INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
HALL, Carolyn Irwin, 1945-
THE EFFECT OF AMPLIFIED ELEMENTARY SCIENCE
READING MATERIALS UPON THE COMPREHENSION
OF UPPER GRADE ELEMENTARY SCHOOL CHILDREN.

The Ohio State University, Ph.D., 1973
Education, curriculum development

University Microfilms, A XEROX Company, Ann Arbor, Michigan
THE EFFECT OF AMPLIFIED ELEMENTARY SCIENCE READING MATERIALS UPON THE COMPREHENSION OF UPPER GRADE ELEMENTARY SCHOOL CHILDREN

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University

By

Carolyn Irwin Hall, B.S.Ed., M.A.

* * * * *

The Ohio State University
1973
ACKNOWLEDGEMENTS

The writer acknowledges with gratitude the help of her committee in completing the study and her graduate program: Dr. Marlin Languis, Dr. James Duncan, Dr. Arthur White, and Dr. Alexander Frazier. Also the writer is grateful to the principal, fourth and fifth grade teachers and students of Evening Street Elementary School, Worthington, Ohio, for their cooperation in carrying out the pilot study and to the principals, fifth grade teachers and students of the following schools of Columbus Public Schools, Columbus, Ohio, for their cooperation in carrying out the study: Alpine, Barnett, Berwick, Courtright, Devonshire, Forest Park, Homedale, Maize, Northtowne, Parkmoor, Valley Forge, and Walden. A special thanks is given to the writer's husband, David; her parents; family; and friends for their support and encouragement during the entire endeavor.
VITA

June 7, 1945 ........ Born, Fort Bragg, North Carolina

1967 ................ B.S.Ed. Southern State College, Magnolia, Arkansas

1967-1968 .......... Elementary Teacher, Magnolia Public Schools, Magnolia, Arkansas

1969 ................ M.A. The Ohio State University, Columbus, Ohio

1969-1970 .......... Elementary Science Resource Teacher, Columbus Public Schools, Columbus, Ohio

1970-1971 .......... Teaching Associate, Early and Middle Childhood Education, The Ohio State University, Columbus, Ohio

1968-1972 ......... University Fellow, The Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Early and Middle Childhood Education

Studies in Elementary Science Education. Professor Marlin L. Languis

Studies in Early and Middle Childhood Education. Professor Alexander Frazier

Studies in Curriculum and Instruction. Professor James K. Duncan
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION TO THE PROBLEM</td>
<td>1</td>
</tr>
<tr>
<td>Rationale</td>
<td></td>
</tr>
<tr>
<td>Problem Statement</td>
<td></td>
</tr>
<tr>
<td>Hypotheses</td>
<td></td>
</tr>
<tr>
<td>Definitions of Terms</td>
<td></td>
</tr>
<tr>
<td>Assumptions</td>
<td></td>
</tr>
<tr>
<td>Limitations</td>
<td></td>
</tr>
<tr>
<td>Overview of Procedures</td>
<td></td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td>25</td>
</tr>
<tr>
<td>Literature Related to the Descriptions of Concepts</td>
<td></td>
</tr>
<tr>
<td>Literature Related to Difficulty of Concepts</td>
<td></td>
</tr>
<tr>
<td>Literature Related to the Written Communication of Elementary Science Concepts</td>
<td></td>
</tr>
<tr>
<td>Literature Related to the Cloze Procedure as a Method of Measuring Reading Comprehension</td>
<td></td>
</tr>
<tr>
<td>III. PROCEDURE</td>
<td>70</td>
</tr>
<tr>
<td>Pilot Study</td>
<td></td>
</tr>
<tr>
<td>Population and Sample</td>
<td></td>
</tr>
<tr>
<td>Instrumentation</td>
<td></td>
</tr>
</tbody>
</table>

iv
# TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td></td>
</tr>
<tr>
<td>Data Collection</td>
<td></td>
</tr>
<tr>
<td>Data Analysis</td>
<td></td>
</tr>
</tbody>
</table>

## IV. RESULTS ........................................ 91

- Number of Subjects
- Reliability Estimates for the Cloze Tests
- Analyses of Variances of the Cloze Test Scores
- Correlations for the Cloze Test Scores and Interview Ratings
- Summary

## V. SUMMARY, CONCLUSIONS, IMPLICATIONS ................ 119

### APPENDIX

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>128</td>
</tr>
<tr>
<td>B</td>
<td>173</td>
</tr>
</tbody>
</table>

### FOOTNOTES ............................................ 207

### BIBLIOGRAPHY .......................................... 208
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Number of Students Participating in the Study by Cells</td>
<td>92</td>
</tr>
<tr>
<td>2.</td>
<td>Reliability Estimates of Cloze Tests</td>
<td>94</td>
</tr>
<tr>
<td>3.</td>
<td>Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of the Sedimentary Rocks Passage by Levels of Interest</td>
<td>96</td>
</tr>
<tr>
<td>4.</td>
<td>Analysis of Variance of Cloze Scores by Amplification for the Sedimentary Rocks Passage and by Interest</td>
<td>97</td>
</tr>
<tr>
<td>5.</td>
<td>Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of the Deciduous Trees Passage by Levels of Interest</td>
<td>99</td>
</tr>
<tr>
<td>6.</td>
<td>Analysis of Variance of Cloze Scores by Amplification for the Deciduous Trees Passage and by Interest</td>
<td>100</td>
</tr>
<tr>
<td>7.</td>
<td>Post Hoc Multiple Comparisons for the Amplification Factor for the Deciduous Trees Passage</td>
<td>101</td>
</tr>
<tr>
<td>8.</td>
<td>Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of the Insect Passage by Levels of Interest</td>
<td>103</td>
</tr>
<tr>
<td>9.</td>
<td>Analysis of Variance of Cloze Scores by Amplification for the Insects Passage and by Interest</td>
<td>104</td>
</tr>
<tr>
<td>10.</td>
<td>Post Hoc Multiple Comparisons for the Amplification Factor on the Insects Passage</td>
<td>105</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>11. Percentages of Subjects Having the Concrete Empirical Props Associated with the Passage Exemplars in their Cognitive Structures</td>
<td>108</td>
<td></td>
</tr>
<tr>
<td>12. Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of Ss Who Had 67 to 100 Percent of the Relevant Concrete Props in their Cognitive Structures</td>
<td>109</td>
<td></td>
</tr>
<tr>
<td>13. Analyses of Variances of Cloze Scores by Amplification of the Ss Who Had 67 to 100 Percent of the Relevant Concrete Props in their Cognitive Structures</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>15. Correlations between Cloze Scores and Interview Ratings</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>16. Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of Sedimentary Rocks Passage by Levels of Interest</td>
<td>184</td>
<td></td>
</tr>
<tr>
<td>17. Analysis of Variance of Cloze Scores by Amplification for the Sedimentary Rocks Passage and by Interest</td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>18. Analysis of Variance of the Simple Effects of Amplification at Levels of Interest and Interest at Levels of Amplification for the Sedimentary Rocks Passage</td>
<td>188</td>
<td></td>
</tr>
<tr>
<td>19. Post Hoc Multiple Comparisons for Amplification at High Interest on the Sedimentary Rocks Passage</td>
<td>189</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>20.</td>
<td>Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification for the Deciduous Tree Passage and by Levels of Interest</td>
<td>190</td>
</tr>
<tr>
<td>21.</td>
<td>Analysis of Variance of Cloze Scores by Amplification for the Deciduous Tree Passage and by Interest</td>
<td>191</td>
</tr>
<tr>
<td>22.</td>
<td>Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of the Insect Passage by Levels of Interest</td>
<td>192</td>
</tr>
<tr>
<td>23.</td>
<td>Analysis of Variance of Cloze Scores by Amplification for the Insect Passage and by Levels of Interest</td>
<td>193</td>
</tr>
<tr>
<td>27.</td>
<td>Reliability Estimates of Cloze Tests</td>
<td>199</td>
</tr>
<tr>
<td>28.</td>
<td>Two by Two Table for Sedimentary Rock Passage</td>
<td>202</td>
</tr>
<tr>
<td>29.</td>
<td>Two by Two Table for Deciduous Tree Passage</td>
<td>202</td>
</tr>
<tr>
<td>30.</td>
<td>Two by Two Table for Insects Passage</td>
<td>202</td>
</tr>
<tr>
<td>31.</td>
<td>Correlations between Cloze Scores and Interview Ratings</td>
<td>204</td>
</tr>
<tr>
<td>32.</td>
<td>Correlations between Cloze Scores and Reading Level Scores</td>
<td>206</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1.</td>
<td>Digraph of Kretschmer paragraph</td>
<td>9</td>
</tr>
<tr>
<td>2.</td>
<td>Digraph of expanded Kretschmer paragraph</td>
<td>11</td>
</tr>
<tr>
<td>3.</td>
<td>Modified digraph of Kretschmer paragraph</td>
<td>11</td>
</tr>
<tr>
<td>4.</td>
<td>Showalter's model for the structure of science</td>
<td>32</td>
</tr>
<tr>
<td>5.</td>
<td>General data matrix for the study</td>
<td>83</td>
</tr>
<tr>
<td>6.</td>
<td>Extended posttest-only control group design used in the study</td>
<td>84</td>
</tr>
<tr>
<td>7.</td>
<td>Flow chart of the interview phase of the study</td>
<td>88</td>
</tr>
<tr>
<td>8.</td>
<td>Data matrix for sedimentary rocks passage</td>
<td>181</td>
</tr>
<tr>
<td>9.</td>
<td>Data matrix for deciduous trees passage</td>
<td>182</td>
</tr>
<tr>
<td>10.</td>
<td>Data matrix for insects passage</td>
<td>183</td>
</tr>
<tr>
<td>11.</td>
<td>The interaction of interest and amplification in the process of comprehending elementary science reading material on sedimentary rocks</td>
<td>186</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION TO THE PROBLEM

The low readability of elementary science textbooks poses a definite problem in their utilization (Carin & Sund, 1970; Bond & Tinker, 1967). Many teachers, particularly those working with grade level and below grade level readers, complain that there are too many difficult words in science texts and that there is too much information packed into too little space in them. Readability studies of elementary science texts tend to support these complaints, for they indicate most science text series are one to two grade levels above that for which they were intended¹ (Mallison, 1964; Ottley, 1965).

This study examines one aspect of the elementary science text utilization problem - that of providing a writing guideline for increasing the readability of elementary science written material by reducing its conceptual difficulty. The problem investigated is to determine the effect of amplification, through the use of exemplars, upon upper elementary children's reading comprehension of written elementary science materials. The rationale of the study is

¹
discussed in four sections: science written materials as a tool in meeting the objectives of elementary science education, utilization of readability formulas to increase readability of written material, amplification of written materials as a means of increasing readability, and the justification for the study.

Science Written Materials as a Tool in Meeting the Objectives of Elementary Science Education

America's schools have been reading schools - probably out of convenience, for it simplifies the budget, curriculum objectives, content, methods of instruction, and methods of evaluation (Hill, 1967). Today, in spite of the influence of the progressive era and of the recent progressive-like curricular thrusts, a majority of teachers still use the text as the primary means of dispensing information (Hilton, 1969). However, even if one bemoans the prevalent practices and wishes for all elementary school science programs to be more activity centered, it is difficult to deny that books can be useful in developing scientifically literate persons in the public schools. Carin and Sund list several ways they feel the reading of science information contributes to the objectives of science in the elementary school.

1. Understanding science words develops from first hand experiences and is deepened and sophisticated by exposure to dictionaries, encyclopedias, textbooks, and tradebooks.
2. In upper grades children use reading to test, verify, and extend the information observed in their limited experiments.
3. Children's science reading may also contribute to recreational enjoyment.
4. Science reading exposes children to the ideas of people other than themselves and their teachers - especially scientists who have conducted experiments impossible to perform in elementary schools.
5. There is 'occupational' value in having children read about the lives and works of famous scientists.
6. Under the teacher's guidance, students can develop the skills necessary to analyze science materials for accuracy and proper interpretation (Carin & Sund, 1970, pp. 158-159).

Obviously, advocating the use of books in the above described manner is not suggesting an elementary science program totally dependent on books. It is suggesting that reading material can be one of several useful mediums. Further this medium is considered useful generally only after elementary children have had first hand experiences. Many of today's "hands on" curriculum projects use the research of Piaget and his followers to insist that elementary school children cannot really learn science concepts unless they interact personally with the actual objects and phenomena which are involved in science concepts (Hurd & Gallagher, 1968). After this point, this writer as well as Carin and Sund (1970) contend that reading material can extend children's knowledge of science, and therefore reading material is a viable tool in elementary science programs. David Ausubel also advocates this approach and has set forth a theory - that of meaningful verbal reception learning - part of which
takes the above sequence into account. He states that once elementary children have concrete empirical props in their cognitive structure, they can comprehend or meaningfully learn verbally or symbolically expressed concepts about those props. This is done by what he calls a subsumption process - in which material enters a cognitive structure, interacts with it, and then is subsumed under the appropriate conceptual system in that cognitive structure (Ausubel, 1963).

Therefore, while probably it is desirable to use reading materials as one of a variety of modes of instruction in an elementary science program, there is a problem with the nature of the available written materials: they do not seem to be well suited for the children for which they were intended.

**Utilization of Readability Formulas to Increase Readability of Elementary Science Written Materials**

Readability measures were developed to put the right book in the right child's hands. The readability formulas are regression equations which utilize counts of linguistic elements such as number of uncommon words, length of sentences, number of prepositional phrases and affixes, in predicting the difficulty of reading a given passage. While the formulas have been useful and are bringing pressure to bear upon textbook publishers, they have serious limitations. Their validity as determined in cross validation studies
runs from .5 to .7, and the standard error is about one grade level (Chall, 1958). Also they do not account for the conceptual difficulty of a passage. For example, "To be or not to be; that is the question." is made up of common words and two short sentences. However, it is obviously rather abstract. Thus one can see it would not be wise to utilize a readability formula as the sole criterion in writing materials.

Two studies tend to bear this out also. Drake (1966) used the Flesch readability formula to rewrite physics material from thirteenth grade level to ninth grade level. Results indicated that students who read the material rated higher in readability did not show any better comprehension of the material than those who read the lower readability material. Marshall (1957) got a similar result when he rewrote tenth grade technical material at fifth grade level using the Dale-Chall readability formula.

A recently developed readability measure called the cloze procedure seems to be a better predictor of reading difficulty than the formulas. In it, every nth word is deleted, and the subject must fill in the blanks with the exact word. The extent to which he can do this is the measure of readability. This procedure appears to take all aspects of readability into account, and therefore it probably is a better measure of readability than the formulas are; but since it does so in such an integrated fashion, it
doesn't offer specific guidelines for writing more readable materials.

Therefore, reading formulas, while suggesting specific ways for increasing readability as measured by them, do not get at all aspects of readability - specifically, conceptual difficulty (Klare, 1963). The cloze procedure as a measure of readability, does seem to take this other aspect into consideration, but it offers no suggestions for increasing readability.

Amplification of Elementary Science Written Materials as a Means of Increasing Readability

This study deals with writing guidelines for conceptual difficulty. At the beginning of this chapter it was stated that educators felt that there are too many difficult words in elementary science texts and that there is too much information packed into too little space in them. Intuitively, it seems reasonable to the writer to hypothesize that the science concepts which are being presented in elementary science textbooks are just not being adequately explained. Or to hypothesize it another way, if the relationship a science concept expressed were described in more detail, children's comprehension would increase; conceptual difficulty would be lessened. To this one might ask if describing a concept in more detail might increase the information load and thereby increase the conceptual difficulty. However, Bruner's (1960) argument for organizing information around
conceptual schemes would tend to counter that argument. He says that by organizing information to be presented to students around concepts, the information can be learned easier because it is more relatable. In a similar vein, Ausubel (1963) suggests that one factor influencing meaningful reception learning is the extent to which the material is discriminable from the concepts that subsume it. To make material more discriminable, one can explicitly point out similarities and differences between the new material and its presumed subsumers in the learner's cognitive structure. Another factor — that of retaining new material — is a function of the stability and clarity of the material's subsumers. Ausubel thinks that probably using exemplars, repeating, and using multi-contextual exposures will enhance the clarity and stability of the subsumers. Therefore, it is logical to assume that a passage describing a concept in more detail would not be more difficult to comprehend. Also, several studies suggest that increasing the length of a passage on a certain topic increases children's comprehension.

Wilson (1944) wrote three versions of "The History of Paper," "Paper Today," and "The Paper Industry" in 300 words, 600 words, and 1200 words. She found that even though the Gray-Leary readability formula rated the 1200 version two grades higher than the 300 word version, children's comprehension was significantly higher on the longer version. From her study, Wilson noted "test items that
measured comprehension of concepts farthest from the experience of children showed the highest percentage of accuracy for the group which read the detailed version" (Wilson, 1944, p. 4). Williams (1968) rewrote sixth grade science text material by substituting nontechnical words, amplifying technical vocabulary through the addition of phrases and sentences, and rephrasing to make thoughts more clear and distinct. He found that the sixth grade students who read the rewritten material scored significantly higher on the comprehension test he gave.

In the preceding studies, the different amounts of detail in passages and the different versions of material could be considered different degrees of amplification. The amplification, however, has been loosely defined. To more precisely operationalize amplification, the writer suggests two things - Kretschmer's (1972) work, in which some of Piaget's concrete operations are defined in digraph theory and a well known pedagogical method - using examples - which is explained in Ausubel's theory of meaningful reception learning.

The Kretschmer approach, originally from Frase (1969), holds that the level of conceptual difficulty of a passage is dependent upon the level of inference which a person can or must make in order to comprehend the passage. For example, from Kretschmer:
Two common types of mammals are rodents and carnivores. Rodents have teeth adapted for gnawing, while the teeth of carnivores are designed for tearing. Rodents are a diverse group, including small species hardly longer than an inch, to larger species two to three feet in length. Species of rodents live in both eastern and western areas of the United States.

The digraph of the preceding passage is shown in Fig. 1.

![Digraph of Kretschmer paragraph.](image)

Digraphs enable one to conceptualize the structure of something through the use of points and arrows. In this case that something is the concept "Mammals with teeth adapted for gnawing are called rodents." The points represent classes of animals, and the arrows represent the relationship "includes" (due to function of teeth). The arrow between two points is considered a path of length one; and in this case, to say "A rodent is a mammal" is a literal assertion. Two arrows are considered a path of length two; and in this case, to say "A small western rodent is a mammal" is a first degree inference. Therefore, the above passage has the potential of being conceptually difficult to a first degree inference.
Using examples of or illustrating a particular topic is an age-old pedagogical principle. It is often used by teachers, writers, speakers, anyone who is trying "to get something across." Hough and Duncan (1970), in discussing principles that guide the use of direct communication strategies, suggest that one way to increase comprehension of the main point or "figure" is to clarify the nature of the figure and its relationship to the learner's background. One way to do this clarification is by illustrating the main point. Similarly, Dale and Chall (1956) and Flesch (1943), in discussing how to make reading material more readable, have both suggested using "concrete" examples. Ausubel has perhaps theorized most clearly how examples or illustrations are used in the process of learning. As previously mentioned, he thinks one way to enhance the clarity and stability of subsumers - one of the important factors in meaningful reception learning - is to use examplars. Therefore, it is with respect to his particular theory that the writer will describe the use of examplars as a method of amplification. For example, "two kinds of small western rodents are kangaroo mice and prairie dogs" could be added to the Kretschmer paragraph. The resulting digraph is shown in Fig. 2.
In terms of Ausubel's (1963) theory, the idea of kangaroo mice or prairie dogs having gnawing type teeth may possibly be relevant concrete props in the cognitive structure of the reader thereby rendering the statement "A prairie dog is a mammal" a literal assertion rather than a second degree inference. Therefore, what at first might have appeared to be making a passage conceptually more difficult is actually making it easier. A digraph showing the process might look like that shown in Fig. 3.

Naturally it would be difficult to assume that one or more exemplars would effectively reduce the potential conceptual difficulty of a passage for a great many readers, such as
the audience for an elementary text series. But it is reasonable to suggest that the greater the number of exemplars in a passage on a given concept, the more likely it is that one or more of the exemplars would be in the cognitive structure of the reader - thereby making the passage more readable.

**Justification for the Study**

In the preceding paragraphs, the writer has established that there is a reading difficulty problem associated with many current elementary science reading materials which should be solved because it is desirable to use reading materials in an elementary science program. It has further been established that using the controlling factors of readability formulas as guidelines for controlling readability of elementary science material is not adequate because they do not control conceptual difficulty - a definite aspect of readability. Therefore, the writer concluded that there is a definite need for studies which would suggest some guidelines for controlling conceptual difficulty of written elementary science materials. Review of literature on amplification of written material, use of examples and illustrations, and Ausubel's theory of meaningful reception learning, and some thought on the matter suggested that perhaps one significant factor relating to conceptual difficulty of written elementary science material is the level of amplification of the concept presented. Therefore, in order to
establish the importance of this factor in the process of comprehending science concepts, this study was undertaken to determine systematically if amplified versions of passages on particular concepts could significantly increase elementary children's reading comprehension scores on those passages.

Although the results from such a study as described above can not be generalized to all science concepts, it seemed to the writer to be an important step toward establishing guidelines concerning conceptual difficulty for writers of elementary science material - be they professional writers or classroom teachers!

**Interest as a Blocking Variable**

From work with elementary children, the writer has become aware that some children, when keenly interested in a particular topic, will successfully read materials which are above the level of their normal reading materials or will read more successfully material which is at their normal reading level. This phenomenon is noted in reading texts (Russell, 1961) and has been the topic of several studies. Fader and McNeil (1968) developed English in Every Classroom, a program designed to provide motivation for reading and writing in children at Maxey's Boys Training School. They found, among other things, that their students improved more than the control group in attitude toward books, magazines, and newspapers and also improved more in their
capacity to absorb paragraph meaning. Bernstein (1955) studied the relationship between adolescent reading interests and comprehension. She found it to be significant and concluded that interest seems to operate to enhance reading efficiency.

Thus it seems reasonable to hypothesize that motivation as implied from the interest level of children, would tend to confound the effects of amplification on comprehension scores on a particular passage; In order to control for this variable, the writer used it as a blocking variable in this study to determine the importance of the level of amplification of written materials in the process of comprehending those materials.

Extent of the Presence of Concrete Props in Subjects' Cognitive Structures

The rationale for the above described approach to the problem is based somewhat on Ausubel's notion that elementary children can comprehend verbally presented materials if they have the necessary concrete props in their cognitive structures. This notion is intuitively true, but it is without adequate scientific support. This study also would not lend scientific support because the necessary concrete props in the cognitive structures of the subjects is not one of the study variables. As such the study results would not indicate if the subjects had the necessary concrete props in their cognitive structures.
Therefore, while it was not the purpose of this study to test Ausubel's notion concerning the necessity of the presence of concrete props; it did seem important to extend the breadth of the study to include a means by which this factor could be scientifically described. To do this, the writer empirically determined the extent of the presence of the relevant exemplars as concrete props in the cognitive structures of the subjects and then correlated these measures with the measures of comprehension of the amplified written material.

Problem Statement

Main Problem

Will amplifying passages, that treat three selected elementary science concepts through the use of exemplars, affect the reading comprehension of upper elementary school children on those passages?

Subproblems

Is level of interest in three selected elementary science concepts related to the reading comprehension of upper elementary school children on passages that treat those concepts?

What is the relationship between level of upper elementary school children's reading comprehension on passages
that treat three selected elementary science concepts and extent of presence of concrete-empirical props associated with those concepts in the cognitive structures of those children?

Hypotheses

1. There will be no significant differences (.05) between the cloze test scores of subjects who read high, moderate, or low amplification versions of passages that treat three elementary science concepts.

2. There will be no significant differences (.05) between the cloze test scores, on passages that treat three elementary science concepts, of subjects who have received high, moderate, or low interest scores.

3. There will be no significant interaction (.05) between the amplification levels and the interest levels.

4. There will be no significant correlations (.05) between cloze test scores of subjects on passages that treat three elementary science concepts and interview ratings. The interviews are designed to assess the extent of the presence of concrete-empirical props which are associated with the three elementary science concepts treated in the passages in the cognitive structures of the same subjects.
Definition of Terms

1. **amplified passage** - reading material which has been written to include exemplars of the concepts described
   a. **low amplification** - inclusion of only one exemplar
   b. **moderate amplification** - inclusion of three exemplars
   c. **high amplification** - inclusion of five exemplars

2. **reading comprehension score** - understanding of the concept contained in the written material, as measured by cloze tests.

3. **cloze test** - a test which measures reading comprehension by requiring students to fill in missing words which have been systematically deleted from written material

4. **science concept** - relationship between ideas. In this study science concepts will be limited to real classificational concepts, real concepts being generalizers of perceptions (Showalter, 1969) and classificational concepts being those concepts that facilitate description of phenomena (Pella, 1966). The three concepts used were small animals with six legs are called insects; trees which lose their leaves once a year are called deciduous; and rocks which are formed from little pieces of other rocks are called sedimentary.

5. **upper elementary school children** - in this study upper elementary school children are defined as children in
the fifth grade who are between the ages ten years, zero months and eleven years, zero months, who have reading abilities within the range 4.0 to 5.0 grade levels as measured by the California Comprehensive Reading Test, Level 2, Form Q, and who are from schools having no families receiving Aid-to-Dependent-Children funds.

6. interest - a desire to know more based on a clearly defined goal connected to that interest (Jenkins, 1969). In this study interest will be measured by an interest inventory constructed by the writer (see Appendix A).
   a. low interest - in this study low interest will be defined as the lower third of the range of scores obtained on the interest inventory
   b. moderate interest - in this study moderate interest will be defined as the middle third of the range of scores obtained on the interest inventory
   c. high interest - in this study high interest will be defined as the upper third of the range of scores obtained on the interest inventory

7. interest inventory - an instrument developed by the writer which measures the level of interest in each of the three elementary science concepts by requiring the students to circle the number of the statements which are true for them. There are twelve pairs of statements for each concept, one positively stated and one
negatively stated.

8. **conceptual difficulty** - comprehensibility of a passage as inferred by the depth of inference required to comprehend the material, depth of inference being inferred by digraph analysis (Frase, 1969 and Kretschmer, 1972)

9. **exemplar** - an example of the concept treated in the passage. In this study, the concepts were chosen with the criterion that they were likely to be referents to concrete-empirical props in the cognitive structures of the subjects (see Chapter III).

10. **concrete-empirical prop** - a cognitive support resulting from direct nonverbal contact with objects or situations involved in a concept (Ausubel, 1963). In this study the presence of concrete-empirical props, in the cognitive structures of the subjects, which are associated with the exemplars in the concept passages, will be determined by an interview (see Appendix A).

11. **interview** - a standardized personal interview developed by the writer which determines the extent of the presence of concrete-empirical props associated with the exemplars in the passages in the cognitive structures of the subjects. There are fifteen basic questions, one for each concept exemplar, in the form "can you describe a ________?"

12. **cognitive structure** - "... an individual's organization, stability, and clarity of knowledge in a particu-
lar subject matter field at any given time" (Ausubel, 1963, p. 26)

Assumptions

1. Elementary children can learn science concepts by reading about them if they have the necessary concrete props concerning the concepts in their cognitive structures.

2. Levels of conceptual difficulty of written passages can be inferred by the depth of inferences required to comprehend the material, depth of inference being illustrated by digraph analysis.

3. Real classificational concepts are within the comprehensibility of the concrete operations child.

4. Piaget's logical operations, which include classificational operations, can be represented in written passages.

Limitations

1. Generalizations concerning the independent variable - amplification - are limited to interpolations within the range of the levels of amplification used in the proposed study - one, three, and five exemplars.

2. Generalizations concerning real classificational concepts are limited to the three concepts used in the study.
3. Generalizations concerning the population from which the sample was taken - children who are between the ages ten years, zero months, and eleven years, zero months and who have a reading ability within the range 4.0 to 5.0 grade levels - are valid only to the extent that the sample was representative of that population. These generalizations are further limited by the fact that the most recent reading achievement scores on the sample were made approximately one year prior to the time of the study.

4. Generalizations concerning the independent variable - interest - are limited to the levels of interest which are used in the study.

5. The only inferences concerning causality which can be made are those concerning amplification, since it is the only active independent variable in the study. Interest was an assigned independent variable, and therefore generalizations concerning it are only descriptive.

6. Generalizations about the relationship between cloze test scores and interview ratings are also only descriptive. These generalizations are limited further by the possibility of the interview ratings being positively biased as a result of rote learning which might have taken place in the reading of the passages.
Overview of Procedures

Pilot Study

All the following procedures were piloted prior to the formal study (see Appendix B for a complete report).

Design

The design of the study was an extension of the post-test-only control group design. It was a three by three treatment by levels design repeated for each of the three elementary science concepts. There were two independent variables: (1) amplification of the elementary science reading material, an active variable, of which there were three levels - high, moderate, and low amplification and (2) interest in the elementary science concepts used in the study, an assigned variable, of which there were three levels - low, moderate, and high interest. The dependent variable was reading comprehension.

The design of the interview phase of the study was that of a simple survey. Two variables were involved: (1) reading comprehension as measured by cloze tests, and (2) extent of the presence of concrete-empirical props associated with the concept exemplars as determined by interviews.

Population and Sample

The population used in this study consisted of elementary children who were between the ages ten years, zero months and eleven years, zero months and whose reading
ability was between 4.0 and 5.0 grade levels. The sample was confined to 135 children from twelve schools in the Columbus Public School System who were designated to be of approximately the same socio-economic background and who were willing to participate in the study.

**Instrumentation**

There were three instruments used in the study: (1) a cloze test to measure reading comprehension of each of the amplified passages, (2) an interest inventory to measure interest in each of the elementary science concepts, and (3) a standardized interview to measure the extent of the presence of concrete-empirical props associated with the elementary science concept exemplars used in the study.

**Treatment**

The treatment consisted of amplified passages on three different elementary science real classificational concepts - one on deciduous trees, one on sedimentary rocks, and one on insects. For each concept, there were three levels of amplification - low, moderate, and high - consisting respectively of one, three, and five exemplars of the given concept. All the passages had the same readability level - fourth grade; and all had the same potential conceptual difficulty - first level inference.

**Data Collection**

The subjects all took the interest inventory first and were blocked into the three interest groups. Then the three
different levels of amplified passages and cloze test combinations were administered to the subjects separately in three sessions. Finally the interviews were given to each subject individually.

**Data Analysis**

A two-way analysis of variance was used to test the first, second, and third hypotheses of the study. The fourth hypothesis of the study was tested by calculating a Pearson Product Moment Correlation Coefficient.
CHAPTER II

REVIEW OF RELATED LITERATURE

The review of literature related to the study of elementary children's comprehension of elementary science written materials treating selected concepts is presented in this chapter. The chapter is divided into four sections. The first section deals with literature related to the descriptions of concepts. It provides a partial basis on which to define and select the elementary science concepts used in the study. The second section deals with literature related to the difficulty of concepts. It provides the rest of the basis on which to define and select the elementary science concepts used in the study. The third section deals with literature related to the written communication of concepts. It provides the theoretical background for the study. The fourth section deals with literature related to the use of cloze tests as a method of measuring reading comprehension. It provides a basis for justifying the use of cloze tests as the instrument for the measurement of reading comprehension in the study.
Overview

This study examines one aspect of readability - conceptual difficulty. In considering how to control the conceptual difficulty of a passage, one runs into a problem - agreement on the nature of a concept.

Psychologists, while having different notions about how concepts are attained, seem to agree basically on a definition of concept. For example, John Carroll describes concepts as "...properties of organismic experience - more particularly, they are the abstracted and often cognitively structured classes of 'mental' experience learned by organisms in the course of their life histories" (Carroll, 1968, p. 301). Bruner, Goodnow and Austin suggest that "...the working definition of a concept is the network of inferences that are or may be set into play by an act of categorization" (Bruner, Goodnow, & Austin, 1956, p. 244). More simply, Earl Hunt wrote a concept is "...any identifying response to members of a set of not completely identical stimuli" (Hunt, 1962, p. 6). Gagne sums the many definitions up by saying

All the definitions have some general properties in common, and I judge these to be as follows.
1. A concept is an inferred mental process.
2. The learning of a concept requires discrimination of stimulus objects (distinguishing 'positive' and 'negative' instances).
3. The performance which shows that a concept has been learned consists in the learner being able to place an object in a class. The common examples of concept learning which would presumably be acceptable to psychologists might include: learning chair as a class of objects, learning red as a property of objects detachable from particular objects, learning classes of direction or position such as up, down, right, or left. From the standpoint of the psychologist, therefore, the notion of a concept as an inferred process which enables the individual to classify objects is both prominent and widely accepted (Gagne, 1966, p. 83).

As previously mentioned in Chapter I, one of the main practices in public education in our society seems to be to dispense information. The curriculum revolution of the late fifties and sixties used concepts as one of the key structuring devices for the information to be dispensed (Michaelis, Grossman, & Scott, 1967). In looking at some of the new elementary science materials, one sees example concepts given as "interaction," "object," and "property" (Science Curriculum Improvement Study, 1967). Other examples are given as "mechanical energy is conserved" and "energy is the ability to do work" (Copes, 1967). Obviously, educators' concepts include those which are too complex to be used in the classic concept formation experiments conducted by psychologists. Yet educators still refer to them as concepts. Further, these same concepts may be labeled by others, depending on their orientation, as facts, generalizations, principles, conceptual threads, etc. Because educators have generally had a broader definition of concept than psychologists, it has been necessary to specify kinds and
levels of concepts in order for them to be usable by curriculum developers and persons involved in educational research. This section of Chapter II will recount some of the literature reviewed by the researcher in order to define, select, and justify the concepts used in this study.

**Descriptions of Concepts on a Horizontal Plane**

Rapparlie (1971) suggests all concepts can be divided into three broad areas according to which aspect you are paying the greatest attention. The broad areas are those of (1) structure, (2) function, and (3) quality. Concepts which focus attention on the structural aspects are those that order and organize the universe of the individual. Generally, the nominal words stand for structural concepts. An example of a structural concept is "dog." Concepts which focus attention on the functional aspects of the referent are primarily concerned with doing and the consequences of doing. From the functional line of concepts, come concepts which lead to the formulation of principles of behavior. This is the cause and effect area of relationships. In terms of language, this group would be basically verbal. Concepts which pay the greatest attention to the qualities of the referent such as shape, color, morality, gracefulness serve the role of defining. They are also used as a prior step to the correct use of structural concepts.

In a similar fashion, but looking specifically at the area of science, Smith (1966) also says concepts differ
intrinsically in kind. She, however, distinguishes only two kinds. Some are "class-type" constructs such as reptile and mammal. This kind seems to correspond to Rapparlie's structure category and quality category. The other concepts seem to be summarizers of experience which are unrelated to the class-type categories. An example is "snow melts when it gets warm." This corresponds somewhat to Rapparlie's concept category related to function.

Pella (1966) also, looking specifically at science, suggests there are three systems in an epistemology of science concepts - classificational, correlational, and theoretical. Classificational concepts make up that system of concepts that facilitates description of phenomena. An example is "an insect is an animal with six legs and three main body divisions." Correlational concepts make up that system of concepts that facilitates prediction. An example is "a force is a push or pull which tends to change the motion of a body." Theoretical concepts make up that system of concepts that facilitates the explanations of phenomena. An example is "an atom is the smallest particle of an element possible and is composed of electrons, protons, neutrons, and other particles." The first two systems correspond respectively to the structure and class-type concepts and the function and summarizer-of-experience concepts formulated by Rapparlie and Smith. Pella however goes a step farther saying that the first two systems are abstractions
based on direct human experience with natural phenomena, while the third system is not based on direct human experience with natural phenomena.

So, here were three somewhat similar views on how concepts could be classified on a horizontal plane. How are these category systems of value to a curriculum designer? Rapparlie (1971) suggests different kinds of concepts may be acquired more efficiently through different kinds of methods. Therefore, knowledge of what kind of concept development is desired might be very essential to a curriculum developer. Another value which this writer sees is that if schools are to dispense knowledge and if the logical structure of knowledge is a conceptual structure (Bruner, 1963) and if concepts are of two or more types, then curriculum developers should include both or all types. Thus, knowledge of horizontal types of concepts would help provide a balance in the knowledge dispensed by schools.

Descriptions of Concepts on a Vertical Plane

Novak (1966) thinks it would be better to accept the broad definition of concept and its synonyms and construct a taxonomy of conceptual levels analogous to Bloom's Taxonomy of Educational Objectives. At the lowest levels would be the concepts referred to by experimental psychologists; and at the highest levels would be concepts of the type referred to by the biologist when he talks of the concept of "evolution." [It should be noted that in measuring
attainment of particular concepts, Pella and his associates have actually used Bloom's Taxonomy (Pella & Struass, 1969).] Novak's idea is a description of concepts on a vertical plane, and it's value to the curriculum developer is obvious. Knowledge of conceptual levels would help in properly sequencing the curriculum. Actually this has been going on for sometime through the so-called "spiral" curriculum in which each level of the spiral introduced a more sophisticated level of concept. An example of a set of conceptual levels is the following devised by Brandwein (Carin & Sund, 1970).

**Conceptual Scheme:** The universe is in constant change.

- **level 1:** There are daily changes on earth.
- **level 2:** There are regular movements of the earth and moon.
- **level 3:** There are seasonal and annual changes within the solar system.
- **level 4:** The motion and path of celestial bodies is predictable.
- **level 5:** Universal gravitation and inertial motion govern the relations of celestial bodies.
- **level 6:** Nuclear reactions produce the radiant energy of stars.

Showalter (1969) also prefers the general definition of concept - a generalizer of experiences, and upon that he has built a model for the whole structure of science. This model is made up of levels of generalizations, but each level has its own name (see Fig. 4).
At the first level of the model are perceptions. Perceptions are actually not generalizers of experience, but they are the basis of the generalizers. They are the multitude of encounters between a person's mind and phenomena outside a person's mind. The second level consists of real concepts. These real concepts are generalizers of perceptions. In language terms, a real concept is a word. An example is "dog," "time," "weight" etc. At the third level are facts. In language terms, a fact is a sentence. The fact describes a relationship between two or more real concepts. An example is "a drop of water formed a sphere." The fourth level is made up of laws. A law is similar to a fact in that it is a statement. However, a law subsumes several facts, and
often laws are expressed as mathematical equations. An example is Galileo's law: \( s = \frac{1}{2}gt^2 \). The first four levels of Showalter's model are empirical, that is they correspond to something in the real world. The last three are not; they are imaginary. In this sense, Showalter's model is similar to Pella's categories. Pella, too, had a concept category - theories - for concepts not based on direct human experience. The fifth level is composed of invented concepts. It is also represented by a single word like real concepts are, but it does not correspond to something in the real world. An example is "energy" and "atom." At the sixth level are principles. They describe a relationship between two or more concepts of which at least one is an invented concept. At the seventh and highest level of generality are theories. A theory is a system of related statements which may be principles or combinations of principles. An example is the "theory of evolution."

**Summary**

In this section, educational concepts have been broadly defined as "generalizers of experiences." The writer has distinguished two ways of describing concepts - one on a horizontal plane and one on a vertical plane. Horizontally, three views were discussed, those of Rapparlie, Smith, and Pella. Rapparlie suggested that concepts can be divided into three areas: those related to structure, quality, and function. Smith suggested only two: class-type and summarizers-
of-experiences. Pella suggested three: classificational, correlational, and theoretical. On a vertical plane, two views were discussed, those of Novak and Showalter. Novak suggested concepts can be thought of as being somewhere on a taxonomy of conceptual levels. One of Brandwein's schemes for a spiral curriculum was given as an example. Showalter suggested a similar model made up of levels of generalizations, going from perceptions to theories.

All these views help one to describe more specifically what kind of concept is being dealt with in any given situation. It enables one to distinguish one type of concept from another and thus improve communication and decision making about science concepts. Therefore, the first step in a study examining one aspect of readability of elementary science written materials - that of conceptual difficulty - seems to be to specify from the multitude of existing science concepts, the kind of concepts which will be dealt with. Taking from the views presented, this writer chose to deal only with real concepts and only with classificational concepts in this study.

Literature Related to Difficulty of Concepts

Overview

The preceding section has recounted how concepts are described qualitatively. Though the concept system de-
scribed by Novak, Brandwein, and Showalter are done so in a qualitative manner, they are also thought of in a quantitative fashion: how relatively difficult are the concepts one to another? This is also a question with which curriculum developers and educational researchers have to deal. Therefore, this section of Chapter II will recount the remainder of the literature reviewed by the research in order to define, select, and justify the concepts used in this study. In reviewing this literature, the writer distinguished a number of factors which are related to the difficulty of concepts. These will be discussed below.

**Number of Attributes**

Obviously, in the hierarchial systems described by these educators, the lowest levels are considered less complex and hence less difficult than the higher levels. A number of persons, in reviewing literature on concept attainment, have pointed out that research does show that the more attributes or characteristics a stimulus has, whether relevant or irrelevant or both, the more difficult the concept is to master (Seels, 1970; Rapparlie, 1970, and Archer, 1966).

**Abstraction**

Not only can concept difficulty be increased by increasing the numbers of attributes, but some people state that an increase in abstraction will cause an increase in difficulty (Rapparlie, 1970 and Smith, 1966). Abstraction
could be defined as the state of being outside the realm of human experience, not empirical. Thus Pella's category of theoretical concepts and Showalter's top three levels of generalizations - invented concepts, principles, and theories - would be more abstract than the other categories and hence more difficult.

**Human Developmental Characteristics**

Not only do concepts themselves have characteristics which make them considered more difficult to learn; but the persons who are to learn concepts also have developmental characteristics, which, when associated with certain concepts, cause those concepts to be considered more difficult. For example, if one considers the theories of Piaget (Flavell, 1963), abstract concepts, as previously defined, would be difficult if not impossible for elementary school age children to learn. Elementary children are generally within what Piaget calls the concrete stage of cognitive development (six + to eleven + years of age). In this stage, Piaget has identified several operations which can be done. These operations are restricted in that they must be done on real physical objects or, in Ausubel's (1963) interpretation, on mental representations of real physical objects. These operations have the property of grouping or groups. This property would enable elementary children to deal with classificational concepts as described by Pella. However, correlational, as well as theoretical concepts could not be
dealt with, according to Piaget, until the stage of formal operations is reached (twelve years). Here, in Piagetian theory, a child is finally able to deal with two variables simultaneously (correlational concepts) and to do hypothetic-deductive reasoning (theoretical concepts). It should be noted that research by Pella and Struass (1969) indicated some correlational and theoretical concepts could be comprehended by elementary children at the knowledge, comprehension, and application levels of Bloom's taxonomy. Also, Flavell (1963) points out that some of Piaget's theory (logical operations four and eight) is not strongly supported by empirical research. Nevertheless, this writer feels it is more clear that classificational concepts are within the comprehensibility of elementary school children than are correlational or theoretical ones. Curriculum developers should and are finding this notion helpful in deciding where to place concepts in the curriculum.

**Conceptual Background**

Another aspect of conceptual difficulty has to do with the conceptual background of the person learning the concept. Frase (1969) suggested conceptual difficulty increases as the depth of inferences needed by the individual to comprehend the concept increases. Obviously this idea could apply to any concept whether empirical or not. It applies specifically only to the particular individual's realm of experience. Thus a classificational concept could be more
difficult if it is far from a person's particular realm of experience. If no inferences are needed, evidently the concept is close to the person's realm of experience. If deep inferences are required, the concept is far from the person's realm of experience. Thinking similarly, Brandwein, Novak, and Showalter devised their hierarchies intending for higher levels to be built on lower levels. Therefore, if a person has somehow skipped a prior level, the concept to be learned would be very difficult if not impossible to learn. AAAS's Science-A Process Approach was developed on such a hierarchy of concepts (written as processes). Evaluation of this program (Walbesser, 1968) indicated students who had previous work with Science-A Process Approach materials achieved more objectives than those who had not had such previous work. Similarly, Ausubel states that if a concept to be learned is more, rather than less, subsumable within a person's existing cognitive structure, then the concept is more easily learned (1963). Ausubel (1963) has cited a number of studies which support this notion of the importance of background knowledge in concept learning - those of Poulton, Brousfield, Coher and Whilemaker, Underwood, Haslerud, and Saugstad. Research by Ausubel and Fitzgerald (1962) also supports this notion. They found students who scored in the upper third of the range of scores on a comprehension pre-test on a subject, learned a passage related to that subject significantly better than those who had
scored in the middle and lower third of the range of scores on the comprehension pre-test. Thus it appears to this writer that knowledge of conceptual background of the prospective learner would be of great importance to the curriculum developer in appropriate placing of concepts to be learned.

**Interest**

Frase (1969), mentioned earlier, has done some research which touches on still another aspect of conceptual difficulty. That is the individual's intent to process or interest in processing particular information. He attempted to cause increases in information learning by using particular orienting directions. In his research, he did find that by giving subjects orienting directions in the form of questions, he could cause significant increases in memory of information. Other research supports this idea, because in reviewing literature on conceptual difficulty, Rapparlie stated "...as the salience of (concept) attributes decreases, difficulty increases" and "...the closer the concept is related to the self concept of the individual, the easier it becomes to learn" (Rapparlie, 1966, p. 663). Also, if one can assume that comprehending reading materials is comparable to learning concepts, then several studies in reading which were mentioned previously support this notion. Fader and McNeil (1968) and Bernstein (1953) both studied the relationship between reading interests and reading comprehen-
sion and found interest seems to positively affect reading ability. Knowledge of this final aspect of conceptual difficulty would be of value to the curriculum developer in two respects. First, it would help him in selecting interesting concepts to be included in the curriculum. Second, it would help him in anticipating the necessity of planning for motivation development in the curriculum.

Summary

In this section, five factors which contribute to the difficulty of concepts were discussed. These were number of attributes, degree of abstraction, level of human development, extent of conceptual background, and interest. The first two dealt with concept characteristics: the more attributes a concept has and the more abstract a concept is, the more difficult it is to learn. The latter three factors dealt with concept characteristics in conjunction with human characteristics. The level of cognitive development a person has dictates what kind of concepts he is capable of learning; according to Piaget, the lower the level of development is, the lower the level of abstraction must be and the fewer the number of attributes must be. The extent of conceptual background also dictates how easily a concept can be learned; according to Ausubel, the less background in a particular concept which a person has, the more difficult it will be for him to learn the concept. Interest also dictates how easily a concept can be learned; generally, the
more interest a person has in a concept, the easier it will be for him to learn it.

As in section one, these ideas also help one to describe more specifically the kind of concepts which are being dealt with in a particular situation. Taking from the above information about factors important to concept difficulty, the writer made the following choices. Concepts with only two main attributes would be used. In the sedimentary rock concept, they are "squeezed together" and "made of small bits of other rocks." In the deciduous tree concept, they are "whose leaves change colors" and "fall off once a year." In the insect concept, they are "small animal" and "with six legs." The concepts are, as previously stated, "real." This limits them to the empirical realm. The writer wished to work with elementary children, and elementary children are generally within the concrete stage of cognitive development. Therefore the age ten years to eleven years was chosen to make sure these children were not still in the pre-operational stage and probably were not into the formal operations stage. Since it is clear classificational concepts can be handled at the concrete stage of cognitive development, classificational concepts were chosen for this study.

The last two factors, conceptual background and interest were not known at the outset of the study. It could not be said if the concepts were of interest to the learners
and to what extent and if the concepts were ones for which the learners had some conceptual background. So, the writer developed and administered an interest inventory and then blocked on interest in the study design in order to control that factor. Design-wise it seemed unwise to measure the conceptual background factor prior to the experiment. Therefore, the writer developed and administered an interview after the experiment in order to determine the conceptual background of the learners. The state of this background was then described by correlating the interview ratings with the cloze test scores.

Literature Related To The Written Communication Of Elementary Science Concepts

Overview

The preceding section described the nature of concepts. This section will recount literature related to one mode of concept learning - written communication of concepts. It should first be acknowledged that there is some disagreement as to whether or not concept learning through verbal channels actually is possible. This question seems to stem from a philosophic disagreement on whether concepts are associated with images or language. Postlinguistic philosophers have, however, tended to compromise the two viewpoints saying they both are legitimate aspects of concepts (Harre, 1966). Some psychologists, however, still continue to
differentiate. There are studies related to concept learning, and there are studies related to verbal learning. However, Underwood after discussing the relationships between concept learning and verbal learning states

...the study of concept learning is the study of the acquisition and utilization of common associates to different objects and events. The study of the development of associative responses is the study of verbal learning as is also the study of the implications of their elicitation after development. Given this orientation, it is sometimes difficult to make a distinction between concept learning and verbal learning (Underwood, 1966, p. 62).

That acknowledgement aside, this writer points out that this study deals with concepts in a particular form - verbal, visual, symbolic. These concepts are associated with language; they are perceived through the sense modality of vision as printed words representing systems of perceptions (Seels, 1970). In this section the writer will not attempt to recount all the psychological literature on verbal learning and/or concept learning, but just one selected aspect which seemed to be especially related to the written communication of concepts - Ausubel's theory of meaningful verbal learning. Besides this psychological literature, there is literature dealing with the effective writing of informational material. Selected aspects of this area will also be recounted.

Guidelines for Writing Good Informational Literature

Colby in discussing what makes a good children's
When children's literature was in its infancy, many authors tried to disguise facts as diction. Inanimate objects were personified as 'Mother West Wind' or 'Little Sara Seedling.' Anthropomorphism is now deplored and has largely disappeared from good science writing. Though concepts may be difficult, if they are worth presenting, they should be presented without condescension or patronization (Colby, 1967, p. 16).

Irving and Ruth Adler, authors of many science books for children have this to say about writing:

Our primary goal is to present scientific ideas so simply that they can be followed and understood by an unsophisticated reader. The science writer, too, must decide what to leave out and what to put in. He must select and organize his material so that the essential idea is developed logically and clearly and is not obscured by unnecessary detail. Of course he must be careful not to distort an idea when he simplifies it. The key to presenting a complex idea simply and accurately is to break it up into its constituent parts and then to present the parts one at a time in the proper sequence (Adler, 1965, p. 524).

Rose Wyler has also made some points on current science writing:

The seasons' new elementary science books will appeal to teachers and librarians as useful adjuncts to primary grade curriculums. Parents seeking scientific adventures with their children may not respond to them, nor will children without prior interest in the given subject. Most of the books assume the reader is seeking more information, thus do not even try to engage young readers through problems that concern them; the leads or hooks, so to speak, are lacking. The main virtue of the books is their simplicity....They thus teach children something
important, the satisfaction of finding out. Yet one cannot but wish for more than this— for books that are not only readable but communicate the thrill of discovery and the excitement of science (Colby, 1967, p. 18).

Fitz-Randolph (1969) suggests the essentials in writing informational materials are interest and understandable presentation, plus a wealth of facts. The material must be made interesting for the intended reading audience, and it should be written in language the average reader in that audience can understand.

Charlotte Huck, an authority on children's literature, has developed a list of criteria for informational books which can be used to sum up the preceding comments on writing science materials as well as to add some more.

Summary of Criteria for Informational Books

I. Accuracy and Authenticity
   1. What are the qualifications of the author?
   2. Are facts accurate?
   3. Is the book realistic?
   4. Are facts and theories clearly distinguished?
   5. Does the text avoid stereotype?
   6. Is the book up to date?
   7. Are significant details omitted?
   8. Are differing viewpoints presented?
   9. Is anthropomorphism omitted?
  10. Are phenomena given teleological explanation?

II. Content
   1. Is the coverage of the book adequate for its purpose?
   2. Is the book within the comprehension and interest range of the age for which it is intended?
   3. Do experiment books lead to understanding of science?
   4. Are experiments and activities safe and feasible?
5. Does the book present interrelationships of facts and principles?
6. Is the book fresh and original?
7. Does the book help the reader understand the methods of science?

III. Style
1. Is the information given directly?
2. Is the text interesting and appropriate for the age level intended?
3. Do vivid language and appropriate metaphor create interest and understanding?
4. Does the style create the feeling of reader involvement?
5. Is the language clear and simple or heavy and pedantic?
6. Is there an appropriate amount of detail?
7. Does the book encourage curiosity and further study? (Huck & Kuhn, 1968, p. 474)

Clearly, most of the preceding suggestions for effective writing of science materials have been somewhat subjective. There have been no objective yardsticks for measuring writing effectiveness mentioned. Many authorities in reading (Dale & Chall, 1949 and Klare, 1954) hold that writing effectiveness cannot be guided totally by scientific measures. They say writing style is partially an art and therefore cannot be so controlled. However, some aspects of writing can be subjected to more objective guidelines.

Dale and Chall (1949) specify this technique for writing readable materials: Consider three aspects of writing - defining the audience, defining the purpose and analyzing the material. Aspects of defining audience includes finding out how well a group reads, what are the groups' interests, and how much does the group know about the subject in question. This first aspect can be accomplished by using read-
ing tests, reviewing research on children's interests, and by using information surveys. Defining the purpose involves finding out exactly how the material will be used. This can be accomplished by consulting curriculum developers or curriculum guides. The third aspect of writing readable material, analyzing the material, involves using the readability formulas described in Chapter I. However, Dale and Chall suggest ignoring the formulas in writing the first draft; they say first use the more subjective writing guidelines. Then afterwards, revise the draft in the light of the formula guidelines.

Flesch, (1949) another formula developer, has given writing suggestions very similar to that of Dale and Chall. His 'Rx' for readability goes like this. First, know your prospective readers' minds; have something to say; organize that something around an interesting, perhaps original theme; start the passage with the idea of getting attention; continue writing while being concrete, vivid, human, and if appropriate, dramatic. All the while keep in mind and use the readability formula (his) factors - simple words, shorter sentences and more personal references.

Klare (1954), in a similar fashion, has outlined seven steps to be used in writing: 1) get as clear a picture as possible of the reader, 2) write a first draft without regard to the formulas, 3) use a formula to rate the draft, 4) compare the rating to that needed by the readers, 5) re-
vise the draft if necessary, 6) rate the second draft using the formula, and 7) get readers's reactions to the final draft.

Chapter I has discussed why the readability formulas should not be used exclusively in writing. The preceeding literature indicates they do, however, have a place. Klare put it this way.

Our opinion is that writing well is an art, and undoubtedly always will be. On the other hand, we do not believe that science, in its broadest sense, need be incompatible with art. Readability formulas represent the application of scientific knowledge to a very limited aspect of the art. Nevertheless, it is an important part of art, its effectiveness. Because writers usually do not write only for other writers, but for large groups of readers, each must study effectiveness (Klare, 1954, p. 144).

As just stated writing is usually for reading. True reading, according to Bond and Tinker is reading with comprehension.

Basically comprehension depends upon facility in the use of concepts or meaning evolved through experience. To be of use in reading, the concepts acquired through experience must be attached to words or groups of words as symbols of their meaning....Then when a reader recognizes a word or group of words, perception of the printed symbol stimulates the recalling or construction of meanings for which the symbol stands.... (that is)Meaning may be derived directly from past experience or it may consist of a new constructed meaning which results from combining and reorganizing meanings already possessed by the reader. The author brings known ideas together
in such a way that the reader senses a new relationship and therefore gains a new idea, concept, or sensory image (Bond & Tinker, 1967, p. 269).

Reading comprehension, reading authorities agree, is a psychological process (Simons, 1971) which then is scientifically explainable and manipulatable. This opinion of reading would imply a more scientific approach to writing is possible and actually more logical. The problem is that there is no complete theory of reading comprehension at the present. There are some eight approaches being used presently, none of which is adequate. The 'readability' approach is one of these (Simons, 1971). It involves the search for those characteristics of written material that are related to reading comprehension. However, the limitations of this approach have already been pointed out in Chapter I. Another of these approaches is the 'models' approach. In this approach, theoretical models of the reading comprehension process are constructed and from them testable hypotheses are generated (Simons, 1971). One such model is that of David Ausubel. This model is more from the field of educational psychology than reading psychology specifically. However, as stated in Chapter I, it appeared to the writer to offer a viable approach to reading comprehension. Therefore this theory and its implications for written communications will be recounted below.
Ausubel's Theory of Meaningful Verbal Learning

It was previously mentioned that some psychologists distinguish between concept learning and verbal learning. Ausubel is one of these people. However, if one compares his definition of meaningful verbal reception learning with the definition of concept learning used in this study, this distinction disappears. Ausubel (1963) defines reception learning as a situation in which the content is presented to rather than discovered by the learner. The learner must only comprehend the material in such a way that it is functionally reproducible for future use. The 'content' Ausubel refers to includes such things as ideas, concepts, facts, values, and skills. According to Ausubel, in order for content learned by reception to be retained over a long period of time as an organized body of knowledge, it must have been meaningfully learned. A meaningful learning process occurs when some potential meaning inherent in the external world becomes converted into an individualized psychological state. Meaning implies some form of representational equivalence between language and mental content. A symbol, spoken or written, has meaning if it evokes the same concrete image as the object it signifies. Thus it would appear to this writer that Ausubel thinks concepts, as defined in this study, can be learned in a verbal learning situation if that situation is meaningful. And, although Ausubel doesn't make any distinctions about differences in
the meaningful verbal learning process through the modalities of sight or hearing, this writer assumes successful reading comprehension is equivalent to the meaningful verbal learning of particular content. It should be noted that much of Ausubel's work with this theory has involved only the modality of sight (Ausubel, 1960; Ausubel & Fitzgerald, 1961; & Ausubel & Fitzgerald, 1962).

Ausubel's model for a theory of comprehension is called the subsumption process. This model assumes the existence of a cognitive structure that is hierarchically organized in terms of highly inclusive conceptual traces under which are subsumed traces of less inclusive subconcepts as well as traces of specific information. The major organizational principle is that of progressive differentiation of trace systems of a given body of knowledge from regions of greater to lesser inclusiveness, each linked to the next higher step in the hierarchy through the process of subsumption. Meaningful reception learning occurs as potentially meaningful material enters a cognitive field and interacts with and is appropriately subsumed under a relevant and more inclusive conceptual system. There are two kinds of subsumption - derivative and correlative. Derivative subsumption occurs if learning material constitutes a specific example of an established concept. Correlative subsumption occurs if material is an extension, elaboration, or qualification of previously learned material. In the initial stage of sub-
sumption, learning and retention are facilitated: Meaning is acquired and made available, and availability of meaning is maintained. The second stage of subsumption, the oblit­erative stage, occurs when a conceptualizing trend in a cog­nitive structure becomes so strong that it becomes less burdensome to retain a single inclusive concept rather than a larger number of specific facts. Thus forgetting of the specific facts occurs.

According to Ausubel, a person's cognitive structure in a particular subject matter field at any given time is regarded as the major factor influencing learning and retention of meaningful new material. There are three variables of cognitive structure which facilitate maximum meaningful reception learning: 1) availability in the cognitive structure of relevant subsuming concepts at an appropriate level of inclusiveness to provide maximum anchorage, 2) discrim­inability of new learning material from the conceptual sys­tem which subsumes it, and 3) stability and clarity of those subsumers. If a cognitive structure does not have these three characteristics, learning and retention are inhibited.

Ausubel feels there is already support for his con­tention that the availability of relevant subsumers facili­tates learning. He has cited many studies which he feels indicate this, some of which deal with spontaneous ante­cedent organizers and some of which deal with mediational organizers. One example is the study of Paulton's. He
showed that memory of short meaningful statements varies directly with the subjects' degree of background in a particular area. The more background a subject had, the more successful he was in learning. In another example, Brousfield, Cohen, and Whitemarsh found that associative clustering in the recall of words is greater if the words themselves are relatively familiar. In a third example, Underwood and Richardson found that concepts are more easily acquired if the specific instances from which they are abstracted are frequently rather than rarely associated with their defining attributes, and if subjects have more rather than less relevant information about the nature of the attribute (Ausubel, 1963).

Ausubel found a similar result in two of his own studies mentioned previously. In one (Ausubel & Fitzgerald, 1961), undergraduate students who scored above the median on a Christianity test did significantly better in learning and retaining new information on Buddhism than students who scored below the median. In the other (Ausubel and Fitzgerald, 1962), undergraduate students who scored in the upper third of the range of scores on a test in endocrinology did significantly better in learning and retaining new materials on endocrinology than students who scored in the middle or lower third of the range. Actually, Ausubel feels that in the field of education, one can't always count on relevant subsumers being available. Therefore he has
advocated the use of organizers built right into expository materials. Since the organizers are placed before the main body of material to be learned, he calls them advance organizers. His work with such advance organizers has also given support to his theory. In one study (Ausubel, 1960), undergrads who were given advance organizers did significantly better in learning and retaining unfamiliar material on metallurgy than undergrads who were not given the advance organizers. In another study (Ausubel & Fitzgerald, 1961), an advance organizer also facilitated learning of new material on Buddhism, but only for subjects who scored below the median on the Christianity test. A third study (Ausubel & Fitzgerald, 1962) showed advance organizers significantly facilitated initial learning and retention on the subject of endocrinology, but this time only for those subjects who scored in the lower third of the range of scores on a test of verbal ability (SCAT).

Ausubel also feels the role of discriminability, stability, and clarity of subsumers is somewhat supported by research. He has cited the work of Calletine and Warren; Hull; and Shore and Sechrest as indicating that the defining attributes of a concept are learned more readily when the concept is encountered in large numbers of different contexts. That is, when their subsumers are more clear and stable, the attributes are anchored easier and retained longer. He also feels his own work (Ausubel & Fitzgerald,
1961 and Ausubel and Fitzgerald, 1962) mentioned previously supports this. Students who scored higher on tests of background knowledge must have had subsumers which were clearer, more stable and more discriminating.

Ausubel feels that by strengthening these above three described aspects of cognitive structure, learning and retention can be facilitated. Therefore, Ausubel's ideas about how to acquire adequate cognitive structure, based on his theory of subsumption, are guidelines for effective communication of concepts. Basically, Ausubel has suggested two things: 1) organize material around big concepts and principles that have the widest inclusiveness, explanatory power, generalizability, and relatibility to the subject matter content of the discipline and 2) use programmatic methods of presenting and ordering the sequence of subject matter that best enhances the clarity, stability, and discriminability of cognitive structures. This latter suggestion is done by organizing material so that big concepts, maybe in the form of advance organizers, come first, then smaller and smaller ones. It is also done by writing so that relationships can be explicitly explored, similarities and difference pointed out, and real and apparent differences reconciled. It can also be done by giving many examples or by using multicontextual exposures (Ausubel, 1963).

Summary

This section has presented two approaches to the
written communication of concepts - a literary approach and a scientific approach. Writers and authorities of children's literature such as Adler, Huck, Dale, and Chall, maintain that effective communication through the written word is partially an art and partially a science. They would have a person write first as a art and then modify the writing through the use of readability formulas. Experts in the field of reading believe reading comprehension is a scientifically explainable process and therefore effective writing is also. However, as Simons has pointed out, there is no one complete, adequate theory of reading comprehension. Ausubel's theory of meaningful verbal learning was presented as one of these partial theories. His model for reading comprehension is called the subsumption process. Ausubel feels, based on research, that there are three factors affecting this subsumption process: availability of subsumers in a learner's cognitive structure, discriminability of new material from the system that subsumes it, and the stability and clarity of the subsumers. To increase the effectiveness of written communication of concepts, Ausubel has made several suggestions, one of which is to give many examples in writing about concepts.

These two approaches helped to shape the treatment passages used in this study. From the literary approach, many suggestions were used. Prior to the actual writing, library research was done to assure that the facts to be
used in the writing were true. Also, an interest inventory was administered in an attempt to identify science topics of interest to fifth graders. In the actual writing, a simple, direct approach was attempted, anthropomorphism was omitted, and an attempt to involve the reader through the use of the phrases "you may have seen" and "if you were to" was made. After writing, the passages were subjected to the Dale-Chall readability formula and revised several times in order to get the material down to the fourth grade reading level.

From the scientific approach came the independent variable of the study - amplification. Amplification, as explained in Chapter I is achieved through the use of exemplars. Ausubel's theory of meaningful verbal learning gives a theoretical basis for the use of exemplars in the written communication of concepts. He indicated that by giving many examples, the clarity and stability of subsumers would be enhanced, thereby facilitating meaningful verbal learning. Hence came the hypothesis that the more exemplars one gives in a written passage, the better the reader should comprehend it.

**Literature Related To The Cloze Procedure As A Method Of Measuring Reading Comprehension**

**Overview**

When Wilson Taylor introduced the cloze procedure, he designated it "a new psychological tool for measuring
effectiveness of communication" (Taylor, 1953, p. 415). The heart of this procedure is the cloze unit. The cloze unit is

...any single occurrence of a successful attempt to reproduce accurately a part deleted from a message (any language product) by deciding from the context that remains, what the missing part should be (Taylor, 1953, p. 416).

The cloze procedure is defined as

...a method of intercepting a message from a transmitter (writer or speaker) mutilating its language patterns by deleting parts, and so administering it to receivers (readers or listeners) that their attempts to make patterns whole again potentially yield a considerable number of cloze units (Taylor, 1953, p. 416).

The word cloze is derived from the Gestalt psychology term closure which applies to the human tendency to complete a familiar, but not quite finished pattern by mentally closing the gaps. Taylor argues that this closure principle also applies to language.

Given 'chickens cackle and __________ quack,' almost anyone can instantly supply 'ducks.' If that word is really the same as the one omitted, the person scores one cloze unit for correctly closing the gap in the language pattern. Note that the sentence pattern is a complex one made up of many sub patterns. One must know not only the meaning (i.e. patterns of symbol meaning relationships) and forms (patterns of letters) of all the five words, but also the meanings of given combinations of them - plus the fact that the sentence structure seems to demand a term parallel to 'cackle' but associated with ducks instead of chickens.
In other words, one must guess what the mutilated sentence means as a whole, then complete its pattern to fit that whole meaning (Taylor, 1953, p. 416).

Taylor acknowledges that the main contributions to the notion of cloze procedure come from the ideas of "total language context" as discussed by George Miller; Charles Osgood's dispositional mechanisms; and the notions of statistical random sampling. In the total language context concept, any language behavior depends on total context. The total context of any language includes everything that tends to motivate, guide, or hinder that behavior. It includes verbal factors such as grammatical skills and knowledge of symbols and non verbal factors such as fears, desires, past experience, and intelligence. Taylor states that for fifty years, research has indicated this to be true (1953, p. 418). Osgood's notion of dispositional mechanisms is that out of each person's unique experiences, his own set of language habits is developed. These habits reflect the redundancies and transitional probabilities of the language patterns this person's verbal skills involve. Redundancy has to do with the number of words used to express a concept. The more words, the more redundancy. "A man is coming this way" is more redundant than "man coming." Transitional probabilities have to do with the likelihood of some words appearing in certain patterns more than in others. "Merry Christmas" is a more probable pattern than "merry
birthday." To the extent that the person's language habit set corresponds to the sets of others in his culture, he can communicate easily. Statistical random sampling, as applied to the cloze procedure, suggests that if enough words in a passage are struck out at random, the blanks will come to represent proportionally all the kinds of words that occur.

When Taylor (1953) first developed the cloze procedure, he looked on it as a new approach to the measurement of readability. Since that time, it has also been suggested and studied as a technique for measuring comprehension and for teaching reading skills (Jongsma, 1970). This section of Chapter II will only recount the literature related to cloze as a technique for measuring comprehension. This literature was reviewed to provide a basis for justifying the choice of cloze tests as the main measuring instrument in this study.

**Cloze as a Measure of Reading Comprehension**

Several studies have examined the validity of the cloze procedure by correlating the cloze test scores with other measures of comprehension - usually multiple choice questions - over the same material. This pertains to specific comprehension. Other studies have examined the validity of the cloze procedure by correlating cloze scores with standardized reading test scores. This pertains to general comprehension. Since this study dealt with specific comprehension, only the studies dealing with it will be recounted.
The first was done by Taylor (1957). He undertook the study after reasoning that if a passage is readable, it is also understandable; and therefore, cloze scores which measure readability should also measure comprehension. Using a 3,240 word technical article describing the Air Force system of supply, he developed parallel forms of a comprehension test, validated and made them reliable. He also developed eight cloze tests representing a twenty percent sample of the article by systemically deleting any ten words from each passage. He gave 152 Air Force trainees the cloze tests followed by one form of the comprehension test. Then the subjects were given unmutilated copies of the article to read and study. Afterwards, the second form of the comprehension test was given followed by the cloze test. Correlations between the cloze test scores and the comprehension test scores ranged from .70 to .80. These correlations were all significant at the .001 level. Taylor concluded that the cloze tests used in this study were valid indicators of reading comprehension.

The next study on specific comprehension was that of John Bormouth's (1962). In this intensive study, Bormouth used three 250 word passages dealing with literature, social studies, and science. Each passage was rewritten three times to correspond to the 4.5, 5.5, and 6.5 grade levels, as measured by the Dale-Chall readability formula. Cloze tests were developed over each of these passages by deleting
every fifth word. Also nine multiple choice tests based on the passages were developed. The test items included seven major comprehension skills - vocabulary, facts, sequence, relationships, main ideas, inference, and author's purpose. The cloze tests were given to 150 students in grades four, five, and six. Then unmutilated passages were given to the students followed by the comprehension tests. All the scores on the nine cloze tests correlated positively and most at the .001 level of significance with each of the seven comprehension skills as measured by the comprehension test. These results led to the conclusion that cloze test scores tend to measure the same comprehension abilities as those measured by a multiple choice reading comprehension test. Seven years later, Bormouth (1969) replicated this study and got similar results - high correlations between cloze test scores and multiple choice comprehension tests on the same passages.

Ransom (1968) conducted a study comparing the cloze procedure with an informal reading inventory for students in grades one through six. Both the cloze tests and the informal reading inventory were based on basal readers. The correlations between independent, instructional and frustration reading levels of both tests were all significant at the .01 level except on the first grade results. Again the results indicated the cloze procedure is a valid measure of reading comprehension, in this case as measured by an in-
formal reading inventory.

The four studies mentioned above dealt with a concurrent measure of construct validity. This measure of construct validity is, of course, dependent upon the validity of the instrument to which the instrument in question is being correlated. Taylor and Bormouth did report satisfactory validities on the comprehension tests they used. However, any reading comprehension test's validity is still theoretically suspect because of a lack of adequate theory, as previously mentioned. Bormouth, in his 1969 study, went a step farther and established another measure of construct validity of that particular cloze test. He conducted a factor analysis and found that one factor accounted for most of the variance. He called that factor "reading comprehension ability." Weaver and Kingston (1963) also dealt with factor analysis of the cloze procedure. In their study three factors were identified: 1) a verbal comprehension factor, 2) a cloze factor, and 3) a rote memory, flexible retrieval factor. Since the cloze tests loaded higher on the cloze factor, the researchers concluded cloze tests are more related to each other than to the other factors in the study. It should be noted that there were great procedural difference between this study and the Bormouth study, and therefore perhaps the two should not be compared. However Jongsma in discussing these two studies concluded
One can see the evidence is conflicting. There is no conclusive research on the construct validity of the cloze procedure. The fact is that the processes one must go through in completing a cloze test are relatively unknown. If one accepts the high positive relationships between cloze tests and tests of reading comprehension, then perhaps the identification of the processes underlying cloze is closely tied with the processes of comprehension itself (Jongsma, 1970, p. 14).

Kretschmer (1972) argues that another objection to the cloze procedure is that it is at best a measure of literal comprehension. That is because the subject's task is to complete the sentence, not to analyze a whole passage for understanding. This writer disagrees with this position pointing to the theory by Miller previously mentioned. Miller (Taylor, 1953) feels research shows any language behavior is influenced by total language context. This would imply to this writer that a person would go beyond the sentence level in completing a cloze test if the cloze test included more than one sentence. He would utilize the whole context, not just part.

At any rate, the question of cloze validity leaves the writer with the circular argument: "Measures should be based on theory, but psychological theory should be based on empirical observations that rely, in turn on accurate measures" (Kretschmer, 1972, p. 15). However, in view of the fact that cloze tests do seem to be measuring the same things as comprehension tests, it appears one can qualifyingly say cloze tests are valid measures of comprehension.
As far as reliability goes, the previously mentioned studies which dealt with validating the cloze procedure as a measure of comprehension, also indicated the cloze tests used in those studies were adequately reliable. The reliability in the Taylor (1957) study was .88, a test-retest coefficient. In the Bormouth (1962) study, a reliability of .83 was found, an internal consistency coefficient.

**Cloze Procedure Methodology**

As Jongsma (1970) has pointed out, the cloze procedure is somewhat misleading, because cloze tests construction has not always been the same in the research done with it. In the beginning, Taylor (1953) rather arbitrarily chose every nth word as long as there were sixteen deletions in the entire cloze test. This method did prove successful in measuring readability. In his later study (Taylor, 1957), he compared the effects of deleting 'easy' (structural) words, 'hard' (semantic) words, and 'any' words. He found all three systems discriminated equally well in measuring comprehension. Greene (1965) obtained a similar result comparing the deletion of every nth word with the deletion of only 'context' words. Therefore this writer, as well as Weber (1971), concluded that it is most efficient to use the easiest and least time consuming method which is the 'any word' method. Rankin (1958) also considered the methodology question. He hypothesized that different kinds of deletion systems measured different kinds of comprehension. The 'any
word' system samples words influencing structural meaning. Structural meaning is that meaning gained by noting morphological and syntactical clues apart from particular vocabulary words, that is the interrelationships between ideas. The noun-verb deletion system samples words which influence lexical meaning. This is the meaning of individual words. Rankin correlated the cloze scores from these two types of tests with two subtests of the Diagnostic Reading Test which measured what he considered lexical and structural comprehension. The correlations were positive and significant at the .05 level leaving Rankin to conclude that his hypotheses were correct. Structural meaning as defined here appears to be similar to the type of comprehension involved in this study.

The use of the every fifth word deletion system seems to be the most commonly used system in research (Jongsma, 1970). It was one of those first tested and found successful by Taylor (1953). Research by MacGintie (1961) also supports the every fifth word system. He tested fifteen different deletion systems and found that more than a five word context did not help in filling in missing words, but less than four words hindered the process.

Taylor was also first in suggesting how many deletions should be included in a cloze test. As previously mentioned, he first suggested sixteen. Later (Taylor, 1957) he revised it to fifty. He felt this latter figure insured a more
representative sample. Since that time most of the passages used in studies involving children have been of approximately 250 words (Potter, 1968). However, it was not until Bormouth's (1964) study indicated that cloze tests of less than fifty items tend to be unreliable that there was some empirical basis for that particular aspect of cloze methodology.

According to Potter (1968), the most widely researched aspect of cloze methodology has been that of scoring. The results of this research consistently indicates counting only exact replacements of the deleted word as correct is the most practical and objective method of scoring. Two examples follow. The first such study was done by Taylor (1953). He compared counting only exact words with counting exact words and synonyms. Naturally the latter method resulted in higher scores, but the proportion of the total score associated with each passage remained the same. Therefore Taylor concluded that the time consuming and somewhat unreliable practice of counting synonyms was unprofitable. Ruddell (1963) decided to repeat the above study but using cloze scores of elementary children instead of adults. He found the same results: the scoring of synonyms in cloze tests did not significantly add to their discriminatory power.

According to Jongsmma (1970) very little research has been done on the aspects of cloze test administration. What
has generally been done is to give the subjects cloze tests based on passages they have never seen. One exception to this is the Taylor study (1957) mentioned previously, in which he had pre and post cloze tests. However, the correlations between the pre-cloze test and post-cloze and post-comprehension tests were so high that this writer and well as Weber (1971) concluded it is more practical just to give only the pre-cloze test.

Summary

This section has discussed the cloze procedure as a method of measuring specific reading comprehension. The studies of Taylor, Bormouth, and Ransom examined cloze validity. These studies indicated that while cloze tests have correlated highly with comparable comprehension tests, factor analysis results have been conflicting. Studies which examined cloze procedures were also discussed. Results of these indicate that the most efficient word deletion method is the any-word method. McGintie tested several any-word systems and found the every fifth word system to be the best. A study by Bormouth indicated that cloze tests must be of at least fifty items to be satisfactorily reliable. In studies examining cloze scoring procedures, the most practical and objective method was found to be that of counting only exact replacements. Little research has been done on cloze test administration, but one study indicated it is most practical just to give subjects cloze tests.
which they have never seen.

In view of the above studies, the writer concluded that cloze tests were qualifyingly valid measures of reading comprehension. It was also decided to use the every fifth word deletion system, to score only exact replacements, and to give the subjects cloze tests which they had never seen before. And, since some of the cloze tests used in the study were of less than fifty items, it was decided to do reliability checks on them.
CHAPTER III

PROCEDURE

The study design, population and sample, instrumentation, treatment, data collection techniques and data analysis techniques are presented in this chapter. Following is a brief chronological overview of the procedures.

In March, 1972, a pilot study was carried out at Evening Street Elementary School, Worthington, Ohio, using the fourth and fifth grades. The purpose of the pilot study was to try out the study procedures and to establish and reliability and validity of the study instruments. Results of the pilot study indicated that further study under more controlled conditions was warranted (see Appendix B for complete details of the pilot study).

In September, 1972, the writer secured the permission of Columbus Public Schools to carry out the study. Schools with no Aid-to-Dependent-Children were located; and twelve of these schools agreed to cooperate with the writer. In these schools there were 135 children having the required characteristics: age between ten and eleven years and a reading grade level between fourth and fifth. Beginning
in October, 1972, the interest inventories were administered. This took two weeks. The subjects were blocked on interest and randomly assigned to one of the three treatment groups - low, moderate, or high amplification. The treatments were administered in six weeks. The reading comprehension tests, cloze tests, were administered simultaneously with the treatments. In the final two weeks of data collection, the interviews to determine the extent of concrete props in the cognitive structures of the subjects were administered.

Following the collection, the cloze test data was subjected to a two-way analysis of variance. In addition, the cloze test scores and the interview ratings were correlated using Pearson Product Moment Correlation Coefficients.

Pilot Study

As part of the piloting procedures, the researcher informally assessed fifth grade children's science interest at four elementary schools. The assessment was done by administering a modified version of Richardson's Science Interest Inventory for Upper Elementary School Children (Richardson, 1971). This procedure gave the researcher some basis for choosing the three concepts to be used in the treatment passages of the study (see Appendix A for a copy of the modified version of Richardson's Inventory).

The formal pilot was conducted at Evening Street Elementary School, Worthington, Ohio in March and April, 1972.
First, the population was given the interest inventory. From these scores, reliability and validity were determined. Then 36 members of the population, six from each class, were blocked on two levels of interest, and both groups received the three treatment-tests for the three concepts. A two-way analysis of variance was computed to determine if there was a significant difference in the treatment groups, if there was a significant difference in the interest groups, and if any significant interaction occurred. Scheffe's Method was used to determine the locations of significances. Also from the cloze scores for these passages, a Kuder Richardson #20 was computed to determine the reliabilities of the cloze tests. Also from the pilot, the researcher determined (1) if the tests and treatments could be easily conducted within the time limits and (2) if the standard instructions were adequate for the situation.

As a corollary to the pilot study, the researcher informally tested part of her theory: those subjects who have more concrete props associated with exemplars in their cognitive structure will score higher on the cloze test than will those who have fewer concrete props associated with exemplars within their cognitive structure. This was done in the following manner. Eight members of the defined population from each class were given an objective test to determine the extent to which the exemplars treated in the passages were already in the cognitive structure of each sub-
ject as concrete props. The mean score for each class on each subtest of the objective test was used with the mean cloze test score for each concept in each class to make a two by two table. Fisher's Exact Probability test was used to determine the probability that the results occurred by chance (see Appendix A for a copy of the Objective Test).

A second study to test the above stated theory statement was also conducted. This was done in the following manner. The 36 members of the defined population who had taken the cloze tests were interviewed by the researcher to determine the extent to which concrete-empirical props were in the cognitive structure of each subject. The interview rating for each subject on each concept was correlated with the cloze test scores for each concept for each subject. Pearson Product Moment Correlation Coefficients were computed to determine the relationship between the two sets of scores (see Appendix A for a copy of the interview instrument).

As a result of the pilot studies, the following decisions about the study procedures were made.

1. The concepts chosen for the cloze passages were of interest to the subjects for which it was intended.

2. The interest inventory was suitably reliable, valid, and readable.

3. The directions for the interest inventory and the cloze tests should be changed slightly to make them more
clear.

4. All the cloze tests were suitably reliable except for one, and it should be appropriately modified.

5. The procedure for administering the cloze tests should be changed from one to three sittings.

6. The n for each cell should remain at 15.

7. All the concept passages should remain.

8. A third blocking level for interest should be added.

9. A fourth hypothesis - concerning the correlation between comprehension and the presence of concrete-empirical props - should be added.

10. The reading ability level range should be changed from 4.0 through 6.0 grade levels to 4.0 through 5.0 grade levels.

(See Appendix B for a complete report of the pilot study.)

Population and Sample

The population for the problem and subproblems was fifth grade elementary children in Columbus Public Schools, Columbus, Ohio, who were between the ages ten years, zero months and eleven years, zero months, whose reading ability scores according to the California Comprehensive Reading Test given October, 1971, fell within the range 4.0 to 5.0 grade levels, and who were from approximately the same socio-economic background. A sample of 135 was obtained in the following manner. The researcher selected twelve
schools in the far east and far north ends of Columbus, Ohio from the 22 which were listed in the Columbus School Profile (Merriman, 1971) as having no children from families receiving Aid-to-Dependent-Children funds. Then all the children in each of these schools with the appropriate age and reading ability were included in the sample.

This sample of 135 subjects took an interest inventory to determine their interest in each of the elementary science concepts used as the topics for the written passages in the study. The sample was then blocked on interest. This was done by assigning the subjects with the upper third of scores in the interest inventory to one group, the subjects with the middle third of the scores on the interest inventory to a second group, and the subjects with the lower third of scores on the interest inventory to a third group. Then each subject of these three groups was randomly assigned to one of the three treatment groups - low, moderate, and high amplification. There were three of these elementary science concepts used as topics for the written passages in the study. The sample took an interest inventory on each of the three concepts. Each subject was blocked and randomly assigned to a treatment group three times - one for each of the three written passages. Then, each subject was randomly assigned to one of six passage orders: R(sedimentary rocks) T(deciduous trees) I(insects), RIT, TRI, TIR, IRT, and ITR. This determined the order in which the
subjects received their three treatment passages.

In the interview phase of the study, all 135 subjects received the same interview. Within each class participating in the study, the subjects were randomly assigned to the order in which they were interviewed.

Instrumentation

**Interest Inventory**

An interest inventory to measure interest toward the three elementary science concepts was developed by the researcher. To do this, parts of models set up by Heimbeiger (1970) and Jenkins (1969) were used. In Heimbeiger's model, the subjects were presented with a series of two statements and asked to select the statements which indicated how they felt. One statement expressed a negative attitude, and the other expressed a positive attitude. The researcher's interest inventory was similar, but the positive and negative statements of the groups were randomly ordered. By having this arrangement, the subjects were unable to mark one or the other of the positive and negative statements in a group. Instead they circled the number of statements which were true for them and left blank those which were false for them. Under perfect testing conditions, this would mean a negative statement which was left blank was an indicator of a positive attitude. In this study, the writer had no way of knowing if this were true; it is possible some items
were overlooked or skipped instead of being purposefully marked blank. However, since there were no indications to the contrary, the writer assumed they were purposefully "marked" blank and therefore counted blank negative statements as indicators of positive attitudes. The number of statements positively related to the concepts described in the passages which were chosen was a measure of the interest of the subject concerning the concept. In Jenkin's model, each interest inventory item was indirectly related to the treatment. This researcher followed the same procedure: the two statements in each group were indirectly related to the concepts described in the treatment passages. An example follows:

Concept - Small animals with six legs are called insects.

1. I like to look at bugs.
2. I don't like to look at bugs.

The inventory consisted of 72 items - 24 for each concept. Reliabilities of the three subtests, as determined by the Kuder Richardson Formula 20, were .928, .874, and .817. Validity of the instrument subtests was determined in three ways. An Item Analysis showed the Point Biserial Correlation Coefficients of all the items to be significant at the .05 level or better. These are measures of the relationship of information given by each item to the information reflected by the total score. Secondly, factor analysis showed the three subtests to consist of relevant factors. The main
ones are (1) interest-in-rocks subtest -- interest in "artistic" aspects of rocks, interest in "intriguing" aspects of rocks, and "general rock hound" interests; (2) interest-in-insects subtest -- interest in "non-goal oriented activities associated with insects," interest in "goal-oriented activities associated with insects," and interest in "activities with insects which do not involve direct contact;" (3) interest-in-trees subtest -- interest in "large numbers of trees," interest in "collecting and doing activities with leaves," and interest in "being outdoors." Thirdly, the writer specifically worked at developing an inventory which measured interest in rocks, trees, and insects. Therefore, from a subjective, but logical point of view, there can be confidence in the content validity of the instrument. (See Appendix A for a copy of the inventory and Appendix B for a detailed explanation of the instrument development.)

Cloze Tests

To measure reading comprehension, cloze tests on the treatments were constructed by deleting every fifth word and putting a standard sized blank of twelve typewriter spaces in the place of the deleted word. As noted in Chapter II, research suggests that cloze tests are valid measures of reading comprehension. Research also suggests cloze tests are most reliable with at least 50 items (Potter, 1968). However, the treatment passages for this study ranged from 150 to 350 words because the researcher felt this was most
suitable for the attention span of fifth graders. With a deletion system of every fifth word, the resulting cloze tests were from 30 to 70 items in length. Therefore, the Kuder Richardson Formula 20 was used in the pilot study to determine the reliabilities of the tests. The coefficients were found to be .69, .88, .37, .67, .90, .86, .76, .73, and .92. The cloze test with the reliability of .37 was examined, and appropriate changes were made to increase the reliability (see Appendix B for a more detailed explanation and Appendix A for a copy of the cloze tests).

**Personal Interview**

To determine the extent of the presence of concrete-empirical props, which are associated with the three concepts in the study, in the cognitive structures of the readers, a standardized personal interview instrument was developed. This instrument used somewhat open-end items which enabled the interviewer to determine two things: (1) could the students distinguish the concept example under question from other concept examples and (2) did the students know the concept example's characteristic which enabled it to be classified as the concept in question. There were fifteen basic questions, one for each concept exemplar, in the form "Can you describe a ________?" Besides these basic questions, there were five to six prompting questions for each concept.

This instrument lacks any measured reliability and
validity. However, several things do contribute to its reliability and validity. The writer was the sole interviewer and strove to be consistent in mannerisms and attitudes toward the children throughout all the interviews. Prior to the study, the researcher had the training experience of a pilot study. The interview instrument was developed by the writer for the above mentioned specific purposes and therefore has content validity. Because it is a standardized instrument, instrument decay is controlled. Also, the researcher's experience and ability as an elementary teacher enabled her to establish rapport with the subjects and thereby obtain the desired information (see Appendix A for a copy of the interview instrument).

**Treatment**

The treatment consisted of three levels of amplification of three different passages describing elementary science real classificational concepts. The choice of the particular real classificational concepts was based somewhat on the researcher's (a former elementary teacher) knowledge of elementary children's interests and on an informal assessment of Columbus Public School fifth grade children's science interest. That assessment indicated no great differences in interest in six major science areas - physical science, chemical science, earth science, space science, botanical science, and zoological science (see Appendices
A and B for more details). The concepts which were chosen are:

1. Small animals with six legs are called insects.
2. Trees which lose their leaves once a year are called deciduous.
3. Rocks which are formed from little pieces of other rocks are called sedimentary.

The levels of amplification - low, moderate, and high - have been previously described as having one, three, and five exemplars. The one exemplar passages consisted of approximately 150 words, the three exemplar passages of 250 words, and the five exemplar passages of 350 words. The number of exemplars used were chosen because the researcher felt five in 350 words was at the limit of the attention span of an average fifth grader. The exemplars of each concept were chosen by the researcher. The basis of decision was the researcher's informal knowledge of the frequency of occurrence of the exemplars in the Columbus, Ohio, area and on the researcher's informal knowledge of the probability of fifth grade children in Columbus encountering the exemplars. Exemplars, whose chances of being encountered by fifth grade children in Columbus, Ohio, were considered great, were chosen. The exemplars used were:

1. sedimentary rocks - sandstone, limestone, shale, conglomerate, Grand Canyon
2. deciduous trees - maple, redbud, willow, buckeye,
sassafras

3. insects - butterfly, roach, ant, grasshopper, walking stick

All passages had the same readability level of fourth grade as measured by the Dale Chall readability formula. They also had the same potential conceptual difficulty as measured by digraph analysis - first level inference.

Design

The main study was a three by three treatment by levels design which was repeated with three different elementary science concepts. The general data matrix for the study is shown in Fig. 5.

Independent variable 1: amplification of written passages describing science concepts

B₁ - low amplification: one exemplar
B₂ - moderate amplification: three exemplars
B₃ - high amplification: five exemplars

Independent variable 2: interest of the children toward each of the science concepts

A₁ - low interest: the lower third of scores on the interest inventory
A₂ - moderate interest: the middle third of scores on the interest inventory
A₃ - high interest: the upper third of scores on the interest inventory
Fig. 5. General data matrix for the study.
Dependent variable: reading comprehension
X - cloze test score

This design is an extension of one of the three true experimental designs described and strongly recommended in Stanley and Campbell (1963) – the posttest-only control group design. See Figure 6.

R X
I1
01
R X
I2
02
R X
I3
03

Fig. 6. Extended posttest-only control group design used in the study.

This design controls most sources of invalidity.

Internal Validity

1. Maturation was generally controlled because all three treatments were administered in proximity - within a period of six weeks.

2. General history was controlled in that all three treatment groups experienced the same general historical effects; but unique intrasession history was not, as all the subjects could not be treated and tested at the same time or in the same place. The researcher did, however, control for this somewhat by administering the treatments and tests only in the mornings.

3. Selection bias was controlled, as the subjects were blocked on interest and then randomly assigned to the
treatment groups.

4. Regression and testing presumably did not affect the study, as there was no pretest.

5. Experimenter bias might have affected the validity of the results, as the researcher administered the treatments and the tests. However, since the interest groups were not distributed in an orderly fashion in each testing group, the researcher was not aware of the interest standing of any particular subject at the time of treatment administration. Also, at each treatment session, each subject received a different treatment level with each repeated measure. This further prevented the researcher from behaving differently toward any of the treatment groups. In addition, the tests were scored in the same order as administered to prevent further experimenter bias.

6. The arrangements described in #5 also prevented a "Hawthorne" effect in the subjects from affecting the internal validity of the study.

7. Instrumentation decay was controlled by using standardized instruments with satisfactory reliability.

8. Interference of treatments possibly may have affected the validity, as each subject received three treatments. Ordering effects were controlled by randomly assigning the subjects to the six possible orders of treatments.
External Validity

1. The external validity of the study may have been decreased in that the sample had particular geographical characteristics which members of the target population can not possibly have.

2. Interaction of the characteristic described in #1 and treatment may have affected the validity of the study.

3. External validity should have been increased through the use of repeated measures which were not exact replicas (Campbell & Stanley, 1963).

4. Multiple-treatment interference may have affected the external validity of the study in that there were three treatments administered to each subject.

5. A "Hawthorne" effect which might have affected external validity was controlled somewhat by giving the interest test, treatment and cloze tests to all fifth grade students in each of the schools selected for use in the study.

6. Novelty and disruption effects may have affected the external validity somewhat, as the subjects were not accustomed to many changes in their daily school routine.

7. An experimenter effect may have affected the external validity of the study.

8. Interaction of pre-testing, in the form of the interest inventory, and treatment may have affected this study.
However, this probably was minor because the interest inventory items were only indirectly related to the treatment passages.

9. Interaction of history and treatment may have affected the external validity in that events such as two student deaths, Halloween, Thanksgiving, Fire Prevention Week, and the first snow occurred during the ten week testing period.

For the interview phase of the study, a simple survey design was used. See Fig. 7.

Data Collection

During October, November, and December, 1972, the researcher administered the tests and treatments to the 135 subjects in the sample. Each school was assigned to the times for taking the inventory and test-treatments according to teacher preference. First, the interest inventory was given in the mornings of ten consecutive days. The interest inventories were then scored. For each concept the subjects were blocked and randomly assigned to one of the three treatments. Each subject was previously randomly assigned to one of six concept orders: R(rocks) T(trees) I(insects), R I T, T R I, T I R, I R T, I T R. During the mornings of the next six weeks, the researcher administered the treatment-tests. The treatment-tests were administered every two weeks at the same time of day (morning) for most of the
Fig. 7. Flow chart of the corollary study design.

classes. Seven variations from this schedule did occur due to trips, programs, and holidays. The testing and treatment administration took place mostly in the subjects own classrooms. In two schools, where the subjects were in split grade classes, the fifth grade students took the treatment tests in another room. After the data was collected, the
scores of subjects who did not complete the interest inventory were thrown out. Immediately after the administration of the treatment-tests, the interviewing began and was completed in seven days. The order of interviewing subjects within classes was random. Interviewing was done mostly in the subject's own classroom by the researcher. In two schools, the students were interviewed in the hall and in an empty classroom.

Data Analysis

Hypotheses 1, 2, and 3 were tested with a two-way analysis of variance. This was done for each of the three groups of data - one on the sedimentary rock passage, one on the deciduous tree passage, and one on the insect passage. It is apparent that a three-way design would have been advantageous in this study; it would have eliminated the lack of independence in two groups of data, and it would have allowed an analysis of the variance among the data on the three passages. However, the sample size needed for such a design (405) would have been unrealistic in a logistical sense.

Hypothesis 4 was tested by correlating the cloze test scores and the interview ratings using Pearson Product Moment Correlation Coefficients. The data from each concept passage and from each level of amplification was correlated separately with its corresponding interview rating.
data. This resulted in nine correlations.
CHAPTER IV

RESULTS

This chapter presents the results of the analysis and the discussion of the results. They are presented in four sections. The sections are (1) number of subjects, (2) reliability of cloze tests, (3) analyses of variances of the cloze test scores, and (4) correlations for the cloze test scores and interview ratings.

Number of Subjects

Table 1 shows the final number of subjects out of the original 135 whose interest inventories, cloze test scores, and interview ratings were analyzed. The differences in the n's in the cells from the original n of 15 per cell resulted from normal absenteeism. The resulting differences in cell n's were considered acceptable, and therefore no attempt was made to equalize the n's by random deletion.

Reliability Estimates for the Cloze Tests

The reliability estimates for the nine cloze tests used in the study were determined first with the pilot
Table 1

Number of Students Participating in the Study by Cells

<table>
<thead>
<tr>
<th>Passages</th>
<th>Low Amplification</th>
<th>Moderate Amplification</th>
<th>High Amplification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>low interest</td>
<td>moderate interest</td>
<td>high interest</td>
</tr>
<tr>
<td>Sedimentary rocks</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Deciduous trees</td>
<td>15</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Insects</td>
<td>13</td>
<td>14</td>
<td>10</td>
</tr>
</tbody>
</table>
study data and then again with the formal study data. A comparison of the results is shown in Table 2. The reliability estimates for the low amplification cloze tests were computed using the Kuder Richardson Formula 20 (Hoyt, 1941). The reliability estimates for the moderate and high amplification cloze tests were computed using even-odd items correlations corrected with the Spearman Brown Formula.

The reliability of the first version of the high amplification sedimentary rock passage, .37, was unacceptable. Consequently the writer revised that test. The results of the second estimate indicate the reliability of the second version of the high amplification sedimentary rock passage (.71) is satisfactory. All the other cloze test reliability estimates are also satisfactory. Therefore, the writer concluded the cloze tests used in the study are reliable measuring instruments.

Analyses of Variances of the Cloze Test Scores

A two-way analysis of variance for each of the three sets of data was computed using the MANOVA program. There was one analysis for the data from the sedimentary rock passage, one for the set of data from the deciduous tree passage, and one for the set of data from the insect passage. Where the F ratios were significant, Scheffe S Method was used to determine the location of the significance.
Table 2

Reliability Estimates of Cloze Tests

<table>
<thead>
<tr>
<th>Cloze Tests</th>
<th>Reliability Estimates&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Pilot Study</th>
<th>Formal Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.69</td>
<td>.61&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sedimentary Rocks</td>
<td></td>
<td>.88</td>
<td>.80&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Low Amplification</td>
<td></td>
<td>.37</td>
<td>.71&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Moderate Amplification</td>
<td></td>
<td>.67</td>
<td>.58&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>High Amplification</td>
<td></td>
<td>.90</td>
<td>.85&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Deciduous Trees</td>
<td></td>
<td>.86</td>
<td>.66&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Low Amplification</td>
<td></td>
<td>.76</td>
<td>.62&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Moderate Amplification</td>
<td></td>
<td>.73</td>
<td>.62&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Insects</td>
<td></td>
<td>.92</td>
<td>.88&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Low Amplification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate Amplification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Amplification</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Internal consistency coefficients.

<sup>b</sup>Estimated with Kuder Richardson Formula 20.

<sup>c</sup>Estimated with even-odd item correlations corrected with Spearman Brown Formula.
**Sedimentary Rock Passage**

The means and standard deviations of cloze test scores on the different levels of amplification of the sedimentary rock passage by levels of interest is shown in Table 3. The F statistics for interest, amplification and amplification X interest interaction on the sedimentary rock passage cloze scores is shown in Table 4.
Table 3
Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of the Sedimentary Rocks Passage by Levels of Interest

<table>
<thead>
<tr>
<th>Interest</th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n  mean  SD</td>
<td>n  mean  SD</td>
<td>n  mean  SD</td>
</tr>
<tr>
<td>Low</td>
<td>13 .502 .093</td>
<td>13 .496 .081</td>
<td>10 .404 .086</td>
</tr>
<tr>
<td>Moderate</td>
<td>13 .479 .085</td>
<td>14 .446 .110</td>
<td>11 .487 .060</td>
</tr>
<tr>
<td>High</td>
<td>14 .490 .158</td>
<td>11 .514 .068</td>
<td>11 .432 .078</td>
</tr>
</tbody>
</table>

Note.—Raw cloze scores are converted to ratio of total points possible.
Table 4
Analysis of Variance of Cloze Scores by Amplification for the Sedimentary Rocks Passage and by Interest

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>2</td>
<td>.002</td>
<td>.001</td>
<td>.110</td>
<td>.896</td>
</tr>
<tr>
<td>Amplification</td>
<td>2</td>
<td>.046</td>
<td>.023</td>
<td>2.429</td>
<td>.093</td>
</tr>
<tr>
<td>Interest x Amp</td>
<td>4</td>
<td>.071</td>
<td>.018</td>
<td>1.856</td>
<td>.124</td>
</tr>
<tr>
<td>Error</td>
<td>101</td>
<td>.960</td>
<td>.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>109</td>
<td>1.079</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—For means and standard deviations, see Table 3.

Examination of Table 4 reveals

1. there are no significant differences between the cloze test scores of subjects who read high, moderate, or low amplification versions of the sedimentary rock passage.

2. there are no significant differences between the cloze test scores, on the sedimentary rock passage, of subjects who have received high, moderate, or low interest scores.

3. there is no significant interaction between the amplification levels and the interest levels.
Deciduous Trees Passage

The means and standard deviations of the cloze test scores on different levels of amplification for the deciduous tree passage by levels of interest is shown in Table 5. The F statistics for interest, amplification, and amplification X interest interaction on the deciduous trees passage cloze scores is shown in Table 6.
Table 5

Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification for the Deciduous Trees Passage by Levels of Interest

<table>
<thead>
<tr>
<th>Amplification</th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>mean</td>
<td>SD</td>
</tr>
<tr>
<td>Low</td>
<td>15</td>
<td>0.546</td>
<td>0.110</td>
</tr>
<tr>
<td>Moderate</td>
<td>14</td>
<td>0.559</td>
<td>0.124</td>
</tr>
<tr>
<td>High</td>
<td>13</td>
<td>0.610</td>
<td>0.091</td>
</tr>
</tbody>
</table>

Note.—Raw cloze scores are converted to ratio of total points possible.
### Table 6

Analysis of Variance of Cloze Scores by Amplification for the Deciduous Trees Passage and by Interest

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>2</td>
<td>.013</td>
<td>.007</td>
<td>.636</td>
<td>.531</td>
</tr>
<tr>
<td>Amplification</td>
<td>2</td>
<td>.186</td>
<td>.093</td>
<td>9.088</td>
<td>.001</td>
</tr>
<tr>
<td>Interest X Amp</td>
<td>4</td>
<td>.034</td>
<td>.008</td>
<td>.828</td>
<td>.510</td>
</tr>
<tr>
<td>Error</td>
<td>106</td>
<td>1.083</td>
<td>.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>1.316</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—For means and standard deviations, see Table 5.

Examination of Table 6 reveals

1. there are significant differences at the .001 level between the cloze test scores of subjects who read high, moderate, or low amplification versions of the deciduous trees passage.

2. there are no significant differences between the cloze test scores, on the deciduous trees passages, of subjects who have received high, moderate, or low interest scores.

3. there is no significant interaction between the amplification levels and the interest levels.
Results of past hoc multiple comparison for the amplification factor for the deciduous trees passage are shown in Table 7. Note that $X_1$, $X_2$, and $X_3$ are respectively low, moderate, and high amplification level means.

Examination of Table 7 reveals

1. there is a significant difference at the .05 level between low and moderate amplification.
2. there is a significant difference at the .05 level between high and low amplification.
3. there is no significant difference between moderate and high amplification.

Table 7

<table>
<thead>
<tr>
<th>Post Hoc Multiple Comparisons for the Amplification Factor for the Deciduous Trees Passage</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi_s = .050$</td>
</tr>
<tr>
<td>If $\Phi i$ (pairwise or compound) &gt; .050, reject $H_0: \Phi i = 0$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$\bar{X}_3$</td>
</tr>
<tr>
<td>$\bar{X}_2$</td>
</tr>
<tr>
<td>$\bar{X}_1$</td>
</tr>
</tbody>
</table>

*significant at the .05 level.
**Insect Passage**

The means and standard deviations of the cloze test scores on different levels of amplification of the insect passage by levels of interest is shown in Table 8. The F statistics for interest, amplification, and amplification X interest interaction on the insect passage cloze scores is shown in Table 9.
### Table 8

Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of the Insects Passage by Levels of Interest

<table>
<thead>
<tr>
<th>Amplification</th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>mean</td>
<td>SD</td>
</tr>
<tr>
<td>Low</td>
<td>13</td>
<td>.476</td>
<td>.077</td>
</tr>
<tr>
<td>High</td>
<td>10</td>
<td>.393</td>
<td>.080</td>
</tr>
</tbody>
</table>

Note.--Raw cloze scores are converted to ratio of total points possible.
Table 9

Analysis of Variance of Cloze Scores by Amplification for the Insects Passage by Levels of Interest

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>2</td>
<td>.012</td>
<td>.006</td>
<td>.657</td>
<td>.511</td>
</tr>
<tr>
<td>Amplification</td>
<td>2</td>
<td>.054</td>
<td>.027</td>
<td>3.123</td>
<td>.048</td>
</tr>
<tr>
<td>Interest X Amp</td>
<td>4</td>
<td>.027</td>
<td>.007</td>
<td>.728</td>
<td>.539</td>
</tr>
<tr>
<td>Error</td>
<td>102</td>
<td>.890</td>
<td>.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>112</td>
<td>.983</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—For means and standard deviations, see Table 8.

Examination of Table 9 reveals

1. there are significant differences at the .05 level between the cloze test scores of subjects who read high, moderate, or low amplification versions of the insect passage.

2. there are no significant differences between the cloze test scores, on the insect passage, of subjects who have received high, moderate, or low interest scores.

3. there is no significant interaction between the amplification levels and the interest levels.

Results of post hoc multiple comparisons for the amplification factor are shown in Table 10. Note that $\bar{X}_1$, $\bar{X}_2$, and $\bar{X}_3$ are respectively low, moderate, and high amplification level means.
Examination of Table 10 reveals
1. there is a significant difference at the .05 level between moderate and low amplification.
2. there is no significant difference between moderate and high amplification.
3. there is no significant difference between high and low amplification.

Table 10
Post Hoc Multiple Comparisons for the Amplification Factor on the Insects Passage

\[ \hat{s} = .050 \]

If \( i \) (pairwise or compound) \( \geq .050 \), reject \( H_0: \hat{i} = 0 \)

<table>
<thead>
<tr>
<th></th>
<th>( \bar{x}_1 )</th>
<th>( \bar{x}_3 )</th>
<th>( \bar{x}_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{x}_1 )</td>
<td>.438</td>
<td>.027</td>
<td>.055*</td>
</tr>
<tr>
<td>( \bar{x}_3 )</td>
<td>.465</td>
<td>.028</td>
<td></td>
</tr>
<tr>
<td>( \bar{x}_2 )</td>
<td>.493</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*significant at the .05 level.

Discussion of Results
Analysis of the cloze test data reveals that the only instance in which amplification caused a significant increase in reading comprehension is on the insect passage
between the moderate and low amplification versions.

The analysis revealed significant differences between the high and moderate amplification versions and between high and low amplification versions of the deciduous trees passage. However, the value of the cell means indicate amplification of this passage caused significant decrease in reading comprehension. Analysis of the cloze test data on the sedimentary rocks passage did not reveal significant differences between any of the levels of amplification, but since the probability level was .093, perhaps a trend is indicated. And the values of the cell means indicate the trend is also toward amplification causing a decrease in reading comprehension.

There are several factors which may account for some or all of the inconsistencies in the results on the amplification factor.

1. Subjects who received the high amplification versions were not observed to check over their work as often as did subjects who received moderate and low amplification versions. It seemed the subjects who read low or moderate amplification versions "gave more thought" to their tests than did subjects who read high amplification versions. This activity might have contributed to the result of higher scores on the low and/or moderate amplification versions than on the high amplification versions.
2. The percentages of students having the concrete empirical props associated with the passage exemplars in their cognitive structures is shown in Table 11.

Examination of Table 11 reveals the concept exemplars were not equally known by the subjects. The writer chose the concept exemplars using the criterion that the chances be good of their being relevant concrete props in the cognitive structures of the subjects. However, it appears the chances were not equally good. In fact the variation in chances is quite large. This indicates moderate amplification, using three exemplars, was not three times as great as low amplification, using one exemplar. And high amplification, using five exemplars, was not five times as great as low amplification, using one exemplar. Thus, the treatment of the study lacks validity. It is perhaps significant to note that the percentages on the exemplars in the insect passage were more similar than the percentages on the exemplars in the deciduous trees passage and the sedimentary rocks passage. This fact may help explain why only the analysis of the insect passage cloze scores indicated amplification caused an increase in reading comprehension.

From the above discussion, one could assume that if only the exemplars had been equally well known, the results might have supported the study hypothesis.
Table 11

Percentages of Subjects Having the Concrete Empirical Props Associated with the Passage Exemplars in their Cognitive Structures

<table>
<thead>
<tr>
<th></th>
<th>Deciduous Tree Passage</th>
<th>Sedimentary Rock Passage</th>
<th>Insect Passage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exemplar Percentage</td>
<td>Exemplar Percentage</td>
<td>Exemplar Percentage</td>
</tr>
<tr>
<td><strong>Low Amplification</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(one exemplar)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. maple</td>
<td>51</td>
<td>sandstone</td>
<td>62</td>
</tr>
<tr>
<td>2. redbud</td>
<td>12</td>
<td>shale</td>
<td>18</td>
</tr>
<tr>
<td>3. sassafras</td>
<td>3</td>
<td>Grand Canyon</td>
<td>12</td>
</tr>
<tr>
<td><strong>Moderate Amplification</strong> (three exemplars)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. buckeye</td>
<td>8</td>
<td>conglomerate</td>
<td>14</td>
</tr>
<tr>
<td>5. willow</td>
<td>21</td>
<td>limestone</td>
<td>3</td>
</tr>
<tr>
<td><strong>High Amplification</strong> (five exemplars)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. maple</td>
<td>51</td>
<td>sandstone</td>
<td>62</td>
</tr>
<tr>
<td>2. redbud</td>
<td>12</td>
<td>shale</td>
<td>18</td>
</tr>
<tr>
<td>3. sassafras</td>
<td>3</td>
<td>Grand Canyon</td>
<td>12</td>
</tr>
<tr>
<td>4. buckeye</td>
<td>8</td>
<td>conglomerate</td>
<td>14</td>
</tr>
<tr>
<td>5. willow</td>
<td>21</td>
<td>limestone</td>
<td>3</td>
</tr>
</tbody>
</table>

Roach 18, Ant 67, Grasshopper 35, Walking Stick 22
With the data available, it was possible to create a situation in which the exemplars would be somewhat equally well known. This was done by analyzing only those cloze scores made by Ss who had 67 to 100 percent of the relevant concrete props in their cognitive structures. From this one could not say moderate amplification is three times greater than low amplification, or high amplification is five times greater than low amplification. One can say, however, that low amplification $< \text{moderate amplification} < \text{high amplification}$. The results of the analyses are shown in Tables 12 and 13.

Table 12

Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of Ss Who Had 67 to 100 Percent of the Relevant Concrete Props in their Cognitive Structures

<table>
<thead>
<tr>
<th>Amplification</th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passage</td>
<td>n Mean</td>
<td>n Mean</td>
<td>n Mean</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>SD</td>
<td>SD</td>
</tr>
<tr>
<td>Sedimentary Rocks</td>
<td>28 .510</td>
<td>8 .471</td>
<td>1 .411</td>
</tr>
<tr>
<td></td>
<td>.112</td>
<td>.112</td>
<td>.411</td>
</tr>
<tr>
<td>Deciduous Trees</td>
<td>19 .575</td>
<td>5 .516</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.106</td>
<td>.050</td>
<td></td>
</tr>
<tr>
<td>Insects</td>
<td>7 .465</td>
<td>14 .493</td>
<td>3 .502</td>
</tr>
<tr>
<td></td>
<td>.106</td>
<td>.050</td>
<td>.092</td>
</tr>
</tbody>
</table>

Note.--Raw cloze scores are converted to ratio of total points possible.
Table 13
Analyses of Variances of Cloze Scores by Amplification of the Ss Who Had 67 to 100 Percent of the Relevant Concrete Props in their Cognitive Structures

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedimentary Rocks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplification</td>
<td>2</td>
<td>.017</td>
<td>.009</td>
<td>.681</td>
<td>.513</td>
</tr>
<tr>
<td>Error</td>
<td>34</td>
<td>.426</td>
<td>.013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deciduous Trees</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplification</td>
<td>1</td>
<td>.014</td>
<td>.014</td>
<td>1.473</td>
<td>.238</td>
</tr>
<tr>
<td>Error</td>
<td>22</td>
<td>.212</td>
<td>.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplification</td>
<td>2</td>
<td>.005</td>
<td>.002</td>
<td>.408</td>
<td>.470</td>
</tr>
<tr>
<td>Error</td>
<td>21</td>
<td>.116</td>
<td>.006</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Examination of Table 13 indicates there were no significant differences between any of the levels of amplification on any of the passages. If one overlooks the great differences in cell means, then it would appear that amplification of passages treating elementary science concepts has no effect on the reading comprehension of Ss who read those passages. Therefore the study hypothesis is apparently not supported at
all. However, a study utilizing a procedure similar to the above, but with more equal cell n's, would have to be conducted before such a statement could be made with confidence.

Further examination of Table 11 reveals that the exemplars with the greatest percentages were the ones first presented in the deciduous trees passage and the sedimentary rocks passage. These two exemplars, having percentages of 51 and 62, made up the low amplification versions of their respective passages. It was the low amplification level that had the highest mean score on the deciduous trees passage and on the sedimentary rock passage. The exemplar having the greatest percentage of the insect passage was the third one. This exemplar, having a percentage of 67, was in the moderate amplification version. It was this level that had highest mean score on the insect passage. This fact may also have contributed to the result of higher cloze scores on the low and/or moderate amplification versions than on the high amplification versions.

The analysis of the cloze test data revealed no significant differences between levels of interest. This result is surprising in view of the pilot study results which indicated significant difference due to interest on the sedimentary rock passage and a trend toward this in the insect passage. It also contradicts popular opinion of
reading authorities (Russell, 1961) and the results of the studies of Bernstein (1955), Fader and McNeil (1968), and Frase (1969). The former two studies indicated that interest enhances reading comprehension. The latter indicated that increased intent-to-process information caused increases in memory of information. The analysis of the cloze test data revealed no significant interaction between amplification and interest. This result is of some interest. Although no previous research reviewed by the writer indicated interaction in situation as this might occur, the pilot study results indicated it would. Interaction on the pilot study results did achieve significance on the sedimentary rocks passage and a trend on the deciduous tree passage. These results indicated amplification might be good for students of low interest and bad for students of high interest. This, however, was not strongly supported by the formal study results. One trend (.124) in the sedimentary rocks passage cloze data does suggest the same relationship, but it is not a strong trend.

A comparison of the amplification level means of the cloze scores for the three passages is shown in Table 14. A breakdown by levels of interest was not considered necessary because the interest factor was not significant in the analysis of variance results.
Examination of Table 14 clearly shows, as mentioned and discussed previously, the inconsistencies in the results of the analyses of the cloze test data for the three passages. For the insect passage, cloze scores increased as amplification increased except between moderate and high amplification. For the sedimentary rocks and deciduous trees passages, the cloze scores decreased as amplification increased. Examination of Table 14 also reveals that the cloze scores for the deciduous tree passage were consistently higher than the cloze scores for the sedimentary rocks or deciduous trees passages. No reasons for this observation are apparent to the writer.

Cloze scores below .40 are considered to indicate reading material which is below the instructional level. As
indicated in Table 14, the amplification level means of the cloze scores for all three passages were well above .40. Further examination of individual cloze scores indicated less than nine percent of the Ss scored below .40. The writer does not consider this percentage to be large enough to have confounded the results.

Having three passages in this study instead of just one was meant to increase external validity. However, because the cloze scores of the three passages were so inconsistent, no generalizations about other concepts can be made. Therefore, the use of three passages has served to prevent erroneous generalizing. If only one passage had been used - for example the deciduous tree passage - erroneous generalizations concerning the effect of amplification upon reading comprehension might have been made.

Correlations for the Cloze Test Scores and Interview Ratings

To determine the relationship between the cloze test scores and the interview ratings, Pearson Product Moment Correlation Coefficients were computed using the Biomedical Computer Program, Correlation with Item Deletion (BMD03D). Coefficients for sets of cloze scores on each of the levels of amplification in each passage and the corresponding interview ratings were computed. The results are shown in Table 15.
Table 15

Correlations between Cloze Scores and Interview Ratings

<table>
<thead>
<tr>
<th>Passage</th>
<th>Low Amplification</th>
<th>Moderate Amplification</th>
<th>High Amplification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedimentary Rock</td>
<td>0.258</td>
<td>0.073</td>
<td>0.001</td>
</tr>
<tr>
<td>Deciduous Tree</td>
<td>0.063</td>
<td>0.115</td>
<td>0.173</td>
</tr>
<tr>
<td>Insect</td>
<td>0.122</td>
<td>0.086</td>
<td>0.161</td>
</tr>
</tbody>
</table>

Examination of Table 15 reveals there are positive correlations between the cloze test scores on the insect, deciduous trees, and sedimentary rocks passages and the interview ratings, but none of these correlations are significant. This finding is of interest, as it does not allow
any support for Ausubel's theory that children can meaningfully learn written material if they have the necessary concrete props. The results of this study seem to indicate the presence of concrete props in the cognitive structures of upper elementary children is not positively associated with reading comprehension. This is the same result found with the pilot study data.

In order to insure that the above correlation was correct, the writer obtained scatter diagrams for each of the nine sets of data using the Biomedical Program, Correlation with Transgeneration (BMD02D). Inspection of these diagrams revealed no curvilinear relationships. Therefore, the assumption of linearity of regression is fairly well satisfied.

However, one factor which probably confounded the results is lack of instrument validity. The interview instrument did not distinguish between exemplars which were actually seen, exemplars which were read about and exemplars which were just heard about. Obviously, only the first type could be considered concrete props. During the course of interviewing, the writer became aware of this distinction and asked many children, "How do you know?" The answers did fall into all three of the above categories. However, the writer does not know the proportions of the answers falling into these categories as no records of those answers were made.
Chapter IV presented the results of the analysis and the discussion of the results.

The final number of subjects who participated in the study was such that unequal cell n's resulted. However, the cell n differences were considered acceptable.

Reliability estimates of the nine cloze tests used in the study indicated all the cloze tests were reliable measuring instruments.

Analyses of variances of the three passages' cloze scores revealed an inconsistency in the effect of amplification upon reading comprehension. For the insect passage, cloze scores significantly increased as amplification increased from low to moderate. For the deciduous tree passage, cloze scores significantly decreased as amplification increased from low to moderate and from low to high. For the sedimentary rock passage, no significant differences were noted, but a trend toward amplification causing a decrease in cloze scores was observed. No significant differences between interest levels nor significant interaction were observed in any of the three passages' cloze score data. Two factors were suggested as causing the inconsistencies in the results. One was the subjects' completing the low and moderate amplification versions more carefully than the high amplification versions. The second was the treatment's
lack of validity: the chances were not equally good of the concept exemplars being relevant concrete props in the cognitive structures of the subjects.

Correlations between cloze scores and interview ratings were all positive, but none were significant. The one factor suggested for this lack of positive significance was the interview instrument's lack of validity.
CHAPTER V

SUMMARY, CONCLUSIONS, IMPLICATIONS

Summary

It is desirable to use written materials in an elementary science program. However, since most currently available elementary science written materials are too difficult for elementary children, a problem exists. Readability formulas have been utilized to raise the readability level, but this has not resulted in higher reading comprehension levels.

In an attempt to solve this problem, the writer reviewed literature on communicating more effectively through writing. In doing so, it became apparent that writing more comprehensible materials involved first knowing how children comprehend reading material. Also since elementary science tends to deal in concepts, the problem became specifically, how do children comprehend and/or learn elementary science concepts through written materials. This, of course, led to the issue of whether concept learning is the same as verbal learning. In the first place, no one knows how children comprehend written material; there are several theories
concerning the process. In the second place, the writer rationalized that if verbal learning involved comprehending a concept which was verbally stated, it was indeed concept learning. With that settled, Ausubel's theory of meaningful verbal learning was used as a theoretical basis for reading comprehension in the study.

Ausubel's model for reading comprehension is called the subsumption process. There are three factors affecting this subsumption process—availability of subsumers in a learner's cognitive structure, discriminability of new material from the system that subsumes it, and the stability and clarity of the subsumers. Ausubel feels that to increase effectiveness of written communication one must enhance these three factors. One way to do this is to give many examples when writing about concepts. From this idea came the study hypothesis: Amplification of elementary science written materials through the use of exemplars will increase the reading comprehension of upper elementary children who read the material.

Three amplified versions of written passages dealing with three elementary science concepts were developed. Guidelines for effective writing of children's informational literature were utilized, such as maintaining an appropriate readability level, making sure the facts were accurate, and using personal references when possible. Also the concepts dealt with in the passages were selected so as to control
concept type and difficulty: Only real, classificational concepts were used.

To determine if the amplified passages caused an increase in reading comprehension, cloze tests on the amplified passages were developed. These were administered to 135 fifth grade children in Columbus Public Schools who were between ten and eleven years old and whose reading level ranged from 4.0 to 5.0 grades. Because literature had indicated interest might serve to enhance reading comprehension, the subjects were blocked on interest. The interest levels were determined by an interest inventory developed by the writer.

The cloze test scores were analyzed using a two-way analysis of variance. Results showed a significant difference between low and moderate amplification on the sedimentary rocks passage and a significant difference between low and moderate and low and high amplification on the deciduous trees passage. However on the latter passage, the amplification level means indicated amplification caused a decrease in cloze scores. There were no significant differences on the amplification in insect passage, but a trend similar to results in the deciduous trees passage was suggested. Results also indicated no significant interaction between interest and amplification and no significant differences between interest levels on any of the three passages.

Part of Ausubel's theory takes Piaget's theory of
cognitive development into account. He says that elementary children in the concrete stage of development can learn verbally stated material if they have the necessary concrete props in their cognitive structures. An interview instrument was also developed by the writer to determine the extent of the presence of concrete props in the cognitive structures of the subjects which were associated with the concept exemplars used in the study. The data from the interviews were correlated with the cloze test scores to determine the relationship between the two. No significant correlations were found.

In an attempt to explain the inconsistency in the amplification factor, the writer calculated the percentage of subjects having concrete props in their cognitive structures which were associated with the concept exemplars in the passages. The percentages were found to vary greatly. Since the criterion of all the exemplars having equally good chances of being relevant concrete props in the cognitive structures of the subjects was not met, moderate amplification was not actually three times as great as low amplification and high amplification was not five times as great as low amplification.

Conclusions

Hypothesis 1: There were significant differences (.05) between the cloze test scores of subjects who
read high, moderate, or low amplification versions of passages that treat three elementary science concepts. Specifically, there were significant differences between the cloze test scores of subjects who read low and moderate amplification versions of the sedimentary rock passage and who read low and moderate and low and high amplification versions of the deciduous tree passage.

Hypothesis 2: There were no significant differences (.05) between the cloze test scores, on passages that treat three elementary science concepts, of subjects who have received high, moderate, or low interest scores.

Hypothesis 3: There was no significant interaction (.05) between the amplification levels and the interest levels.

Hypothesis 4: There were no significant correlations (.05) between cloze test scores of subjects and interview ratings.

As indicated above, there were some significant differences in the data from this study. However, the computation of percentages of subjects having the concept exemplars as concrete props in their cognitive structures indicated the treatment of the study was somewhat invalid. Thus the writer feels any conclusions regarding hypotheses 1, 2, or
3 of the study are precluded by that invalidity and should not be made.

Hypothesis 4 is not affected by the treatment invalidity. Therefore, conclusions concerning it are somewhat credible. Since the correlations were positive, but not significant, the writer concluded that, in this study, cloze test scores are not strongly associated with the presence of concrete props related to the concept exemplars.

In this study, the constructs measured and manipulated are of different types, yet they all involve related aspects of the concept meanings. The treatment, amplification through the use of exemplars, attempted to impart an aspect of the meaning related to the concepts. The interview instrument attempted to measure an aspect of the meaning - the pre-concept meaning held by each subject. The interest inventory attempted to measure an aspect of the meaning as indicated by subjects' motivation to learn a particular concept. Finally, the cloze test attempted to measure the aspect of the meaning gained through interaction of the amplification meaning, the motivation meaning, and the pre-concept meaning held by the subjects. It is very difficult to know if these four constructs, one treatment and three instruments, manipulated and measured meaning in a comparable manner. The question is: Did what was imparted through amplification actually interact with what was measured by the interest inventory and the interview instrument? And
if so, was this interaction actually measured by cloze tests? The writer did make an accounting of the validity of the instruments used in the study, but in the face of such a complex phenomena as reading comprehension, one cannot be completely confident of validity determined by conventional means. Therefore, even if the treatment employed in the study had been valid, there is still the ultimate question of instrument validity. And this question would undoubtedly shroud the validity of further conclusions.

Implications

The study, of course, should be repeated, but with a more valid treatment. This can be accomplished by making an a-priori survey of children's concrete props related to particular concepts. From this information, one could then develop written passages truly of varying levels of amplification. Or the information could be used as a blocking variable. Also the reading level scores for the study sample should be more recent than those used in this study. Reading levels determined from scores made a month or two prior to a study would be much more valid than those determined from scores made a year prior to a study.

Several changes in study methodology were suggested by the results. (1) Because children who had low and moderate amplification versions seemed to complete their tests with more care than those children who had high amplification
versions, perhaps the length of the treatment-cloze tests should be reduced. In this way a fatigue-frustration factor may be reduced. (2) Since during the third treatment administration, children were observed to be using context more diligently then during the first and second treatment administrations, perhaps the subjects should be given several warm-up or practice cloze tests prior to the actual study. In this way the becoming-test-wise effect may not be so prevalent in the third treatment-test administration. (3) Finally, since the interview failed to distinguish between exemplars actually seen, exemplars heard about and exemplars read about, the question "How do you know?" should be added to the interview instrument.

It appears from the results that in order to use the idea of amplification through the use of exemplars as a guide to writing more comprehensible elementary science materials, a list of familiar concrete props will be needed. This will take a great deal of survey work. However, it is similar to a list already in existence - The Dale List of 3000 Familiar Words which is used with the Dale-Chall readability formula. Therefore, the writer considers the idea possible. The results of such a survey would also help to determine if a number of common familiar concrete props for children in the United States does exist. It could be that groups of common concrete props exist in various areas of the country, geographically and socio-economically. If
this is the case, then perhaps national science textbooks and trade books are not logical tools for elementary education; rather books written for regions or localities are.

A suggestion for further research would be to repeat the study but omit giving an interview to measure the extent of concrete props in the cognitive structures of the subjects. Instead unusual exemplars of concepts could be chosen and presented as concrete props to the study sample prior to administering the cloze tests. This would eliminate all the interview instrument invalidity questions and also enable the researcher to specifically assess the effects of concrete props in the reading process. It should be noted that this research procedure would be consistent with methodology advocated in most of the new elementary science curriculums.

The possibility of using exemplars to increase comprehension in other mediated materials is another area for further research. For example, would using more examples through the use of illustrations increase comprehension of printed materials? Knowledge of this might be very important for children whose reading level is low, but whose experiential background (concrete props) is fairly extensive.
Modified Richardson' s Science Interest Inventory

Instructions

"Boys and girls, I am here because I would like to find out what things you like to do. I know everybody enjoys doing some things, but not others; and I would like to know what it is you enjoy doing. This is not a test. You will not be graded on your answers. I only want to know how you feel about certain things. To answer the questions, you need only a pencil." Pass out answer sheets. "Look at page one. On it you see 38 numbers. For each number on your paper, I will read part of a sentence. You will complete the sentence by writing a, b, c, or d. You will write a if you would like very much to do what the sentence says. You will write b if you would like to do it somewhat. You will write c if you would dislike somewhat to do it, and you will write d if you would dislike very much to do it. I will read each sentence twice while you decide which letter to write. Find example one near the top of your sheet. Listen carefully while I read the sentence part which goes with that number. I would a, b, c, or d to visit the Center of Science and Industry." Repeat. "Now which letter will you write if you would like very much to visit the Center of Science and Industry?" Wait for response. "Yes, you will write a. If you would dislike somewhat to visit the Center
of Science and Industry, what letter would you write?" Wait for response. "Yes, you will write c. If you would dislike very much to visit the Center of Science and Industry, which letter will you write?" Wait for response. "Yes, you will write d. Now which letter will you write if you would like somewhat to visit the Center of Science and Industry?" Wait for response. "Yes, you will write b. Now, beside example one, write the letter which best describes how you feel about the sentence part." Walk around to see that all are following directions. "Good, are there any questions?" Wait for response. "Now we shall begin with the other numbers. Do not talk or ask questions once we begin. If you should break your pencil lead, raise your hand and I'll bring you another one. Alright, find number one and listen while I read the sentence part. Then write a, b, c or d to show how you feel about it."

Part One Interest Inventory Items

1. watch animals at the zoo. Repeat.
2. tinker with broken machines. Repeat.
3. be shown how to detect chemicals in unknown substances.
4. tour a company that makes rocket engines.
5. be taught how to predict the weather.
6. collect rocks.
7. watch a T.V. program or taking care of trees.
8. do experiments dealing with sound, heat, or light.
9. collect bugs.
10. have someone show me how to make a battery.
11. write a science fiction story about life on Mars.
12. read about how mountains were formed.
13. watch caterpillars develop into butterflies.
14. collect leaves.
15. have someone tell me how to take care of trees, bushes and other plants.
16. see a film about a mountain being formed.
17. read about explorations of the moon.
18. be told how different chemical elements were discovered.
19. see the planet Mars through a telescope.
20. tour an atomic power plant.
21. build an electric motor.
22. grow plants such as vegetables and flowers.
23. see a film about radiation belts in outer space.
24. listen to someone explain why baby animals look so much like their parents.
25. read about how the various organs of my body work.
26. tour a factory which makes plastic things.
27. make a display showing the things that are made from coal, limestone, or other natural resources.
28. see a rock collection containing gems, minerals, ores, and metals.
29. read about how to raise small animals such as cats, dogs, rabbits, etc.
30. see a film about how to build or repair various things.
31. be taught the names of common rocks.
32. visit a greenhouse or flower garden.

Instructions Continued

"That is the end of the first part. Turn to page two. This part is very similar to the first part. But this time, I will read a whole sentence; and you will choose the answer which best describes what you have done. You will write a if you do the activity very often or whenever possible. If you do the activity sometimes - more than twice - you will write b. You will write c if you do the activity rarely - only once or twice. And if you never do the activity, write d. Now find example two at the top of the second page. Listen while I read the sentence which goes with that number. I have picked wild flowers." Repeat. "Now if you pick wild flowers often or whenever possible, what letter would you write?" Wait for response. "Yes, you will write a. If you pick wild flowers only rarely - just once or twice, what letter will you write?" Response. "Yes, you will write c. If you never pick wild flowers, what letter will you write?" Response. "Yes, you will write d. Now if you have sometimes picked wild flowers - more than twice, which letter will you write?" Response. "Yes, you will write b. Now, beside example two, write the letter which best describes what you do concerning picking wild flowers." Walk around
to see if everyone is following directions. "Good, are there any questions?" Wait for response. "Now we shall begin with the other numbers. Find number one and listen while I read the sentence. Then write a, b, c, or d to show what you do."

Part Two Interest Inventory Items

33. I have taken apart mechanical things, such as a clock to see how they work. Repeat.
34. I have collected books or pictures of animals. Repeat.
35. I have read about how to build or repair various things.
36. I have collected magazine articles about the space program.
37. I have wondered what makes the stars appear to twinkle.
38. I have built animal shelters such as bird houses, dog houses, rabbit hutches, etc.
39. I have watched how birds, fishes, or other animals behave.
40. I have wondered how the universe was formed.
41. I have visited a zoo or other places where animals were shown.
42. I have wondered what causes salt to melt ice.
43. I have put together the parts of a machine and have had it work.
44. I have made simple mechanical repairs on bicycles, toys, water faucets, etc.
45. I have read about the possibility of life somewhere else in the universe.
46. I have raised a garden.
47. I have gone up to a strange machine to look it over.
48. I have made drawings of trees, flowers, and other plants on my own time.
49. I have read about atomic energy.
50. I have wondered how the mountains were formed.
51. I have read about how animals behave.
Example 1. __________________

I would a, b, c, or d to

a. like very much
b. like somewhat
c. dislike somewhat
d. dislike very much

1. __________________  20. __________________
2. __________________  21. __________________
3. __________________  22. __________________
4. __________________  23. __________________
5. __________________  24. __________________
6. __________________  25. __________________
7. __________________  26. __________________
8. __________________  27. __________________
9. __________________  28. __________________
10. __________________ 29. __________________
11. __________________ 30. __________________
12. __________________ 31. __________________
13. __________________ 32. __________________
14. __________________ 33. __________________
15. __________________ 34. __________________
16. __________________ 35. __________________
17. __________________ 36. __________________
18. __________________ 37. __________________
19. __________________ 38. __________________
Example 2.

- a. Very often (whenever possible)
- b. Sometimes (more than twice, but not very often)
- c. Very rarely (one or two times)
- d. Never
Directions for Science Interest Inventory

"Boys and girls, I am here because I would like to find out what things you like to do. This is not a test. You will not be graded on your answers. I only want to know how you feel about certain things. To answer the questions you will need only a pencil. I will now pass out the interest inventories. Do not open them until I tell you. When you receive the inventory, write your name in the place provided. Write your first and last names." Pass out the inventories. "This inventory has 72 statements. For each statement, you will indicate if it is true or false for you. If the statement is true for you, you circle the number of the statement. If the statement is false for you, do not circle the number. Leave it blank. Look at example one on the first page. I would like to have a guinea pig for a pet. If that statement is true for you, that is you would like to have a guinea pig for a pet, how would you indicate it on this paper?" Wait for response. "Yes, you would circle the number one. If that statement is false for you, that is you would not like to have a guinea pig for a pet, how would you indicate it on this paper?" Wait for response. "Yes, you would not circle the number. You would leave it blank. Now, you go ahead and indicate on your papers if example one is true or false for you. If it is true for you, circle the number. If it is false for you, leave it blank. Look at example
two. I don't care much about feeding animals. If that statement is true for you, that you don't care much about feeding animals, how would you indicate it on your paper?" Wait for response. "Yes, you would circle the number two. If that statement is false for you, that is you do like to feed animals, how would you indicate it on your paper?"

Wait for response. "Yes, you would leave it blank. Now, you go ahead and indicate on your papers if the example two is true or false for you. If it is true, you circle the number. If it is false, you leave it blank. Look at example three. I would like to fool around with electrical gadgets. If that statement is true for you, that is you would like to fool around with electrical gadgets, how would you indicate it on this paper?" Wait for response. "Yes, you would circle number three. If that statement is false for you, that is you would not like to fool around with electrical gadgets, how would you indicate it on this paper?"

Wait for response. "Yes, you would not circle the number. You would leave it blank. Now you go ahead and indicate on your papers if example three is true or false for you. If it is true, you circle the number, if it is false, you leave it blank. Look at example four. I wouldn't care if I had a magnet or not. If that statement is true for you, that is you would not care if you had a magnet or not, how would you indicate it on your paper?" Wait for response. "Yes, you would circle the number four. If that statement is false
for you, that is you would like to have a magnet, how would you indicate it on your paper?" Wait for response. "Yes, you would not circle the number. You would leave it blank. Now, you go ahead and indicate on your papers if example four is true or false for you. If it is true, you circle the number. If it is false, you leave it blank. Are there any questions? Be sure to read each statement carefully. When you finish turn your paper over and sit quietly at your desk. You may begin."

Directions for Scoring the Interest Inventory

To score the science interest inventory, count every positive response a correct answer. This way, if a subject leaves a negatively stated statement blank, he is indicating a positive response and is given credit for a correct response. The correct answers are totaled, and this total is the measure of interest a subject has in a particular subject area of science - rocks, trees or insects.
Science Interest Inventory

Example 1. I would like to have a guinea pig for a pet.
Example 2. I don't care much about feeding animals.
Example 3. I would like to fool around with electrical gadgets.
Example 4. I wouldn't care if I had a magnet or not.
1. In the summer, I like to go out at night and see the bugs that fly around lights.
2. I don't care about making a leaf collection.
3. I would like to hear about the bugs that lived a long time ago.
4. I would not like to see a film on rocks very much.
5. I like to play in the leaves in the fall.
6. It would be interesting to gather and press leaves.
7. In the summer, I don't much like to go out at night and see the bugs that fly around lights.
8. I wouldn't care about seeing a volcano.
9. It would be fun to do art work with rocks.
10. If I found a cocoon, I would keep it to watch.
11. I would like to see rocks that glow in the dark.
12. I like an artificial Christmas tree best.
13. I wouldn't care much about visiting the petrified forest.
14. I don't have a favorite tree.
15. I think it would be fun to collect rocks.
16. I would like to have a large crystal.
17. It would not be much fun to collect bugs
18. I wouldn't care much about visiting a coal mine.
19. I would like to see pictures of bugs from far-off places.
20. I would like to see a film on rocks.
21. I have a favorite tree.
22. I would not care much about going in a cave.
23. I would not like to see movies about bugs very much.

go on to the next page
24. It would not be very interesting to gather and press leaves.
25. I don’t think bugs are very interesting.
26. I don’t care about playing in the leaves in the fall.
27. I wouldn’t like to go on a field trip to see where bugs live very much.
28. I would like to hunt for fossils.
29. I don’t care much about looking at rocks under a magnifying glass.
30. I do not care if I had a large crystal or not.
31. It wouldn’t be much fun to do art work with rocks.
32. I wouldn’t like to be a rock scientist when I grow up.
33. It is not much fun to do art work with leaves.
34. I wouldn’t like to see pictures of plants in other parts of the world very much.
35. When I grow up, I will plant trees in my yard.
36. I would like to go on a field trip to see where bugs live.
37. It is fun to do art work with leaves.
38. I would like to help plant a burned forest.
39. I do not like to walk in the woods.
40. I would like to visit the petrified forest.
41. I like to watch bugs.
42. It would be fun to watch bugs grow up.
43. I would like to make a leaf collection.
44. I would like to look at bugs under a magnifying glass.
45. I would like to go in a cave.
46. I wouldn’t care much about hearing about the bugs that lived a long time ago.
47. I would like to see a volcano.

Go on to the next page
48. I would like to visit a coal mine.
49. It would not be much fun to watch bugs grow up.
50. I wouldn't care much about looking at bugs under a magnifying glass.
51. I do not care about lying down under trees and looking up at the leaves.
52. I would not care much about having a bug cage.
53. I might like to be a rock scientist when I grow up.
54. I do not think it would be much fun to collect rocks.
55. I do not like to watch bugs.
56. I would not care about seeing rocks that glow in the dark.
57. I wouldn't care much about seeing pictures of bugs from far-off places.
58. I would not like to help plant a burned forest.
59. I would like to look at rocks under a magnifying glass.
60. I like to walk in the woods.
61. I like to lie down under trees and look up at the leaves.
62. It would be fun to collect bugs.
63. I would like to see movies about bugs.
64. When I grow up, I probably won't plant trees in my yard.
65. I would like to see pictures of plants in other parts of the world.
66. I would like to have a bug cage.
67. I do not care about picking up rocks and looking at them.
68. I would not like to hunt for fossils very much.
69. I like a live Christmas tree best.
70. If I found a cocoon, I would not keep it.
71. I like to pick up rocks and look at them.
72. I think bugs are very interesting.

end
Directions for Cloze Test

First Administration

"Boys and girls, I am going to give you a science test. It is a test, but how well you do on it will not affect your grade in school. Only I will know how you do. To take the test you need a pencil. Your teacher will now pass out the tests. Do not look at the test until I tell you to do so."

Pass out the tests. "In a few minutes you will read a passage. As you read you will notice that some of the words have been left out. In place of these words, you will find a blank. You are to try to guess what words have been left out. Look at the example on the board. Our school is on ________ Street. It is for ________ in grades kindergarten to ________. What words do you think were left out in the three blanks." Pause. "What word do you think was left out of the first blank?" Wait for response. "Yes, the word ________. What word do you think was left out of the second blank?" Wait for response. "Yes, children (or kids or students). What word do you think was left out of the third blank?" Wait for response. "Yes, the word six. Good." "Look again at the blanks. As you can see, the blanks are all the same size, but the word you write may be short or long. It could be a word like students, or it could be a word like six. Notice that only one word is used for each blank. You will find some blanks very easy and
some very hard. You will be able to fill in the blanks easier if you use the other words in the passage to help you decide what the best word is. Do not be afraid to guess what a word might be. Try to fill in all the blanks. If you don't know how to spell a word which goes in one of the blanks, just spell it as well as you can. Are there any questions? You have 25 minutes to complete the test. Work carefully and check over your work. Some of you have short tests, and some have long tests. Therefore some of you may finish before others. If you do finish before others, sit or work quietly at your seat until the 25 minutes is up. Do not get out of your seat. Do not talk to or in any way disturb other people until time is called. All right, you may begin."

Second Administration

"Boys and girls, I am going to give you another test. It is a test, but how well you do on it will not affect your grade in school. Only I will know how you do. To take the test you will need a pencil. Your teacher will now pass out the tests. Do not look at the test until I tell you to do so." Pass out the tests. "In a few minutes you will again read a passage. As you read you will again notice that some of the words have been left out. In place of these words, you will find a blank. You are to try to guess what words have been left out. You may remember the example we did on
the board last time. Our school is on _______ Street. It is for _______ in grades kindergarten to _______. The words you thought were left out of the blanks were ________, ________, and ________. Remember the blanks are all the same size, but the word you write may be short or long. It could be a word like students or a word like six. Remember that only one word is used for each blank. You may remember that some blanks were easy and some were hard. This time try to use the other words in the passage to help you decide what the best word is. Do not be afraid to guess what a word might be. Try to fill in all the blanks. Remember if you don't know how to spell a word, just spell it the best you can. Are there any questions? Again, you have approximately 25 minutes to complete the test. Work carefully, and check over your work. Again some of you have short tests and some have long tests. And this time some of you may have the same test or parts of the same test you had before. Therefore some of you may finish before others. If you do finish before others, sit or work quietly at your seat until the 25 minutes is up. Do not get out of your seat. Do not talk or in any way disturb other people until time is called. All right, you may begin."

Third Administration

"Boys and girls, I am going to give you another science test. Remember it is a test, but how well you do on it
will not in any way affect your grade in school. Only I will know how you do. To take the test, you will need a pencil. Your teacher will now pass out the tests. Remember, do not look at the tests until I tell you to." Pass out the tests. "Again you are going to read a passage in which some of the words have been left out. In place of those words you know there will be a blank. Your job, of course, is to fill in the blanks with the words you think make most sense in the sentences. Remember the blanks are all the same length, but the words you put in the blanks may be of different lengths. Remember there should be only one word in each blank. Remember, some of the blanks are hard and some are easy. However, you will find it easier to fill in the hard blanks if you use the other words in the passage to help you decide what the best word is. Remember, try to fill in all the blanks, and don't be afraid to guess what a word might be. Remember if you don't know how to spell a word, just spell it the best you can. Are there any questions? Again you have approximately 25 minutes to complete the test. Work carefully and check over your work. Again some of you have short tests and others have long tests, and again you may have the same test or parts of the same test you have had before. Therefore some of you will probably finish before others. If you do finish before others, remember, sit or work quietly at your seat. Do not get up,
talk, or in any way disturb other people until time is called. All right, you may begin."

Directions for Scoring the Cloze Tests

Only the exact words are counted as correct. Synonyms are not counted. Any decodable mispellings are allowed. The correct answers are totaled and converted to percentages.
There are many rocks _________ earth. These rocks were
___________ all made in the _________ way. Some rocks were
___________ from very little bits _________ other rocks.
Sometimes water _________ these small bits of _________ into
creek and river _________ Sometimes the wind blows _________
into valleys. There they _________ up. The pile becomes _________
heavy. Finally the ones _________ the bottom are squeezed _________
into a rock. A _________ made this way is _________ sedimentary.

One sedimentary rock _________ may have picked up _________
sandstone. It is stone _________ of sand. The sand _________
been stuck together. If _________ rub sandstone, some of _________
little bits of sand _________ come off. Old Man’s _________ is
made of sandstone. _________ walls of the cave _________ very
sandy. The sandstone _________ have seen may have _________
brown, gray, or red. _________ can be all those _________.
Some sandstone is used _________ building. There are some _________
at Ohio State University _________ of sandstone.
There are many rocks on earth. These rocks were all made in the same way. Some rocks were from very small bits other rocks. Sometimes water these small bits of into creek and river. Sometimes the wind blew into valleys. There they up. The pile becomes heavy. Finally the ones the bottom are squeezed into a rock. A made this way is sedimentary.

One sedimentary rock may have picked up sandstone. It is stone of sand. The sand been stuck together. If rub sandstone, some of little bits of sand come off. Old Man's is made of sandstone. walls of the cave very sandy. The sandstone have seen may have brown, gray, or red. can be all those. Some sandstone is used building. There are some at Ohio State University of sandstone.

There is well known place in United States which is of sedimentary rocks. It Grand Canyon. You may seen a picture of great big ditch. It wide, colored lines running from top to bottom. these wide colored lines sedimentary rocks.

Another sedimentary you may have picked is shale. Most of time it is a gray or brown rock. shale is wet, it like mud. This is shale was once mud, tiny pieces of earth mud were squeezed together make a rock.
There are many rocks in the earth. These rocks were all made in the same way. Some rocks were from very small bits other rocks. Sometimes water these small bits of into creek and river. Sometimes the wind blows into valleys. There they up. The pile becomes heavy. Finally the ones the bottom are squeezed into a rock. A made this way is sedimentary.

One sedimentary rock may have picked up sandstone. It is stone of sand. The sand been stuck together. If rub sandstone, some of little bits of sand come off. Old man's is made of sandstone. walls of the cave very sandy. The sandstone have seen may have brown, gray, or red. can be all those. Some sandstone is used building. There are some at Ohio State University of sandstone.

There is well known place in United States which is of sedimentary rocks. It Grand Canyon. You may seen a picture of great big ditch. It wide, colored lines running from top to bottom. these wide colored lines sedimentary rocks.

Another sedimentary you may have picked is shale. Most of time it is a gray or brown rock. shale is wet, it like mud. This is shale was once mud. tiny pieces of earth mud were squeezed together make a rock.
Another kind of soil in Texas is a conglomerate. It looks like several sizes of rocks all stuck together. You may have seen conglomerates around where roads are built. Or you may see them in places gravel is being dug.

The state capital building in Columbus is built sedimentary rock. The rock is limestone. As you might guess, this is stone made into limestone. Limestone is what in sea shells.

So, can be made of pieces of sea shells. of the time, these are too small to see. Sometimes, however, you can see the sea shells.
You may have noticed there are two kinds of _________. The leaves of one ________ change colors in the _________. The color of the _________ may change from green _________ yellow. It may change _________ green to brown. Then _________ leaves drop off.
The _________ limbs are bare all _________. In the spring, new _________ leaves grown on the _________ again. These trees, whose _________ change colors and drop _________ every year, are called deciduous.

_________ deciduous tree you _________ have seen is the _________. Its leaf looks something ________ a hand. It has _________ points. These leaves make _________ maple tree one of _________ most beautiful we see _________ the fall. Before dropping _________, the maple leaves turn _________ yellow, orange and red. _________ people take rides around _________ city just to see _________ maple trees. In the _________ some maple trees can _________ cut to give maple _________.
You may have noticed there are two kinds of trees. The leaves of one kind change colors in the fall. The color of the leaves may change from green to yellow. It may change green to brown. Then the leaves drop off. The limbs are bare all winter. In the spring, new leaves grow on the tree again. These trees, whose leaves change colors and drop off every year, are called deciduous.

Deciduous trees you have seen in the city. Its leaf looks something like a hand. It has points. These leaves make the maple tree one of the most beautiful in the fall. Before dropping, the maple leaves turn yellow, orange and red. People take rides around the city just to see maple trees. In the fall some maple trees can cut to give maple sap.

Sassafras is another deciduous you may have seen. It is so small you think it is just a bush. The leaves of sassafras look like mittens. Of all, the twigs taste like beer. In the fall, leaves of the sassafras turn orange and red before off.

One of the deciduous trees is the redbud. Many people plant them in their yards. In the spring, the redbuds are covered with tiny pink flowers. Soon flowers become seed pods that look like green beans. You may have noticed the shaped leaves on the tree. These leaves turn yellow before dropping off.
You may have noticed two kinds of

The leaves of one kind change colors in the fall. The color of the leaves may change from green to yellow. It may change green to brown. Then leaves drop off. The limbs are bare all winter. In the spring, new leaves grow on the tree again. These trees, whose leaves change colors and drop every year, are called deciduous.

deciduous tree you have seen is the maple. Its leaf looks something a hand. It has points. These leaves make maple tree one of most beautiful we see the fall. Before dropping, the maple leaves turn yellow, orange and red.

people take rides around city just to see maple trees. In the some maple trees can cut to give maple.

Sassafras is another deciduous you may have seen. it is so small think it is just bush. The leaves of sassafras look like mittens. of all, the twigs the sassafras taste like beer. In the fall, leaves of the sassafras orange and red before off.

One of the deciduous trees is the redbud. Many people plant them their yards. In the spring, the redbuds are covered tiny pink flowers. Soon flowers become seed pods look like green beans. may have noticed the shaped leaves on the tree. These leaves turn yellow before dropping off.
deciduous tree which is known as the buckeye. Other deciduous trees, the leaves of which is brown, lose its leaves in fall. You may know buckeye gets its name from its nut. On the brown nut is a spot. This spot looks like an eye.

The weeping willow is a deciduous tree may like. Often you see on the banks of the banks of the whole tree hangs. The long leaves almost the ground. It looks if it is sad. In the winter, thin bare branches rattle when the wind blows.
Some of the smallest insects we can see are insects. You cannot see insects. They are everywhere on earth. When are grown up, they three pairs of legs — in all. These six are joined. That means leg can bend in places. The six legs not always all alike. there are always six on every insect.

One of insect that can is called a butterfly. butterfly has beautiful wings are black, yellow, brown, orange. You may have some butterflies flying near flowers. Butterflies eat the liquid found in flowers. you were to look a butterfly very closely, would see it had joined legs — three on side of its soft . When a butterfly lands a flower, it uses its back four legs. two front legs are close to its body.
Passage 3

Some of the smallest ___________ we can see are ___________
insects. You can not ___________ seeing insects. They are ___________
everywhere on earth. When ___________ are grown up, they ___________
three pairs of legs - ___________ in all. These six ___________ are
jointed. That means ___________ leg can bend in ___________ places.
The six legs ___________ not always all alike. ___________, there are
always six ___________ on every insect.

One ___________ of insect that can ___________ is called a
butterfly. ___________ butterfly has beautiful wings ___________ are
black, yellow, brown, ___________ orange. You may have ___________
some butterflies flying near ___________ flowers. Butterflies eat the
___________ liquid found in flowers. ___________ you were to look
___________ a butterfly very closely, ___________ would see it had
___________ jointed legs - three on ___________ side of its soft
___________. When a butterfly lands ___________ a flower, it uses
___________ its back four legs. ___________ two front legs are
___________ close to its body.

___________ insect you may have ___________ is the roach.
Roaches ___________ often seen in houses. ___________ like to eat
food ___________ is left out at ___________. Most people do not
___________ reaches around and will ___________ to kill them.
Roaches, ___________ other insects, have six ___________ legs.

An insect which ___________ fun to watch is ___________ ant.
Ants are always ___________. Their six little jointed ___________
ever seem to stop. ___________ may have seen ants ___________ in and
out of ___________ small hole in the ___________. This is the door
___________ the ant house. Many ___________ live and work together
___________, one house.
Some of the smallest __________ we can see are __________ insects. You can not __________ seeing insects. They are __________ everywhere on earth. When __________ are grown up, they __________ three pairs of legs - __________ in all. These six __________ are jointed. That means __________ leg can bend in __________ places. The six legs __________ not always all alike. __________, there are always six __________ on every insect.

One __________ of insect that can __________ is called a butterfly. __________ butterfly has beautiful wings __________ are black, yellow, brown, __________ orange. You may have __________ some butterflies flying near __________ flowers. Butterflies eat the __________ liquid found in flowers, __________ you were to look __________ a butterfly very closely, __________ would see it had __________ jointed legs - three on __________ side of its soft __________. When a butterfly lands __________ a flower, it uses __________ its back four legs. __________ two front legs are __________ close to its body.

__________ insect you may have __________ is the roach. Roaches __________ often seen in houses. __________ like to eat food __________ is left out at __________. Most people do not __________ roaches around and will __________ to kill them. Roaches, __________ other insects, have six __________ legs.

An insect which __________ fun to watch is __________ ant. Ants are always __________. There six little jointed __________ never seem to stop. __________ may have seen ants __________ in and out of __________ small hole in the __________. This is the door
Many live and work together one house.

An interesting insect is the grasshopper. I may have walked through tall grass and made grasshoppers hop. They hop high. Their two big jointed legs are made for jumping. Their four jointed legs are smaller. Sometimes make a loud noise.

The walking stick an odd insect. It like a green or stick with six jointed. This makes the walking hard to tell from twigs and leaves. Some like to have walking for pets. They carry on their shoulders.
Directions for Cloze and Objectives Tests
(Pilot Study)

"Boys and girls, I am going to give you a science test. It is a test, but how well you do on it will not affect your grade in school. Only I will know how you do on the test. To take the test you will need a pencil. There are two different tests. Some of you will get one test. Some of you will get the other test. I am going to tell you about both tests before any of you begin. So, since you do not know which test you will get, you must listen carefully to the directions for each test." Walk to test one example written on blackboard. "On test one you will read three passages. As you read you will notice that some of the words have been left out. In place of these words, you will find a blank. You are to try to guess what words have been left out. Look at the example on the board. Our school is on __________ Street. It is for __________ in grades kindergarten to __________. What word do you think was left out of the first blank?" Wait for response. "Yes, Evening was left out." Write Evening in the blank. "What word do you think was left out of the second blank?" Wait for response. "Yes, children, kids, students was left out." Write word in blank. "What word do you think was left out of the third blank?" Wait for response. "Yes, six was left out." Write six in the blank. "Good. Look again at the blanks. As you can
see, the blanks are all the same size, but the word you
write may be short or long, and only one word can be used
in each blank. It could be a word like Evening or a word
like six. If you don't know how to spell a word which goes
in one of the blanks, just spell it as well as you can. Do
your best to fill in every blank. You will find some blanks
very easy and others very hard. You will be able to fill in
the blanks easier if you do not skip any of the reading. Do
not be afraid to guess what a word might be. Try to fill in
all of the blanks. Are there any questions?" Walk to test
two example written on the black board. "Test two consists
of 15 questions like the example on the board. Each ques­
tion has five possible answers under it. You are to circle
the number of the figure which you think best answers the
question. Look at the example. Which of these shows a tri­
gle with three dots in it?" Wait for response. "Yes,
figure four is the correct answer. Why isn't number one the
correct answer?" Wait for response. "Yes, it doesn't have
three dots. Why isn't number five the correct answer?"
Wait for response. "Yes, figure five has three dots, but
it isn't a triangle. Good. Are there any questions? You
have approximately 40 minutes to complete the test. When
you finish, check over your work. Turn over your paper and
sit or work quietly at your desk. Do not get out of your
seat. Do not talk or in any way disturb other people until
everyone is finished. I will now pass out the tests. Do not open them until I tell you to do so." Pass out tests. "You may begin."
SCIENCE TEST
1. Which of these shows a butterfly with the correct number of legs?

2. Which of these shows a roach with the correct number of legs?

3. Which of these shows an ant with the correct number of legs?

4. Which of these shows a grasshopper with the correct number of legs?

5. Which of these shows a walking stick with the correct number of legs?
1. Which of these best shows a maple tree in the winter?

2. Which of these best shows a sassafras tree in the summer?

3. Which of these best shows a redbud tree in the winter?

4. Which of these best shows a hickory tree in the fall?

5. Which of these best shows a weeping willow tree in the winter?
1. Which of these best shows what sandstone is made of?

2. Which of these best shows what the Grand Canyon looks like?

3. Which of these best shows what shale is made of?

4. Which of these best shows what a conglomerate is made of?

5. Which of these best shows what limestone is made of?
Directions for Interview

Be positive at all times. "Hi ___name____. I am going to ask you to describe some things. While you are describing, I am going to be putting some checks on this sheet of paper. OK?"

1. Can you describe a butterfly?

<table>
<thead>
<tr>
<th>prompting questions to be used only if necessary</th>
</tr>
</thead>
<tbody>
<tr>
<td>What does it look like? What color is it?</td>
</tr>
<tr>
<td>About how big is it? How does it move?</td>
</tr>
<tr>
<td>What have you observed it doing? Where have you seen it?</td>
</tr>
<tr>
<td>How many legs does it have?</td>
</tr>
</tbody>
</table>

Repeat the above for roach, ant, grasshopper, and walking stick.

2. Can you describe a maple tree?

<table>
<thead>
<tr>
<th>prompting questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>What does it look like? What shape are the leaves? Where have you seen it?</td>
</tr>
<tr>
<td>Is there anything unusual about it that you have noticed?</td>
</tr>
<tr>
<td>What does it look like at different times of the year?</td>
</tr>
</tbody>
</table>

Repeat the above for willow, buckeye, redbud, and sassafras.

3. Can you describe sandstone?

<table>
<thead>
<tr>
<th>prompting questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>What does it look like? How does it feel? What color is it?</td>
</tr>
<tr>
<td>Where have you seen it? Is there anything else you have noticed about it?</td>
</tr>
<tr>
<td>What do you think it is made of?</td>
</tr>
</tbody>
</table>

Repeat for shale, limestone, conglomerate, and Grand Canyon.
<table>
<thead>
<tr>
<th>Distinguishing Features of Exemplar</th>
<th>Concept Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>butterfly</strong> bright, colored wings, flies around flowers.</td>
<td>six legs</td>
</tr>
<tr>
<td><strong>roach</strong> shiny brown plastic like beetle, seen around food and water at night; crawls</td>
<td>six legs</td>
</tr>
<tr>
<td><strong>ant</strong> lives in ground with lots of other ants, brown, black, red, bites, crawls very fast as if very busy</td>
<td>six legs</td>
</tr>
<tr>
<td><strong>grasshopper</strong> large back legs, jumps, often heard in fields or tall grass, makes buzz or chirping noise; eats grass, wheat, etc, green or brown</td>
<td>six legs</td>
</tr>
<tr>
<td><strong>walking stick</strong> looks like a green or brown stick, crawls on trees, feels prickly if it crawls on you</td>
<td>six legs</td>
</tr>
<tr>
<td>Rocks</td>
<td>Distinguishing Features of Exemplar</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td></td>
<td>sandstone</td>
</tr>
<tr>
<td></td>
<td>feels sandy, old mans cave is made of it</td>
</tr>
<tr>
<td>sandstone</td>
<td></td>
</tr>
<tr>
<td></td>
<td>shale</td>
</tr>
<tr>
<td></td>
<td>smells like mud when wet, usually smooth and flat, dark brown, gray or black</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>conglomerate</td>
<td>little rocks all stuck together with other stuff in between, around gravel quarries and road construction</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>can see sea shells in it, fizzes if coke or vinegar is put on it,</td>
</tr>
<tr>
<td>limestone</td>
<td></td>
</tr>
<tr>
<td></td>
<td>layers or lines running sideways along the sides, very deep ditch, colorful</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Grand Canyon</td>
</tr>
</tbody>
</table>
### Distinguishing Features of Exemplars

<table>
<thead>
<tr>
<th>Concept Parameters</th>
<th>Distinguishing Features of Exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>got sugar or syrup from it, leaves with 5 points, bright yellow in the fall</td>
</tr>
<tr>
<td>maple</td>
<td>tree bare in winter</td>
</tr>
<tr>
<td></td>
<td>little pink flowers in the spring, many in yards, not in woods, heart shaped leaves</td>
</tr>
<tr>
<td>redbud</td>
<td>tree bare in winter</td>
</tr>
<tr>
<td></td>
<td>long droopy leaves and branches, often near water</td>
</tr>
<tr>
<td>willow</td>
<td>tree bare in winter</td>
</tr>
<tr>
<td>sassafras</td>
<td>sitten shaped leaves with thums, kind of fuzzy leaves, branches and roots can be boiled or chewed and will taste like root beer</td>
</tr>
<tr>
<td></td>
<td>tree bare in winter</td>
</tr>
<tr>
<td>buckeye</td>
<td>has brown shiny nut with light spot on it. leaves are in groups of five.</td>
</tr>
<tr>
<td></td>
<td>tree bare in winter</td>
</tr>
</tbody>
</table>
Directions for Rating Interviews

If the subjects can distinguish an example of a concept from other examples, one check is given. If the subjects know the characteristic of the example which enables it to be classified as the concept, a second check is given. If the subjects receive both checks for a particular example, they are given one point. The final rating for the subjects is the total number of points received.
Appendix B
Report of the Pilot Study

In February, 1972, prior to the formal pilot, the researcher informally assessed fifth grade children's interest in science in order to establish a firmer base for choosing three concepts, which would be of interest, to be used in the treatment passages on the study. This was done by giving a general science interest inventory to four fifth grade classes, one from each of four schools in the Columbus Public School District, the District from which the study sample was drawn. The four schools included an inner city school, a fringe inner city school, a blue collar suburban school, and a white collar suburban school — as characterized by the elementary science resource teacher of Columbus Public Schools. The interest inventory given was a modification of an Inventory to Measure Within the Domain of Curiosity and Interests in Science of Upper Elementary School Children (Richardson, 1971). The modification included shortening the inventory and administering it orally instead of having it read by the students. The inventory was given orally because readability of the instrument was at grade levels five to six, which the researcher considered too high for some fifth grade students. The shortening occurred as a result of establishing some degree of content validity. Richardson had not intended the instrument to be used as a measure of interest in different science areas, but of general interest
in science. Therefore it was necessary to classify the item into different science areas. The researcher and a science education graduate student both classified the items, and then only those items which were similarly classified by both raters were retained. Informal inspection of the mean scores for each of the six subtests - physical science, chemical science, earth science, space science, botanical science, and zoological science - revealed no outstanding differences. The scores for the subtests were respectively 2.50, 2.08, 2.13, 2.14, 2.00, 2.05 on a scale of 1 to 4, 1 indicating high interest and 4 indicating low interest. The results suggested that fifth grade children in Columbus Public Schools are probably equally interested in all six science areas. On the basis of this, the researcher decided to retain the three concepts previously suggested - one in earth science, one in botanical science, and one in zoological science (see Appendix A for a copy of the modified version of Richardson's Inventory).

The formal pilot was conducted at Evening Street Elementary School in Worthington, Ohio during March and April, 1972. The school records were examined, and only 48 subjects found with the population characteristics - reading level between 4.0 and 6.0 grade levels and age between 10.0 years and 11.0 years. Since more subjects were needed, it was decided to redefine the reading level requirement to any level above 4.0. This brought the total population to 86.
At the principal's request, the instruments were administered to the population along with their classmates in their regular classrooms.

In order to obtain the blocking variable, the interest inventory was given first. All the subjects completed the inventory on the same day at different times. For each of the subtests, an item analysis and a factor analysis was done, the former by the Item Analysis Program developed by the Ohio State University Center for Measurement and Evaluation and the latter by the Biomedical Computer Program, General Factor Analysis (BMD03M).

On the Interest-in-Trees subtest, the estimate of the reliability coefficient of internal consistency, computed by the Kuder Richardson 20 formula, was a satisfactory .817. The phi coefficients, which are measures of the items' ability to detect differences between extreme groups, were all significant at the .05 level or better. The Point Biserial Correlations, which are measures of the relationship of information given by the item and the information reflected by the total scores, were all positive and significant at the .05 level or better. As such, the Point Biserial Correlations can also be considered measures of the validity of the items. Factor analysis of the subtest revealed five factors - all of which the researcher considers to be associated with interest in trees. Items 13, 35, 38, 40, 58, and 64 seemed to measure interest in large number
of trees such as in forests of different kinds. Items 2, 6, 24, 33, 37, and 43 seemed to measure interest in collecting and doing activities with leaves. Items 14 and 21 seemed to measure interest in a particular tree. Items 5, 26, 39, 51, 60, and 61 seemed to measure interest in being outdoors - walking in the woods, playing in the leaves, etc. Items 12 and 69 seemed to measure interest in the characteristics of Christmas trees. Items 33 and 65 did not seem to be measuring any of these specific factors. These five factors accounted for 47% of the variance.

On the Interest-in-Rocks subtest, the estimate of the reliability coefficient of internal consistency, as computed by the Kuder Richardson 20 formula, was a satisfactory .874. The phi coefficients of the items were all significant at the .05 level or better. The Point Biserial Correlations were all significant at the .05 level or better. Factor analysis revealed three factors - all of which the researcher considers to be associated with interest in rocks. Items 3, 4, 5, 8, 13, 18, and 21 seemed to measure interest in artistic aspects of rocks, such as art work with rocks. Items 1, 6, 7, 15, 16, and 17 seemed to measure interest in the more intriguing aspects of rocks, such as going in caves, seeing volcanos, etc. Items 1, 5, 8, 11, 19, 20, 22, 23, 24 seemed to measure general 'rock hound' interests, such as collecting rocks, picking up rocks and looking at them, etc. Items 2, 9, 10, 12, and 14 did not seem to measure any
of these specific factors. These three factors accounted for 37% of the variance.

On the Interest-in-Insects subtest, the estimate of the reliability coefficient of internal consistency, as computed by the Kuder Richardson 20 formula was a satisfactory .928. The phi coefficients of the items were all significant at the .001 level. The Point Biserial Correlations were all significant at the .01 level. Factor analysis revealed four factors - all of which the researcher considers to be associated with interest in insects. Items 1, 3, 8, 10, 11, 18, 20, and 24 seemed to measure interest in nongoal oriented activities associated with bugs, such as general watching and collecting. Items 6, 8, 9, 10, 12, 13, 14, 15, 16, 18, and 19 seemed to measure interest in more goal oriented activities associated with bugs, such as watching bugs grow up or going on a field trip to see where bugs live. Items 7, 16, 17, 21, and 22 seemed to measure interest in activities associated with bugs which did not involve direct contact with them, such as seeing movies on bugs or looking at them under a magnifying glass. Items 4, 6, 13, and 23 seemed to measure interest in perhaps the unusual aspects of bugs, such as watching a cocoon or seeing pictures of bugs from far off places. Item 2 did not seem to measure any of these specific factors. These four factors accounted for 57% of the variance.

From the preceding results, the researcher concluded
the Interest Inventory was satisfactorily reliable and valid. Also from the pilot study, it was concluded that 30 minutes was an appropriate time allotment for administering the inventory. It was found, however, that the instructions were not completely adequate: more examples were needed in order for the children to understand how to indicate their interest, and some remarks were superfluous and could be omitted. Even though the inventory had been written at below fourth grade level, the children had been asked to circle any words or sentences which they did not understand. There were four words and two sentences circled by one to four students. However, the researcher felt these were not sufficient numbers to warrant changes. Therefore the inventory was considered suitably readable.

Once the blocking variable was obtained, the main thrust of the pilot study was carried out. Six students from the defined population in each class were randomly assigned to one of the six treatment orders - I (low amplification) II (moderate amplification) III (high amplification), I III II, II I III, II III I, III II I, III I II. Each of these students had previously been randomly assigned to one of the six concept orders R (sedimentary rocks) T (deciduous trees) I (insects), R I T, I R T, I T R, T R I, T I R. The median for interest on each of the science inventory sub-tests was computed; and those with interest scores above the median were placed in the high interest group; and those
with interest scores below the median were placed in the low interest group.

The researcher had intended to have equal n's in all the cells. However, by mistake, the subjects were randomly assigned to the three treatment groups before blocking, and unequal n's in some of the cells were the result. The resulting data matrices are shown in Fig. 8, 9, and 10. In the major study, the subjects will be blocked prior to random assignment, and each level of interest will be randomly assigned separately.

The treatments and tests were administered simultaneously at different times on two consecutive days for the six classes involved. Analyses of variances of the results of the cloze tests were computed using the MANOVA program. Six subjects did not complete the treatment-tests, and their scores were removed prior to running the MANOVA program. Because of the probability of the resulting F's computed from unequal n's being biased, an F ratio was computed to test the significance of the differences between the variances in each of the three passages. None of the F Ratios were significant at the .02 level, thus suggesting the F's in the analyses of variances were not biased.

The means and standard deviations and F statistics for the sedimentary rock passage cloze scores are shown in Tables 16 and 17. Since the interaction was significant, the means were graphed (see Fig. 11).
<table>
<thead>
<tr>
<th>INTEREST</th>
<th>B₁</th>
<th>B₂</th>
<th>B₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>(x_{111})</td>
<td>(x_{121})</td>
<td>(x_{131})</td>
</tr>
<tr>
<td></td>
<td>(x_{112})</td>
<td>(x_{122})</td>
<td>(x_{132})</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>(x_{117})</td>
<td>(x_{127})</td>
<td>(x_{134})</td>
</tr>
<tr>
<td>A₂</td>
<td>(x_{211})</td>
<td>(x_{221})</td>
<td>(x_{231})</td>
</tr>
<tr>
<td></td>
<td>(x_{212})</td>
<td>(x_{222})</td>
<td>(x_{232})</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>(x_{215})</td>
<td>(x_{225})</td>
<td>(x_{238})</td>
</tr>
</tbody>
</table>

Fig. 8. Data matrix for sedimentary rocks passage.
**AMPLIFICATION**

<table>
<thead>
<tr>
<th></th>
<th>(B_1)</th>
<th>(B_2)</th>
<th>(B_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1)</td>
<td>(x_{111})</td>
<td>(x_{121})</td>
<td>(x_{131})</td>
</tr>
<tr>
<td></td>
<td>(x_{112})</td>
<td>(x_{122})</td>
<td>(x_{132})</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(x_{116})</td>
<td>(x_{126})</td>
<td>(x_{136})</td>
</tr>
<tr>
<td>(A_2)</td>
<td>(x_{211})</td>
<td>(x_{221})</td>
<td>(x_{231})</td>
</tr>
<tr>
<td></td>
<td>(x_{212})</td>
<td>(x_{222})</td>
<td>(x_{232})</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(x_{216})</td>
<td>(x_{226})</td>
<td>(x_{236})</td>
</tr>
</tbody>
</table>

**Fig. 9.** Data matrix for deciduous tree passage.
**Fig. 10. Data matrix for insect passage.**
Table 16
Means and Standard Deviations of Cloze Test Scores
on Different Levels of Amplification of the
Sedimentary Rock Passage by Levels of Interest

<table>
<thead>
<tr>
<th>Interest</th>
<th>n</th>
<th>low amplification mean</th>
<th>SD</th>
<th>Treatment mod amplification mean</th>
<th>SD</th>
<th>n</th>
<th>high amplification mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>7</td>
<td>41.961</td>
<td>13.840</td>
<td>5</td>
<td>41.150</td>
<td>15.181</td>
<td>4</td>
<td>54.282</td>
</tr>
<tr>
<td>high</td>
<td>5</td>
<td>64.998</td>
<td>7.462</td>
<td>5</td>
<td>48.072</td>
<td>11.129</td>
<td>8</td>
<td>49.639</td>
</tr>
</tbody>
</table>

Note.—Raw cloze scores are converted to percentage of total points possible.
Table 17
Analysis of Variance of Cloze Scores by
Amplification for the Sedimentary Rock
Passage and by Interest

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>1</td>
<td>638.456</td>
<td>638.456</td>
<td>5.600</td>
<td>.025</td>
</tr>
<tr>
<td>Amplification</td>
<td>2</td>
<td>328.474</td>
<td>164.237</td>
<td>1.441</td>
<td>.254</td>
</tr>
<tr>
<td>Inter X Amp</td>
<td>2</td>
<td>1081.794</td>
<td>540.897</td>
<td>4.744</td>
<td>.017</td>
</tr>
<tr>
<td>Error</td>
<td>28</td>
<td>3192.212</td>
<td>114.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>5240.945</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—For means and standard deviations, see Table 16.
Fig. 11. The interaction of interest and amplification in the process of comprehending elementary science reading material on sedimentary rocks.

By observation, the interaction was disordinal, and therefore, the main effects cannot be interpreted. To interpret the simple effects of amplification at levels of interest and the simple effects of interest at levels of
amplification, three one-way analyses of variances and two two-way analyses of variances were computed (Winer, 1971, p. 441). As shown in Table 18, a significant F was observed for interest at $B_1$ and $B_2$, thus permitting rejection of the null hypotheses of no differences among interest groups at low and moderate amplification on the sedimentary rock passage. The F for the amplification factor at high interest is significant, but examination of the means (Table 16) indicates the significance is in the opposite direction to that stated in the pilot study hypothesis. The F statistics for B at $A_1$ and A at $B_3$ did not achieve significance, and therefore the null hypotheses associated with them cannot be rejected. These results, of course, are somewhat biased. By computing five analyses of variances on the same set of data, the probability of achieving significance was increased by five times. Therefore, only the F ratio for the simple effects of interest at low amplification, which was significant at the .01 level, can be interpreted with confidence. The results of post hoc multiple comparison for the amplification factor at high interest, employing the Scheffe S Method are shown in Table 19. The mean of low amplification groups at high interest significantly surpassed the mean of moderate amplification at high interest on the sedimentary rock passage.
Table 18
Analysis of Variance of the Simple Effects of Amplification at Levels of Interest and Interest at Levels of Amplification for the Sedimentary Rock Passage

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Effects of Amplification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B at A₁</td>
<td>2</td>
<td>337.93</td>
<td>168.97</td>
<td>1.48</td>
<td>.25</td>
</tr>
<tr>
<td>B at A₂</td>
<td>2</td>
<td>765.90</td>
<td>382.95</td>
<td>3.36</td>
<td>.05</td>
</tr>
<tr>
<td>Simple Effects of Interest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A at B₁</td>
<td>1</td>
<td>1419.37</td>
<td>1419.37</td>
<td>12.45</td>
<td>.01</td>
</tr>
<tr>
<td>A at B₂</td>
<td>1</td>
<td>664.81</td>
<td>664.81</td>
<td>5.84</td>
<td>.05</td>
</tr>
<tr>
<td>A at B₃</td>
<td>1</td>
<td>3.49</td>
<td>3.49</td>
<td>.03</td>
<td>-</td>
</tr>
<tr>
<td>Error</td>
<td>28</td>
<td>3192.212</td>
<td>114.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 19
Post Hoc Multiple Comparisons for Amplification at High Interest on the Sedimentary Rocks Passage

Ψs = 16.77
If Ψi (pairwise or compound) ≥ 16.77, reject H0: Ψi = 0

<table>
<thead>
<tr>
<th>Ordered X1:</th>
<th>X2</th>
<th>X3</th>
<th>X1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>48.072</td>
<td>49.639</td>
<td>64.998</td>
</tr>
</tbody>
</table>

X2 < X1 at p < .05.

The means and standard deviations and F statistics for interest, amplification, and amplification X interest interaction on the deciduous tree passage cloze scores are shown in Tables 20 and 21. Significance was not achieved. Therefore, the null hypotheses associated with them cannot be rejected.

The means and standard deviations and F statistics for interest, amplification, and amplification X interest interaction on the insect passage cloze test scores are shown in Tables 22 and 23. Significance was not achieved. Therefore, the null hypotheses associated with them cannot be rejected.
Table 20
Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of the Deciduous Tree Passage by Levels of Interest

<table>
<thead>
<tr>
<th>Interest</th>
<th>low amp</th>
<th>mod amp</th>
<th>high amp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>mean</td>
<td>SD</td>
</tr>
<tr>
<td>low</td>
<td>5</td>
<td>50.553</td>
<td>12.897</td>
</tr>
<tr>
<td>high</td>
<td>6</td>
<td>62.775</td>
<td>8.800</td>
</tr>
</tbody>
</table>

Note.—Raw cloze scores are converted to percentage of total points possible.
Table 21
Analysis of Variance of Cloze Scores by Amplification for the Deciduous Tree Passage and by Interest

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>1</td>
<td>9.763</td>
<td>9.763</td>
<td>.068</td>
<td>-</td>
</tr>
<tr>
<td>Amplification</td>
<td>2</td>
<td>117.835</td>
<td>58.918</td>
<td>.409</td>
<td>-</td>
</tr>
<tr>
<td>Amp X Inter</td>
<td>2</td>
<td>605.780</td>
<td>300.890</td>
<td>2.103</td>
<td>.141</td>
</tr>
<tr>
<td>Error</td>
<td>28</td>
<td>4032.695</td>
<td>144.025</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>4766.73</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—For means and standard deviations, see Table 20.
Table 22
Means and Standard Deviations of Cloze Test Scores on Different Levels of Amplification of the Insect Passage by Levels of Interest

<table>
<thead>
<tr>
<th>Interest</th>
<th>low amp</th>
<th>Treatment</th>
<th>high amp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>mean</td>
<td>SD</td>
</tr>
<tr>
<td>low</td>
<td>6</td>
<td>46.77</td>
<td>13.02</td>
</tr>
<tr>
<td>high</td>
<td>6</td>
<td>51.07</td>
<td>9.87</td>
</tr>
</tbody>
</table>

Note.--Raw cloze scores are converted to percentage of total points possible.
Table 23

Analysis of Variance of Cloze Scores by Amplification for the Insect Passage by Levels of Interest

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>1</td>
<td>219.751</td>
<td>219.751</td>
<td>1.971</td>
<td>.172</td>
</tr>
<tr>
<td>Amplification</td>
<td>2</td>
<td>222.097</td>
<td>111.048</td>
<td>.996</td>
<td>.388</td>
</tr>
<tr>
<td>Amp X Inter</td>
<td>2</td>
<td>5.352</td>
<td>2.676</td>
<td>.024</td>
<td>-</td>
</tr>
<tr>
<td>Error</td>
<td>28</td>
<td>3010.455</td>
<td>111.498</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>3457.655</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—For means and standard deviations, see Table 22.
Because the factor of interest, amplification, did not achieve significance (except once in the opposite direction); but appeared to be in the hypothesized direction in the low interest groups in the rock and tree passages and in both interest groups in the insect passage, fourteen additional scores were added to the data. In the rocks passage, three scores were added to the low interest, low amplification cell; one score was added to the high interest, moderate amplification cell; and one score was added to the high interest, high amplification cell. In the tree passage, one score was added to the low interest, moderate amplification cell; two scores were added to the low interest, high amplification cell; one score was added to the high interest, moderate amplification cell; and one score was added to the high interest, high amplification cell. In the insect passage, one score was added to the low interest, low amplification cell; one score was added to the low interest, moderate amplification cell; one score was added to the low interest, high amplification cell; and one score was added to the high interest, moderate amplification cell. These were scores made by children who had the defined population characteristics, but who had not been randomly assigned to the treatment orders nor for whom had the treatment concepts been randomly ordered. A comparison of the results are shown in Tables 24, 25, and 26. In the comparisons shown in Tables 24, 25, and 26 the significance of interaction in the
Table 24: F Statistics for the Study Population and the Study Population Plus for the Sedimentary Rock Passage

<table>
<thead>
<tr>
<th></th>
<th>Study Pop</th>
<th></th>
<th>Study Pop +</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>P</td>
<td>F</td>
<td>P</td>
</tr>
<tr>
<td>Interest</td>
<td>5.600</td>
<td>.025</td>
<td>3.322</td>
<td>.077</td>
</tr>
<tr>
<td>Amplification</td>
<td>1.441</td>
<td>.254</td>
<td>2.712</td>
<td>.081</td>
</tr>
<tr>
<td>Amp X Inter</td>
<td>4.744</td>
<td>.017</td>
<td>3.224</td>
<td>.053</td>
</tr>
</tbody>
</table>

Table 25: F Statistics for the Study Population and the Study Population Plus for the Deciduous Tree Passage

<table>
<thead>
<tr>
<th></th>
<th>Study Pop</th>
<th></th>
<th>Study Pop +</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>P</td>
<td>F</td>
<td>P</td>
</tr>
<tr>
<td>Interest</td>
<td>.068</td>
<td>.796</td>
<td>.266</td>
<td>.609</td>
</tr>
<tr>
<td>Amplification</td>
<td>.409</td>
<td>.668</td>
<td>.648</td>
<td>.530</td>
</tr>
<tr>
<td>Inter X Amp</td>
<td>2.103</td>
<td>.141</td>
<td>2.429</td>
<td>.104</td>
</tr>
</tbody>
</table>
Table 26
F Statistics for the Study Population and the Study Population Plus for the Insect Passage

<table>
<thead>
<tr>
<th></th>
<th>Study Pop</th>
<th>Study Pop +</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>P</td>
</tr>
<tr>
<td>Interest</td>
<td>1.971</td>
<td>.172</td>
</tr>
<tr>
<td>Amplification</td>
<td>.996</td>
<td>.385</td>
</tr>
<tr>
<td>Inter X Amp</td>
<td>.024</td>
<td>-</td>
</tr>
</tbody>
</table>

Sedimentary rock passage F statistics decreased slightly but was still significant. In the deciduous tree passage F statistics, significance of interaction increased from .141 to .104. The significance of amplification in the insect passage F statistics increased from .385 to .179. The increase in significance of interaction in the deciduous tree passage F statistics is a reasonable increase due to increased n per cell. However, the increase in significance of amplification in the insect passage F statistics is too great to be attributed totally to increased n per cell. Apparently some other factors associated with the population + are causing the large increase in significance. The only factor which is apparent to the writer is the population +'s lack of random assignment of all Ss. Therefore, only the tree
passage data indicated an increased \( n \) per cell would increase the power of the design. However, it was decided to go ahead and use an \( n \) of fifteen in the formal study in an attempt to increase the overall power of the design and thereby achieve significance in the \( F \) ratios.

The hypothesis of increased comprehension due to amplification applied to all three concepts. The pilot, however, suggested that this may not be the case for the high interest groups on the sedimentary rock passage: amplification may actually cause decreases in comprehension scores. This would imply that amplification is detrimental tool with some people on some science topics. Therefore, in order to gain as much information as possible about amplification with different concepts, it was decided to retain all three concept passages. The results of the pilot study also suggest that interest is indeed correlated somewhat with the cloze score, and blocking on this variable was appropriate. To increase the efficiency of the design even further, it was decided to add a third level to the blocking variable. Feldt (1958) suggests that with the given \( n \), levels of the active variable, and probable correlation of the blocking variable to the dependent variable, power will not be decreased by making this addition. Interest, then will be divided into high, moderate, and low levels. The top third of the interest scores will be high interest, the middle third will be moderate interest, and the bottom third will
be low interest. The design will then be a three by three treatment by levels and will be handled similarly to the two by three design previously considered. In order to retain 15 Ss per cell, it will be necessary to increase the population from the original suggestion of 90 to 135.

Reliability estimates of the nine cloze test were determined using the Kuder Richardson 20 formula for the low amplification passages and the Spearman Brown formula for the moderate and high amplification passages. The results are shown in Table 27.

The lower reliability of the low amplification cloze tests was expected as explained earlier. However, the obtained reliabilities for these low amplification passage cloze tests were considered acceptable. The reliability of the high amplification passage cloze test for sedimentary rocks, .37, was not considered acceptable. Since the reliability estimates of the low and moderate amplification passage cloze tests were acceptable, it was decided to examine the last two exemplars of the high amplification passage for possible ambiguity. These two exemplars were not included in the other two passages and were considered a possible source of the unreliability of the passage. After examination, it was decided that the addition of two contextual clues and two changes in grammatical structure would make the two exemplars less ambiguous and hopefully increase
Table 27
Reliability Estimates of Cloze Tests

<table>
<thead>
<tr>
<th>Passage Type</th>
<th>Low Amplification</th>
<th>Moderate Amplification</th>
<th>High Amplification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedimentary Rock Passage</td>
<td></td>
<td></td>
<td>.69</td>
</tr>
<tr>
<td>Moderate Amplification</td>
<td></td>
<td></td>
<td>.88</td>
</tr>
<tr>
<td>High Amplification</td>
<td></td>
<td></td>
<td>.37</td>
</tr>
<tr>
<td>Deciduous Tree Passage</td>
<td></td>
<td></td>
<td>.67</td>
</tr>
<tr>
<td>Moderate Amplification</td>
<td></td>
<td></td>
<td>.90</td>
</tr>
<tr>
<td>High Amplification</td>
<td></td>
<td></td>
<td>.86</td>
</tr>
<tr>
<td>Insect Passage</td>
<td></td>
<td></td>
<td>.76</td>
</tr>
<tr>
<td>Moderate Amplification</td>
<td></td>
<td></td>
<td>.73</td>
</tr>
<tr>
<td>High Amplification</td>
<td></td>
<td></td>
<td>.92</td>
</tr>
</tbody>
</table>

*aInternal consistency coefficients.

the reliability of the whole passage cloze test (see the revised copy on page 152).

As a result of the pilot study, it was decided that the entire set of three cloze test should not be given in one sitting. It took from 50 to 60 minutes for the pilot
population to complete the set, and fatigue and restlessness were evident in many children after 35 to 45 minutes. In the major study, the cloze tests will be given on three consecutive days at the same time each day, in 30 minute segments. This, of course, will make the entire time of treatment and test administration take longer thus increasing the chances of subject mortality and intrasession history effects. However, it was decided that the effects of fatigue may be greater than intrasession history. And if necessary, some subjects who do not complete all the tests can be deleted. Also from the pilot, it was determined that the directions for administering the test should be modified slightly to make them more clear.

As a corollary to the pilot study, an informal study to answer the following question was conducted. Will those students who have more concrete props associated with exemplars of a concept in their cognitive structure score higher on the cloze tests than those students who have fewer? To answer the question, an objective test on each of the three concepts was constructed (see Appendix A). While the rest of each class completed the cloze tests, the eight remaining members of the population in each class completed the objective test. Then the mean score for each subtest in each class was computed. The means of the cloze test scores of each concept passage for the other six members of the population in each class were also computed. The two sets of
scores were used to make Tables 28, 29 and 30. Fisher's Exact Probability Test was used to determine the probability that the results in the tables were due to chance. The probability of the results being due to chance on the Insect Concept subtest was .20, on the Sedimentary Rock Concept subtest was .45 and on the Deciduous Tree Concept subtest was .20 (but in the opposite direction). The insect and sedimentary rock concept results, while not achieving significance, may indicate a trend which would answer the question positively. However, if the trend exists, it is evident only to the extent indicated by the probabilities of .20 and .45. Therefore, on the basis of the results of this test, the conclusion that students who have more concrete props associated with exemplars in their cognitive structure will achieve higher cloze test scores cannot be made.

A month later, a second study attempting to answer the same corollary question was conducted. This time an interview technique, designed to determine the extent of the presence of concrete props associated with the exemplars of the three concepts was used (see Appendix A). Each of the 36 subjects who had taken the cloze tests were interviewed by the researcher. At the principal's request the interview took place in each subject's classroom. Each interview was given a rating score. The interview rating for each subject in each concept was correlated with the cloze test scores
Table 28

Two by Two Table for Sedimentary Rock Passage

<table>
<thead>
<tr>
<th>Cloze test scores</th>
<th>Below mean</th>
<th>Above mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below mean</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Objective test scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above mean</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 29

Two by Two Table for Deciduous Tree Passage

<table>
<thead>
<tr>
<th>Cloze test scores</th>
<th>Below mean</th>
<th>Above mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below mean</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Objective test scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above mean</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 30

Two by Two Table for Insects Passage

<table>
<thead>
<tr>
<th>Cloze test scores</th>
<th>Below mean</th>
<th>Above mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below mean</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Objective test scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above mean</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>
for each concept for each subject. Pearson Product Moment Correlation Coefficients were computed to determine the relationship between the two sets of scores using the Basic (P-CORR,STAT) program. The results are shown in Table 31. Only three of the nine correlations were positive and significant. Therefore the conclusion that students who have more concrete props associated with exemplars of a concept in their cognitive structure will achieve higher cloze scores cannot be made.

As a result of the above two corollary studies to the pilot, it was decided that the assumption, that children can learn verbally presented elementary science concepts if they have the appropriate concrete-empirical props in their cognitive structures, may not be justified. Therefore to prevent misinterpreting Ausubel's theory as it applies to the main study, it was decided to use the interview technique as a corollary to the main study. The decision to use the interview technique instead of the objective test was based on the researcher's subjective opinion that the interview more clearly determined what the children knew about the exemplars in question.

Since the reading ability range of the population for the pilot had been increased from 4.0 through 6.0 grade levels to anything above 4.0 in order to get a larger number of subjects, it was decided to check the correlations between the reading ability scores and the cloze test scores.
Table 31  
Correlations between Cloze Scores  
and Interview Ratings

<table>
<thead>
<tr>
<th>Passage</th>
<th>Low Amplification</th>
<th>Moderate Amplification</th>
<th>High Amplification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedimentary Rock Passage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.56(^{a})</td>
<td>.78(^{b})</td>
<td>-.05</td>
</tr>
<tr>
<td>Deciduous Tree Passage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-.07</td>
<td></td>
<td>.30</td>
</tr>
<tr>
<td>Insect Passage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-.11</td>
<td>.55(^{a})</td>
<td>.18</td>
</tr>
</tbody>
</table>

\(^{a}\)significant at .05 for one tailed test.  
\(^{b}\)significant at .005 for one tailed test.
To do this, Pearson Product Moment Correlation Coefficients were computed using the Basic (P-CORR,STAT) program. The results are shown in Table 32. All nine correlations were positive and six were significant. Therefore it was decided that in the main study, a tighter control on the reading ability range should be used. Instead of using a range of 4.0 through 6.0 grade levels, a range of 4.0 through 5.0 grade levels will be used.
Table 32
Correlations between Cloze Scores and Reading Level Scores

<table>
<thead>
<tr>
<th>Passage</th>
<th>Low Amplification</th>
<th>Moderate Amplification</th>
<th>High Amplification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedimentary Rocks</td>
<td>.74&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.73&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.54&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Deciduous Tree</td>
<td>.28</td>
<td>.72&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.73&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Insect</td>
<td>.38</td>
<td>.42</td>
<td>.59&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>significant at .05 for one tailed test.
<sup>b</sup>significant at .001 for one tailed test.
FOOTNOTES

1Languis, M. L. Personal communication. 1971.
BIBLIOGRAPHY


COPES Newsletter, 1967, 1, (1).


Hill, W. Content textbook: help or hindrance. J. Read., 1967, 10, 408 - 413.


Merriman, H. O. The Columbus school profile. Columbus, O.: Columbus Public Schools, 1971.


