</td>
<td>6</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>17</td>
<td>17</td>
<td>26</td>
<td>22</td>
</tr>
<tr>
<td>9</td>
<td>26</td>
<td>18</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Appendix IV

The statement with the highest rank of agreement was No. 15 (Appendix IV), "had a student record book which helped reinforce learning by causing the student to record." Teachers apparently liked the record book. However, an evaluation of statement No. 16, "had a student record book which involved too much writing on the part of students," ranked 14 out of 26. This indicates that teachers felt there was too much writing involved in the record book.

The statement ranking second highest was No. 1, "are attractive to students." Teachers were high in their praise for the
attractiveness of the materials to their students and verbally reported that most students really enjoyed using them.

Third high was No. 21, "are a step in the right direction in the preparation of curriculum materials." Teachers were elated at the prospect of having materials which accommodated the reading spread in their classes. This was the feature which convinced most of the teachers that they would like to try the materials.

Fourth and fifth ranked were No. 11, "had a good balance between reading and experiment involvement," and No. 19, "should have had a separate kit of materials for experiments for each teacher." The first of these statements is good endorsement of the materials. The latter indicates that too many times essential materials, such as balloons, wire, etc. had been used up when the materials were passed to the sharing room and class procedure was delayed.

The lowest ranked statement was No. 22, "are too complex for teachers to use." In other words, they felt the materials were not too complex to be used adequately.

The next lowest ranked was No. 5, "had too much reading to it." They disagreed with this, but many of them remarked that some of the research booklets were too difficult as shown by a rank of 18 for statement No. 10, "were hard for students to understand."

Other statements with which they disagreed were No. 23, "call for too much student participation"; No. 14, "involved too much class
discussion"; No. 22, "had too many experiments"; and No. 24, "are less satisfactory than other elementary science materials." In other words they like the greater amount of student participation called for in the materials and they felt the materials were generally more satisfactory than other science materials they had been using.

All of this amounts to the fact that with some good teacher in-service training, the teachers could probably learn to use the materials to greater advantage.
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary of Results

The statements enumerated below are based upon evidence produced in this study. The first three treat the null hypotheses tested and the others are relevant to other findings.

1. There were no significant differences in mean gain scores among treatments involving classes which used science materials written at one, three or five different reading levels.

2. There were no significant differences in achievement gain scores between treatment group four in which students did not discuss their science materials, and the first three treatment groups involving different reading levels.

3. There was a significant adverse influence on achievement gain scores between Treatment 5 and any of the first four treatment groups.

4. There were no consistent significant differences in achievement gain scores among treatments for questions unique to any of the reading levels nor for questions in any of the six special categories.

5. The reading level groups ordered themselves from lowest (A) to highest (E) consistently.
6. The influence of teachers on class mean achievement gain scores showed consistent significance.

7. The teacher questionnaire indicated that the teachers were generally favorable in their judgment of the materials.

Conclusions

In considering the conclusions of this study the reader is reminded that the evaluation reflects only a measurement of knowledge achievement gain for the materials and lacks any information concerning other identified objectives, such as arousing curiosity, stimulation of creativeness or the gain in skills of inquiry, all three of which are strong objectives of the producers of these science instructional materials. This was stated in the beginning as one of the limitations of the study.

The information produced by this study depends directly upon the assumptions expressed earlier: (1) materials are written at the grade levels described, (2) elementary teachers with limited science backgrounds can handle the unique materials in the manner recommended by the authors, (3) classroom teachers followed prescribed instructions, and (4) the criterion test validly and reliably measured achievement gain.
Reading Level

The study shows clearly that achievement gain scores as measured by this criterion test are not influenced significantly by the fact that the materials are written at grade-equivalent reading levels (Table 12). There was only one minor exception. Students in Treatment 1 (students read only at C level) did significantly better than those in Treatment 2 (students read at A, C and D reading levels) on questions unique to reading level C (Table 13, line 3, column 5). However, this does not change the validity of the first statement in this paragraph. On the basis of these results the first null hypothesis (no significant differences whether materials are written at one, three or five different reading levels) is accepted.

Nevertheless, it would be going beyond the evidence to assume that the materials written for a specific reading level did not influence the students in any manner whatsoever. It will depend upon the development of other kinds of measuring instruments before one can determine, for instance, whether students who are given the opportunity to read at their own level are not motivated to read more than those who are forced to read either above or below their reading level. As cited in the second section, there is much to be said for just 'enjoying' what one reads when it is at his reading ability level. For this reason, the teachers' questionnaire which reveals that the students enjoyed
the materials takes on greater significance.

Perhaps the study was searching out the wrong objective when it evaluated only the achievement gain based on a short unit. Admittedly, achievement is one of the most important goals in teaching but perhaps it is related to a sustained presentation of materials at reading ability level, rather than one short unit as was the case in this study.

**Class Discussion**

The influences of Treatment 4 (in which the variable was non-discussion of the materials) on achievement gain scores were not significant in this study except for minor exceptions. Achievement among students in Treatment 1 (all students read at C level only) was significantly better than in Treatment 4 for questions which were conceptual-common and also for the discussion questions (Table 16, lines 9 and 11, column 1). Another exception was the comparison of Treatment 3 (students read at five different levels) and Treatment 4 on factual-common questions (Table 16, line 8, column 3). In these cases the lack of discussion apparently caused the students to do significantly poorer in Treatment 4. Nevertheless, the general results of this study support the second null hypothesis (no significant differences whether or not the materials are discussed), and it is accepted.

A check of the time spent (Table 21) shows that students in
Treatments 1, 2 and 3, (in which classes discussed the materials) averaged seven percent more time with the materials than students in Treatment 4 (in which the materials were not discussed). Greater familiarity on the part of participating teachers, with the unique discussion techniques afforded by these materials seems indicated. Although we must be guided by the statistical data, another look at Figure 2 allows us to speculate that the discussion factor was close to significance and that a little more familiarity with the materials might have allowed the teachers to be more effective in this area.

The Biological Sciences Curriculum Study Group have put great emphasis on the value of discussion and as stated by Schwab (16, p. 430-431) the intangibles of "understanding" and "wanting" developed through class discussion are important and tied to effective achievement.

Experiments

The results of the study show that students in Treatments 1, 2 and 3, which involved specific reading levels, made significantly higher achievement gain scores than those in Treatment 5 (in which the variable was not performing the experiments). Likewise, in Treatment 4, (in which the variable was nondiscussion of the materials), students performed significantly better than those in Treatment 5 (Table 17). On the basis of these comparisons the third null
hypothesis (no significant differences whether students perform experiments or not) is rejected.

In relation to Treatment 5 it will be recalled that students in this treatment spent the least amount of time with the materials and furthermore that there was a .69 correlation between gain scores for all treatments and the amount of time spent with the materials. One is tempted to say that the students who spend the least amount of time with the materials are certain to do poorest on the criterion test. However, the reader should be reminded that only 2 of the 60 items on the criterion test were taken from the experiments (Table 8). This means that students in Treatment 5 had ample opportunity to learn the same materials that students had in other treatment groups. Furthermore, it lends direct implication that performance of the experiments by students in the first four treatment groups reinforced their learning. This is an important finding and strongly suggests that those who prepare science curriculum materials should make experiments a central part rather than only an adjunct to their publications. At the present time commercial producers of textbooks for elementary school science do not make experiments the central part of their materials. However, organizations such as the American Association for the Advancement of Science (AAAS) are trying to do this very thing under Federal grants. This is also true of other curriculum projects such as The Illinois Program which has produced materials
for elementary schools in astronomy.

Questions Unique to Reading Levels or in Special Categories

As shown on Table 8 questions unique to the different reading levels or which fell into special categories as described following Table 8, were analyzed to see if they had special significance. It was presupposed that perhaps students reading at the lowest level would do best on questions unique to materials written at that level. The same is true for questions unique to the other reading levels. However, no such trend is apparent. A very positive trend was for the learners to order themselves, based on gain scores, by having the lowest level readers score lowest and the highest level readers score highest. It is concluded that it makes no significant difference on gain scores whether a question is unique to a reading level or not.

There was also a question about whether students would make higher achievement scores on questions which were unique or common to all reading levels, conceptual or factual, or on those taken from discussion. As shown in preceding section, there was no consistent pattern of significance in any of the categories and it is concluded that the type of question makes no significant difference on achievement gain scores.
Teacher Questionnaire Assessment

As explained in the third section, this questionnaire (Appendix IV) was given to teachers after they had completed using the materials in their classes. All of the questions were positive in nature and were scored and ranked on the basis of the highest agreement with the statement. It is interesting to note that there was general acceptance of the materials by the teachers. This complements the feeling the investigator had about the enthusiasm the teachers had for wanting to try the materials. They were very interested in finding materials designed to meet the reading ability needs of their students and which had a different approach to learning.

Recommendations

The strongest recommendation for a further investigation related to the one completed here would be to measure the effects of these same materials after having organized them so the experiments become the focal point. Based on the results of this investigation, it would seem logical to present the experiments to the students first. Two things would likely happen; (1) lacking "previous information" students would have greater desire to write their own explanations for their results and, (2) students would have greater incentive to read "additional information" related to what they found out in their
experiments. This additional information could be supplied in the research booklets and by the author's recommended reading lists.

Another recommendation would be to devise criterion instruments which measure the effects of materials presented at reading ability level on the other objectives listed by the authors; (1) to arouse curiosity in science, (2) to stimulate creativeness, and (3) to develop skills of inquiry. The materials need to have these objectives measured. This would entail the development of a different kind of evaluation instrument, one not dependent upon reading ability.

Further investigation of class discussion as it is related to learning needs to be carried out. Again, its contribution to the objectives listed above should be determined. Verbalization on the part of the student is another facet of self-involvement which is so important to "understanding."

In view of the fact that the influence of the teacher on class achievement scores shows itself as important, an evaluation needs to be done which measures the significant influence on student achievement scores by two groups of teachers; (1) those who have had special workshops with these materials and in the "multi-level" philosophy, and (2) those who use the materials without workshops. The differential influence of teachers on their classes was very apparent in this study; however, it was probably based upon a particular teacher's interest in science. The measurement should be on teachers who
express a high science interest.

Summary

This particular evaluation needed to be completed because this is the first commercial producer to actually try to build materials to meet individual differences in the science area of the elementary school. Their philosophy is sound and deserves not only this evaluation but those mentioned above as well. They have already effected changes in their presentation based on this evidence. These same kinds of evidences will also give other commercial producers a foundation on which to build effective materials.
BIBLIOGRAPHY


APPENDICES
APPENDIX I

1. Which of the following would not cause a weather balloon filled with helium to fall back to the earth's surface?

   A. Heating the helium
   B. Cooling the helium
   C. Compressing the helium
   D. Releasing the helium

2. What would happen to a half-filled balloon if it were left on a warm radiator?

   A. It would get larger
   B. It would gradually get smaller
   C. It would melt
   D. Nothing would happen

3. Equal amounts of air were put into two balloons. One balloon was placed in a refrigerator; the other balloon was held over a pan of boiling water. Later the two balloons were compared.

   Which of the following describes the balloon that was placed in the refrigerator?

   A. Heavier than the other balloon
   B. Lighter than the other balloon
   C. Smaller than the other balloon
   D. Larger than the other balloon

4. The earth's atmosphere is made of

   A. ice crystals
   B. water vapor
   C. pure oxygen
   D. a mixture of gases
5. The term "dense air" means the molecules in the air are

A. mixed with water vapor  
B. warm  
C. far apart  
D. close together

6. Which of the following is **not** a good example of how air can do work?

A. Sailboat  
B. Windmill  
C. Air conditioner  
D. Parachute

7. Which of the following statements is false?

The air ocean is made up of gases that are

A. colorless  
B. weightless  
C. odorless  
D. tasteless

8. When molecules in the air **are** heated they will

A. increase in size  
B. move faster  
C. Decrease in size  
D. move slower

9. If you have a shovel made of iron, you probably oil it before packing it away. Oiling the iron shovel will help prevent

A. rotting  
B. rusting  
C. dehydration  
D. freezing
10. At which of the following altitudes would you expect your heart to beat the fastest?

A. 10 feet  
B. 100 feet  
C. 1,000 feet  
D. 10,000 feet

11. In a sample of air from your classroom, you would **not** expect to find

A. water vapor  
B. chlorine gas  
C. dust particles  
D. helium gas

12. If cold air suddenly entered a heated room, it would

A. decrease the number of molecules in the room  
B. decrease the air pressure  
C. force the warm air up  
D. immediately rise to the ceiling

13. When molecules in air push against a surface the result is called

A. weight  
B. density  
C. pressure  
D. gravity

14. When a glass of ice water is placed in a warm room for a few minutes, drops of water appear on the outside of the glass.

This water comes from the

A. air  
B. ice  
C. glass  
D. water in the glass
15. Suppose you were holding your finger over a drinking straw half full of water so that the water remained in the straw. If you were quickly flown to the top of a high mountain in a helicopter

A. all the water would stay in the straw
B. the water would be pushed up higher into the straw
C. some of the water would drip out of the straw
D. no water would remain in the straw

16. A stream of air has less sideways pressure when it is

A. moving slowly
B. moving fast
C. cooled
D. heated

17. When you drink a coke through a straw

A. air pressure pushes the coke up the straw
B. the coke rises because it is expanding
C. the weight of the coke pushes it up the straw
D. the temperature of your mouth warms the coke and makes it rise

18. If you open a bottle of cold air in a warm room and want the cold air to stay in the bottle, you should keep the bottle

A. right side up
B. upside down
C. on its side
D. in your hand

19. The gas which makes up about 79% of our atmosphere is

A. helium
B. oxygen
C. carbon dioxide
D. nitrogen
20. You could probably wash your hands in boiling water on the top of a high mountain. This is because the boiling-point of water depends on the

A. temperature of the air
B. amount of water
C. density of the water
D. amount of air pressure

21. Why does air push (exert pressure) in all directions, but a solid pushes (exerts pressure) only downward?

Molecules in air

A. are heavier than those of a solid
B. move more freely in all directions
C. are warmer than those of a solid
D. weigh less than those of a solid

22. If an empty water glass were turned upside down and pushed down into a bucket of water, the glass would

A. have air in it
B. have nothing in it
C. fill with water
D. break because of the pressure

23. One way to make the molecules in air move faster is to

A. heat the air
B. cool the air
C. dry out the air
D. humidify the air

24. Carbon dioxide is useful for fighting small fires because it

A. takes up more space than oxygen
B. is colder than oxygen
C. is more dense than oxygen
D. is lighter than oxygen
25. While playing baseball in a city at a high altitude, a player should be able to hit the ball farther because at the high altitude

A. balls are more "alive"
B. pitchers can't throw the ball as hard
C. the ball parks are smaller
D. there are fewer molecules in the air

26. A fluid will always

A. be wet
B. be colder than things around it
C. have the same shape as its container
D. be round like a drop of water

27. Which of the following was not considered an element by ancient peoples?

A. Sun
B. Air
C. Earth
D. Fire

28. Inside a closed bottle, air presses

A. only against the lid
B. only against the sides
C. only against the bottom
D. in all directions

29. Oxidation is the process of oxygen

A. combining with something else
B. being produced by green plants
C. being weighed
D. being produced by animals
30. Which of the following is a "layer" of the atmosphere?

   A. Troposphere  
   B. Stratosphere  
   C. Ionosphere  
   D. All of the above

31. At room temperature air is

   A. liquid  
   B. solid  
   C. fluid  
   D. molecule

32. People on earth live in an ocean of air. Where in this air ocean do they live?

   A. Near the top  
   B. Near the middle  
   C. Near the bottom  
   D. Near the equator

33. Suppose you put a drinking straw into a glass of water. Now if you put your finger over the top of that straw and lift it from the water, there will still be water inside the straw because

   A. air is lighter than water  
   B. the straw is very narrow  
   C. air is pushing up on the water at the bottom of the straw  
   D. air above the water in the straw holds it in

34. Air is made up mostly of

   A. different gases  
   B. dust particles  
   C. water particles  
   D. oxygen
35. If you weren't wearing a space suit and you fell out of a rocket flying in space, you would be killed by

A. too much pressure on the outside of your body
B. intense cold
C. heat of the sun
D. lack of pressure on the outside of your body

36. If you were riding a bicycle on the highway and a long truck passed close to you at a fast speed, air pressure might push you

A. up into the air
B. into the side of the truck
C. down onto the pavement
D. into the ditch

37. If you removed most of the air from a container, the molecules left in the container would

A. settle to the bottom
B. move slower
C. evaporate
D. spread out

38. It is hard to pull a suction cup off a wet, smooth surface because

A. the vacuum inside holds it tightly against the surface
B. air outside the cup pushes against it and holds the cup against the surface
C. high air pressure inside holds it against the surface
D. the moisture makes it stick

39. You are able to "see your breath" when it is cold outside because

A. it turns to snow
B. it warms the air and makes it visible
C. the oxygen condenses
D. water vapor from your mouth cools and forms droplets
40. It is easier for jets to fly at higher altitudes than at lower altitudes because

A. they run into fewer molecules in the air  
B. there is more oxygen for fuel at high altitudes  
C. there are no winds at high altitudes  
D. they are almost weightless at high altitudes

41. Anything which flows is called a

A. gas  
B. liquid  
C. fluid  
D. solid

42. Blowing more molecules of air into a balloon will

A. increase the pressure in the balloon  
B. decrease the weight  
C. lower the temperature  
D. decrease the density

43. If our air had half as much oxygen as it now does, forest fires would be

A. less likely to start  
B. harder to control  
C. hotter than the ones we have now  
D. more likely to occur on high mountains

44. The gas which causes objects to rust is

A. nitrogen  
B. carbon dioxide  
C. helium  
D. oxygen
45. You would become short of breath more easily running on a mountain top than running at sea level because

A. it is colder on the mountain top
B. you would be breathing in too much nitrogen
C. there is more carbon dioxide on the mountain top
D. there is less oxygen at higher altitudes

46. Which of the following can be most easily compressed?

A. Soil
B. Water
C. Nitrogen gas
D. Tar

47. The "hissing" sound that is made when a can of vacuum-packed coffee is opened is caused by

A. air escaping
B. air rushing in
C. pressure decreasing
D. air in the can contracting

48. A helium filled balloon will rise into the air because heavier air pushes it up. The balloon will stop rising when it

A. reaches the top of the atmosphere
B. gets as large as it can get
C. becomes too cold to rise higher
D. becomes the same weight as air around it

49. If an inflated football were left in a room where the air was gradually being removed, the

A. weight of the football would increase
B. pressure inside the football would increase
C. weight and pressure inside the football would not change
D. size of the football would increase
50. Any push or pull is called
   A. density
   B. force
   C. weight
   D. mass

51. The earth pulls all things toward it. Because of this we say all things on earth have
   A. molecules
   B. atoms
   C. weight
   D. magnetism

52. At sea level, the distance between gas molecules in the air is
   A. about an inch
   B. more than a mile
   C. nearly eighty feet
   D. less than the width of an eyelash

53. If you put an airtight cap on a can in which you had just boiled some water and set it aside in a cool place, the can would
   A. explode
   B. be crushed
   C. cool slower
   D. cool faster

54. The air around the earth gets its heat from
   A. the earth's surface
   B. the sun
   C. volcanoes
   D. factory smoke and car exhausts
55. Suppose you punched a small nail hole near the bottom, the middle and near the top along the side of a tall tin can. If you covered each hole with one of your fingers, filled the can with water, then uncovered all 3 holes at once, the water would go out the bottom hole

A. with most force
B. with least force
C. with the same force as through the other holes
D. none of these

56. If you poured water over half a glassful of dry dirt you would see

A. bubbles of carbon dioxide
B. bubbles of water
C. bubbles of dirt
D. bubbles of air

57. Why is the gas argon used in many light bulbs?

A. because it won't combine with most things
B. because it is very light
C. because it is cheaper than any other gas
D. because it is colorless

58. A gas which helps protect living things from too much ultraviolet light from the sun is

A. helium
B. oxygen
C. ozone
D. nitrogen

59. If you hold the tops of two sheets of notebook paper about 2 inches apart, then blow air between them, they will

A. move farther apart
B. move closer to each other
C. try to fly off
D. tear apart
60. In salt water, a block of wood would float higher (more of the wood out of the water) than in fresh water. This is because salt water is

A. lighter than fresh water
B. warmer than fresh water
C. heavier than fresh water
D. colder than fresh water
APPENDIX II

EXPERIMENTAL TREATMENT--RDX (C). Experimental Group

1. Your class will be one of the experimental groups. Please be, and encourage your students to be, as creative as possible within the restrictions of this experiment and your own good judgment about when the class is learning and when it is marking time.

A. Perhaps you will note these materials work best for a teacher who is more of a wise counselor and organizer, than for one who is used to "telling." Be encouraging to students in the classroom experiments--remember they do not all work perfectly, but when they do not--there is a reason for it. Encourage students to learn from mistakes. Remember you are not expected to know all the answers to the problems--you are only expected to be a creative teacher who (because he or she is a teacher) knows how to help learners learn! Don't be afraid to say "I don't know", but always follow up with, "Let's see if there isn't a way to find out!"

2. Nothing is masked out of your lab, but the research booklets for reading levels A-B and D-E have been removed. You have received extra C research booklets. First survey the materials in the lab. Question yourself about why this or that was included and why it is organized like it is. Read four parts thoroughly: (1) Teacher's Handbook, (2) Teacher's Instructional Aid Booklet, (3) Student Record Book, and (4) Key-Model Booklet. It will not be necessary to read the content of each Student Research Booklet, however, you will probably want to be familiar with each of the experiments in order to help students plan materials, etc. Review the materials in the lab after reading the above--they will now make more sense to you. Recite to yourself the names of the parts--you must be familiar with the names of the parts in order, later, to help your students become familiar with them.

3. Note the time allotment schedule on pages 21-22 of the TEACHER'S HANDBOOK. These times are estimated--you are not bound by them. You should follow this outline of presentation. Will you please write down, opposite the estimated printed times for each part, the approximate time it required your class to complete the parts. I will need it for analysis.
A. The only variation we ask is that you do not encourage the students to read other research booklets outside their reading level--this is for experimental purposes. Some of the criterion questions are tagged to certain reading levels and we want to determine if they are transmitted by processes other than reading.

4. Tell students they will be pretested and post tested--but that the results of these tests will not be used to determine grade. They will be graded only on their participation and co-operation in the experiment.
EXPERIMENTAL TREATMENT--RDX (A-C-E). Experimental Group

1. Your class will be one of the experimental groups. Please be, and encourage your students to be, as creative as possible within the restrictions of this experiment and your own good judgment about when the class is learning and when it is marking time.

A. Perhaps you will note these materials work best for a teacher who is more of a wise counselor and organizer, than for one who is used to "telling." Be encouraging to students in the classroom experiments--remember they do not all work perfectly, but when they do not--there is a reason for it. Encourage students to learn from mistakes. Remember you are not expected to know all the answers to the problems--you are only expected to be a creative teacher who (because he or she is a teacher) knows how to help learners learn! Don't be afraid to say "I don't know", but always follow up with, "Let's see if there isn't a way to find out!"

2. Nothing is masked out of your lab, but the research booklets for reading levels B and D have been removed. You have received extra A and E research booklets. First survey the materials in the lab. Question yourself about why this or that was included and why it is organized like it is. Read four parts thoroughly: (1) Teacher's Handbook, (2) Teacher's Instructional Aid Booklet, (3) Student Record Book, and (4) Key-Model Booklet. It will not be necessary to read the content of each Student Research Booklet, however, you will probably want to be familiar with each of the experiments in order to help students plan materials, etc. Review the materials in the Lab after reading the above--they will now make more sense to you. Recite to yourself the names of the parts--you must be familiar with the names of the parts in order, later, to help your students become familiar with them.

3. Note the time allotment schedule on pages 21-22 of the TEACHER'S HANDBOOK. These times are estimated--you are not bound by them. You should follow this outline of presentation. Will you please write down, opposite the estimated printed times for each part, the approximate time it required your class to complete the parts. I will need it for analysis.

A. The only variation we ask is that you do not encourage the students to read other research booklets outside their reading level--this is for experimental purposes. Some of the
criterion questions are tagged to certain reading levels and we want to determine if they are transmitted by processes other than reading.

4. Tell students they will be pretested and post tested--but that the results of these tests will not be used to determine grade. They will be graded only on their participation and co-operation in the experiment.
EXPERIMENTAL TREATMENT--RDX (A-E)  Control Group

1. Your class will be one of the control groups. Please be, and encourage your students to be, as creative as possible within the restrictions of this experiment and your own good judgment about when the class is learning and when it is marking time.

A. Perhaps you will note these materials work best for a teacher who is more of a wise counselor and organizer, than for one who is used to "telling." Be encouraging to students in the classroom experiments--remember they do not all work perfectly, but when they do not--there is a reason for it. Encourage students to learn from mistakes. Remember you are not expected to know all the answers to the problems--you are only expected to be a creative teacher who (because he or she is a teacher) knows how to help learners learn! Don't be afraid to say "I don't know", but always follow up with, "Let's see if there isn't a way to find out!"

B. You will be furnished with two lists of materials for experimentation. The first list is for common materials which students can bring from home, or which it is reasonable that schools supply. It would be a good idea to get the list of these things collected and stored in the room so you will not be delayed because a student forgets to bring a promised item. The second is one of the materials which SRA is furnishing to your classroom. Please become familiar with these and indicate on the sheet the information asked for in regard to this supply.

2. Nothing is masked out of your lab. First survey the materials in the lab. Question yourself about why this or that was included and why it is organized like it is. Read four parts thoroughly: (1) Teacher's Handbook, (2) Teacher's Instructional Aid Booklet, (3) Student Record Book, and (4) Key-Model Booklet. It will not be necessary to read the content of each Student Research Booklet, however, you will probably want to be familiar with each of the experiments in order to help students plan materials, etc. Review the materials in the Lab after reading the above--they will now make more sense to you. Recite to yourself the names of the parts--you must be familiar with the names of the parts in order, later, to help your students become familiar with them.
3. Note the time allotment schedule on pages 21-22 of the TEACHER'S HANDBOOK. These times are estimated—you are not bound by them. You should follow this outline of presentation. Will you please write down, opposite the estimated printed times for each part, the approximate time it required your class to complete the parts. I will need it for analysis.

A. The only variation we ask is that you do not encourage the students to read other research booklets outside their reading level—this is for experimental purposes. Some of the criterion questions are tagged to certain reading levels and we want to determine if they are transmitted by processes other than reading.

4. Tell students they will be pretested and post tested—but that the results on these tests will not be used to determine grade. They will be graded only on their participation and co-operation in the experiment.
EXPERIMENTAL TREATMENT--RX(A-E) Experimental Group

1. Your class will be one of the experimental groups. Please be, and encourage your students to be, as creative as possible within the restrictions of this experiment and your own good judgment about when the class is learning and when it is marking time.

A. Perhaps you will note these materials work best for a teacher who is more of a wise counselor and organizer, than for one who is used to "telling." Be encouraging to students in the classroom experiments--remember they do not all work perfectly, but when they do not--there is a reason for it. Encourage students to learn from mistakes. Remember you are not expected to know all the answers to the problems which will arise. You are expected to be a creative teacher who (because he or she is a teacher) knows how to help learners learn! Don't be afraid to say "I don't know", but always follow up with, "Let's see if there isn't a way to find out!"

B. You will be furnished with two lists of materials for experimentation. The first list is for common materials which students can bring from home, or which it is reasonable that schools supply. It would be a good idea to get this list of materials collected and stored in the room so you will not be delayed because a student forgets to bring a promised item.

The second list is one of the materials which SRA is furnishing to your classroom. Please become familiar with these and indicate on the sheet the information asked for in regard to this supply.

2. Everything in your lab and other materials regarding any type of class discussion is masked. Your class will do everything except discuss the reading materials or the experiments. You will do the experiments, but will not discuss, nor demonstrate them.

First survey the materials in the lab. Question yourself about why this or that was included and why it is organized like it is. Read (except the masked parts) four parts thoroughly: (1) Teacher's Handbook, (2) Teacher's Instructional Aid Booklet, (3) Student Record Book, and (4) Key-Model Booklet. It will not be necessary to read the content of each student Research Booklet, however, you will probably want to be familiar with each of the experiments in order to help students plan materials, etc.
Review the materials in the Lab after reading the above—they will now make more sense to you. Recite to yourself the names of the parts in order, later, to help your students become familiar with them.

3. Note the time allotment schedule on pages 21-22 of the TEACHER'S HANDBOOK. These times are estimated—you are not bound by them. You should follow this outline of presentation (except where masked). Will you please write down, opposite the estimated printed times for each part, the approximate time it required your class to complete the parts. I will need it for analysis.

A. Ask students not to read other research booklets outside their reading level—this is for experimental purposes. Some of the criterion questions are tagged to certain reading levels and we want to test for them.

4. Tell the students they will be pretested and post tested—but that the results of these tests will not be used to determine grade. They will be graded only on their participation and co-operation in the experiment.
EXPERIMENTAL TREATMENT--RD (A-E) Experimental Group

1. Your class will be one of the experimental groups. Please be, and encourage your students to be, as creative as possible within the restrictions of this experiment and your own good judgment about when the class is learning and when it is marking time.

A. Perhaps you will note these materials work best for a teacher who is more of a wise counselor and organizer, than for one who is used to "telling." Remember, you are not expected to know all the answers to the problems—you are only expected to be a creative teacher who (because he or she is a teacher) knows how to help learners learn! Don't be afraid to say "I don't know", but always follow up with, "Let's see if there isn't a way to find out!"

2. Your class will do none of the experiments, nor the discussion about the experiments. All this part is masked in both your, and the student material. You will do all the other parts, including the discussion of the research booklets.

A. Get acquainted. First survey the materials in the lab. Question yourself about why this or that was included and why it is organized like it is. Read (except the masked parts) three parts thoroughly: (1) Teacher's Handbook, (2) Teacher's Instructional Aid Booklet and, (3) Student Record Book. It will not be necessary to read the content of each Student Research Booklet, however, you will probably want to be familiar with the responses to the discussion questions in the TEACHER'S INSTRUCTIONAL AID BOOKLET. Review the materials in the Lab after reading the above—they will now make more sense to you. Recite to yourself the names of the parts in order, later, to help your students become familiar with them.

3. Note the time allotment schedule on pages 21-22 of the TEACHER'S HANDBOOK. These times are estimated—you are not bound by them. You should follow this outline of presentation (except where masked). Will you please write down, opposite the estimated printed times for each part, the approximate time it required your class to complete the parts. I will need it for analysis.
A. Ask students not to read other research booklets outside their reading level--this is for experimental purposes. Some of the criterion questions are tagged to certain reading levels and we want to test for them.

B. Tell the students they will be pretested and post tested but that the results of these tests will not be used to determine grade. They will be graded only on their participation and co-operation in the experiment.
EXPERIMENTAL TREATMENT--CRITERION TEST CONTROL GROUP

1. Your class will be pretested and post tested when the other groups are. However, you will not use any of the SRA science materials. Your part is that of control.

2. You should carry on with your normal science class periods, but please do not use any units which pertain to weather or the atmosphere. Try to avoid reference and study about them.

3. Your school will receive a complimentary set of the materials following the experiment, and at that time you may use the materials as you see fit.
To classroom teachers:

The materials listed below should be furnished by the students or the school. Teachers should be alert to the times when such items as electric fans are needed so student volunteers can supply them for experiments. Many of the other things can be stock piled in the room.

Alcohol lamp
Blanket (for fire protection)
Bottles (milk and pop)
Cardboard (assorted sizes and shapes)
Cloth pieces (assorted colors and sizes)
Electric fan
Extension cord
Glass jars (pints, quarts, larger, olive, all with lids)
Fire extinguisher
Hammer
Hot pan holders
Marking pencil
Medicine bottles (2-3)
Pencils with erasers
Pie pans (loaned from home)
Pins
Rulers (1/2 doz.)
Salt
Sand (bucket full)
Scissors
Slide projector (for light source)
Spools (thread)
Spoons
Stopwatch
Tape (cellophane and adhesive)
Thermometer
World globe
APPENDIX III

Dear Teachers,

Thank you for your patience! I think we are ready at last. Our first step is to administer the criterion pretest. This is necessary to establish our base. The next step is for each of you to use the lab precisely according to instructions—each one of you has a vital part to play if we are going to be able to analyze the philosophy on which these materials are built. Remember, the four areas are: (1) Multi-level reading materials, (2) Directed discussion which encourages inter-action among student reading levels, (3) Group and individual investigations and (4) Self-evaluation. (We have found the latter to be so intrinsically woven into the other parts that it is impossible to pull it out with this criterion test. Perhaps we will do a special study on it at another time.)

We were quite surprised to see the great variability in reading level among the rooms—we had hoped they would be closer together. Because of this fact we have had to match them as closely as possible and the purely random assignment of experimental parts hasn't been possible in all cases. Please remember how important it is for you to carry out your part according to directions. Each of you will get a sheet of instructions for your part. Some parts of the labs have been masked and so have the teacher's handbooks and teacher's instructional
guide book. These are sort of like an "AFTER CHRISTMAS GIFT--DO NOT OPEN UNTIL FEBRUARY." When you have finished with your part of the experiment and your class has taken the criterion post test, then you may take off the wraps and use whatever part you wish. There will be a complimentary lab for each participating school.

Practically all of the directions will have to be by correspondence because it will be physically impossible for me to be in each room every day--I will make one or two of your sessions, if it is at all possible. Please feel free to call me collect at 757-1421, Ext. 345 if you need anything. I will do my best to supply all the essentials, but of course, there will be some things which you as a creative teacher will have to develop, just as you would if you were not helping in this experiment.

All results will be published by TEACHING RESEARCH and I will see that each school receives a copy. We are off to a good start--your cooperation has been the very best--here's hoping we have a good study under way!

Sincerely,

William H. (Bill) Fryback
Research Assistant

WHF:bl
12/16/63
Dear Teachers:

You are certainly a great group of people. In spite of all the delays, I was most impressed by your enthusiasm for this project. That surely means that it will be carried through well!

I have picked up all the pretest papers except for two or three, which I shall get next week. Some of you have had quite an absence problem and have wondered what to do. You will simply have to go on with the materials and work the absentees into it as well as possible when they return to school. Test them before letting them work with the materials. Put the test material and answer sheet into a manila envelope and mail it to me at O. C. E. If you have a student who is absent more than a total of one week, during the time your class is using the materials, please note the fact on your class list opposite his name.

I know some of you will have misgivings about having to teach the part of the experiment which you drew. First, let me assure you that you were assigned to a particular part by random choice, after your class was matched in a group of five. There was no reason why your class was assigned a particular part, except by random choice.

Some of you may feel you are penalizing your students by not discussing the material with them, or having them experiment, or whatever your part happens to be. Please remind yourself that you are helping in a rather short-term experiment. Follow the directions in the letter which was taped to the SRA Laboratory Box. If anyone of you is in doubt about what part you are to do, please call me.

Remember, you will have an opportunity to re-do any part which you missed. Please tell your students that they are playing a part in the experiment. I am sure it will please them, and make them perform better.

Our post test should come about the first week in March. You should all set that week up as a target date. Those of you who are doing all parts should try to move right along. You should probably figure on a science period each day. After you get through the starter activity, and the first "big idea", things will go faster. It takes a little while for the students to become accustomed to the material.
Those who are not doing the discussing, or the experiments can probably make it on a three or four day a week schedule. You will know how fast you can go after just a little experience.

Follow the schedule on page 21 of the Teacher's Handbook and mark down how much time it actually took you to accomplish the step described. Remember how important the results may be on the break-through for multi-level reading materials in many areas of the curriculum.

Thanks for your good cooperation.

Sincerely,

William H. Fryback

WHF:la
Dear Teachers,

It is a real satisfaction to see things going so well with you. It looks as if most of you will make the finish sometime during the first week of March. If some of you run over until the second week it will be all right. Most of you are finding you speed up after getting through the starter activities. Those of you who are reading all levels, along with discussion and experiments, should not try to move too fast, thereby slighting the materials. Please contact me if there are any serious problems.

A few general questions have come up for which I have some answers.

1. Q- What about students who obviously are not assigned the right reading level--it is too high, or too low?
   A- Impress on them that this is an experiment and ask them to go along with us--even if the assignment is wrong.

2. Q- What can we do with the loose pages in the Student Record Book?
   A- Some have tied strings through the holes to hold them in the book.

3. Q- What if some of the students do not have materials for home activities? Such things as clay?
   A- Encourage them to think of substitutes--like wet mud for clay, etc.

4. Q- How fast should we go?
   A- Try to finish around the 1st of March. You will find you can pace yourselves after finishing the starter activity.

5. Q- Where can we get "Plumber's Friends" that are sanitary?
   A- Use the little rubber suction darts in your experimental kits.

6. Q- Why didn't you tell us when we placed a newspaper over it and struck the end sticking over the end of the table, that it would break? (The yardstick--the only one I had)
   A- Mmmmm Aaaaaaa Welllll****!
Thanks again for your good spirit and good works. I will get all the testing materials to you at the proper time. You will probably not all test at exactly the same time.

If you have any holdover tests--please mail them to me in Monmouth.

Are you keeping:

1. A record of the amount of time you are spending?

2. A record of students who miss more than 5 class periods? (Catch them up as much as possible--but don't worry about the impossible)

3. A list of suggestions about how to make the materials more effective--and general constructive criticisms?

Sincerely,

William H. Fryback
Research Assistant
March 19, 1964

Dear Teachers:

Aren't you happy it is over at last. No one knows better than I, how inconvenient it is to be doing someone else's program. You have been wonderfully cooperative—even to the crippling little last minute requests like putting I.Q. scores on the class lists. I also appreciate the suggestions and constructive criticisms which many of you took the time to write.

I was pleased that many of you were able to get by without having any absences for the testing. All but a few of you held out the necessary testing materials to test those who were absent. I will include, with this letter, the necessary materials for those of you who did not.

About 100% of your absentees will be back right after Spring Vacation—please try to catch them and have your school secretaries send the results to me as soon as possible. SRA is working on the scoring of the tests now. They will return the raw scores to me and I will put them through the computer. I will have some tentative data in June, but it will probably be August before it is complete. Each school will receive a copy of the report when it is finished.

Thanks again for all of your help.

Sincerely,

William H. Fryback
Research Assistant

WHF:la
APPENDIX IV

TEACHER QUESTIONNAIRE

Name__________________________ April 3, 1964
School_________________________

Instructions:

Please answer the following according to this code and according to the way your class used the materials.

AA (Agree)
AD (Agree somewhat)
N (Neutral)
DA (Disagree somewhat)
DD (Disagree)

The SRA Learnings in Science Laboratory materials:

____ 1. are attractive to students.
____ 2. are stimulating to students.
____ 3. hold the interest of students.
____ 4. meet the reading level spread of my class.
____ 5. had too much reading to it.
____ 6. could be used independently by students.
____ 7. have a strong student self-evaluation emphasis.
____ 8. caused some of my slower learners to show an unusual interest in science.
____ 9. caused some of my faster learners to show an unusual interest in science.
10. were hard for students to understand.
11. had a good balance between reading and experimental involvement.
12. had too many experiments.
13. took too much teacher preparation time.
14. involved too much class discussion.
15. had a student record book which helped reinforce learning by causing the student to record.
16. had a student record book which involved too much writing on the part of students.
17. were easy for students to understand.
18. had satisfactory kits of materials for experiments.
19. should have had a separate kit of materials for experiments for each teacher.
20. were better for my students than the science textbook materials which I have been using.
21. are a step in the right direction in the preparation of curriculum materials.
22. are too complex for teachers to use.
23. call for too much student participation.
24. are less satisfactory than other elementary science materials.
25. have a very helpful Teacher's Handbook.
26. have a very helpful Teacher's Instructional Aid Booklet.