1986

Factors Affecting Summary Writing and Their Impact on Reading Comprehension Assessment.

Martha Holcomb Head

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation


https://digitalcommons.lsu.edu/gradschool_disstheses/4239

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This reproduction was made from a copy of a manuscript sent to us for publication and microfilming. While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. Pages in any manuscript may have indistinct print. In all cases the best available copy has been filmed.

The following explanation of techniques is provided to help clarify notations which may appear on this reproduction.

1. Manuscripts may not always be complete. When it is not possible to obtain missing pages, a note appears to indicate this.

2. When copyrighted materials are removed from the manuscript, a note appears to indicate this.

3. Oversize materials (maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or in black and white paper format.*

4. Most photographs reproduce acceptably on positive microfilm or microfiche but lack clarity on xerographic copies made from the microfilm. For an additional charge, all photographs are available in black and white standard 35mm slide format.*

*For more information about black and white slides or enlarged paper reproductions, please contact the Dissertations Customer Services Department.

UMI Dissertation Information Service
University Microfilms International
A Bell & Howell Information Company
300 N. Zeeb Road, Ann Arbor, Michigan 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Head, Martha Holcomb

FACTORS AFFECTING SUMMARY WRITING AND THEIR IMPACT ON READING COMPREHENSION ASSESSMENT

The Louisiana State University and Agricultural and Mechanical Col. PH.D. 1986

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. 
Problems encountered with this document have been identified here with a check mark □.

1. Glossy photographs or pages □
2. Colored illustrations, paper or print □
3. Photographs with dark background □
4. Illustrations are poor copy □
5. Pages with black marks, not original copy □
6. Print shows through as there is text on both sides of page □
7. Indistinct, broken or small print on several pages □
8. Print exceeds margin requirements □
9. Tightly bound copy with print lost in spine □
10. Computer printout pages with indistinct print □
11. Page(s) ________ lacking when material received, and not available from school or author.
12. Page(s) ________ seem to be missing in numbering only as text follows.
13. Two pages numbered ______. Text follows.
14. Curling and wrinkled pages □
15. Dissertation contains pages with print at a slant, filmed as received □
16. Other ________________________________________________________________

__________________________________________

University
Microfilms
International
FACTORS AFFECTING SUMMARY WRITING AND THEIR IMPACT ON READING COMPREHENSION ASSESSMENT

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Education

by

Martha H. Head
B.S., Louisiana State University, 1971
M.A., Louisiana State University, 1982
May 1986
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT .................................................. v</td>
</tr>
<tr>
<td>CHAPTER</td>
</tr>
<tr>
<td>1 INTRODUCTION ................................ 1</td>
</tr>
<tr>
<td>Review of Related Literature ........... 3</td>
</tr>
<tr>
<td>Need for the Study ..................... 5</td>
</tr>
<tr>
<td>Free Recalls and Summaries ........... 5</td>
</tr>
<tr>
<td>Prior Knowledge ....................... 6</td>
</tr>
<tr>
<td>Topic Interest ....................... 9</td>
</tr>
<tr>
<td>Writing Ability ....................... 10</td>
</tr>
<tr>
<td>Summaries as a Measure of Reading Comprehension ..................... 11</td>
</tr>
<tr>
<td>2 METHOD .........................................16</td>
</tr>
<tr>
<td>Subjects .....................................16</td>
</tr>
<tr>
<td>Materials .....................................17</td>
</tr>
<tr>
<td>Target Passage .......................... 17</td>
</tr>
<tr>
<td>Multiple Choice Tests ................... 18</td>
</tr>
<tr>
<td>Interest Inventory ..................... 17</td>
</tr>
<tr>
<td>Writing Sample .......................... 19</td>
</tr>
<tr>
<td>Written Recalls .......................... 19</td>
</tr>
<tr>
<td>Procedures ................................. 20</td>
</tr>
<tr>
<td>Data Collection .......................... 20</td>
</tr>
<tr>
<td>Day One .................................... 20</td>
</tr>
<tr>
<td>Day Two .................................... 21</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day Three</td>
<td>21</td>
</tr>
<tr>
<td>Day Four</td>
<td>21</td>
</tr>
<tr>
<td>Scoring</td>
<td>21</td>
</tr>
<tr>
<td>Interest Inventory</td>
<td>21</td>
</tr>
<tr>
<td>Writing Sample</td>
<td>21</td>
</tr>
<tr>
<td>Multiple Choice Tests</td>
<td>23</td>
</tr>
<tr>
<td>Written Recalls</td>
<td>23</td>
</tr>
<tr>
<td>RESULTS</td>
<td>29</td>
</tr>
<tr>
<td>MANOVA</td>
<td>29</td>
</tr>
<tr>
<td>Words</td>
<td>30</td>
</tr>
<tr>
<td>Idea Units</td>
<td>32</td>
</tr>
<tr>
<td>Number of Important Idea Units</td>
<td>32</td>
</tr>
<tr>
<td>Proportion of Important Idea Units</td>
<td>33</td>
</tr>
<tr>
<td>Efficiency</td>
<td>34</td>
</tr>
<tr>
<td>Rule Use</td>
<td>35</td>
</tr>
<tr>
<td>Regression Analyses</td>
<td>36</td>
</tr>
<tr>
<td>Words</td>
<td>39</td>
</tr>
<tr>
<td>Idea Units</td>
<td>39</td>
</tr>
<tr>
<td>Important Idea Units</td>
<td>42</td>
</tr>
<tr>
<td>Proportion of Important Idea Units</td>
<td>42</td>
</tr>
<tr>
<td>Efficiency</td>
<td>42</td>
</tr>
<tr>
<td>Rule Use</td>
<td>42</td>
</tr>
<tr>
<td>Correlational Analyses</td>
<td>43</td>
</tr>
<tr>
<td>Words</td>
<td>43</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idea Units</td>
<td>43</td>
</tr>
<tr>
<td>Important Idea Units</td>
<td>43</td>
</tr>
<tr>
<td>Proportion of Important Idea Units</td>
<td>45</td>
</tr>
<tr>
<td>Efficiency</td>
<td>45</td>
</tr>
<tr>
<td>Rule Use</td>
<td>45</td>
</tr>
<tr>
<td>4 DISCUSSION</td>
<td>46</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>54</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>61</td>
</tr>
<tr>
<td>A. Definition of Terms</td>
<td>61</td>
</tr>
<tr>
<td>B. Review of the Literature</td>
<td>64</td>
</tr>
<tr>
<td>C. Pilot Study</td>
<td>135</td>
</tr>
<tr>
<td>D. Multiple Choice Instruments</td>
<td>138</td>
</tr>
<tr>
<td>E. Interest Inventory</td>
<td>145</td>
</tr>
<tr>
<td>F. Free Recall and Summary Instruments</td>
<td>147</td>
</tr>
<tr>
<td>G. Data Collection Scripts</td>
<td>150</td>
</tr>
<tr>
<td>H. Letter to Parents</td>
<td>158</td>
</tr>
<tr>
<td>I. Raw Data</td>
<td>160</td>
</tr>
<tr>
<td>VITA</td>
<td>165</td>
</tr>
</tbody>
</table>
ABSTRACT

This study examined several factors that may be involved in the writing of summaries and the impact of these factors on the use of summaries for reading comprehension assessment. Fifty-three seventh-grade subjects of average to above-average reading ability participated in the study. Data collection took place in four sessions over a one-week duration. Levels of prior knowledge and interest in a topic, as well as general writing ability, were assessed before subjects read the target passage. Subjects read a novel passage consisting of a five-page section of a chapter in their regularly-assigned American history text. After reading, subjects wrote, in counterbalanced order, both a free recall and a summary of the passage. After a second reading, subjects answered 15 passage-dependent, multiple-choice questions. In addition, recent scores on a standardized test of reading comprehension were obtained for each subject.

Data from the free recalls and summaries were analyzed in three ways. First, a multivariate analysis of variance, with repeated measures on free recalls and summaries, was performed. Next, a standard regression analysis was undertaken, using scores from the prior knowledge, topic interest, and writing ability measures as predictor variables for the free recall and summary.
scores. Finally, Pearson product-moment correlations coefficients were calculated for both free recalls and summaries with multiple-choice questions scores and scores on the standardized reading comprehension test.

MANOVA results revealed a significant main effect for Task, and follow-up univariate tests indicated significant differences between free recalls and summaries for number of idea units and proportion of important idea units included. The regression analysis revealed that prior knowledge, topic interest, and writing ability may play at least a partial role in the writing of summaries. Finally, correlation analyses indicated that summaries measure at least some of the same aspects of reading comprehension as multiple-choice questions and standardized tests. Implications for instruction and directions for further research are cited.
CHAPTER 1

INTRODUCTION

The assessment of reading comprehension has been approached in a variety of ways, in part because of the difficulty researchers and practitioners have found in operationally defining comprehension (Farr, 1969; Johnston, 1983). In recent years there has been a growing tendency to assess comprehension of a passage by free recall, sometimes as the only dependent measure employed and sometimes as a part of a battery of measures (e.g., Freebody & Anderson, 1983a, 1983b; Smith, 1985; Taylor, 1984; Thomas & Bridge, 1980). One recent study, in fact, has pointed to recall as a more valid measure of postreading comprehension than short-answer questions, with summarized recall found superior to free recall (Taylor, 1984). Taylor explained these results in terms of Kintsch and van Dijk's (1978) notion of the gist that a reader forms while reading a passage. She reasoned that a summary more closely represents this gist than does either a free recall or short-answer question measure. In all three types of measures, however, it seems the reader must have understood the passage in order to produce a free recall, summarize, or answer questions about it.

From this line of reasoning, it might be concluded that readers who produce complete, well-written summaries have comprehended well. But what about readers whose
summaries are not good - either those that contain little or erroneous information or those that are poorly-organized? Can it be concluded that poor summaries are the result of faulty comprehension?

Several other variables, however, could be postulated to affect summary production. For instance, do the readers understand the task, that is, has summary writing been used frequently in the classroom with careful attention given to what good summaries ought to include and how to write them? Second, how familiar are the readers with the topic of the passage to be read and summarized? Could a lack of prior knowledge affect the quality of a summary produced by the reader? Third, are the readers interested enough in the topic to care about producing a good summary, or is the writing merely a task to be performed and completed? Fourth, are the readers also accomplished writers, who can plan carefully what is to be included in a summary and then organize it in a logical, coherent manner? Finally, how adequate are summaries as a measure of comprehension as compared to other measures such as free recall, multiple choice questions, and standardized reading comprehension tests?

These concerns suggest that some factors other than quality of comprehension must be involved in the quality of summaries that are produced after reading. Therefore, this study will investigate the following areas: (a)
similarities and differences between written free recalls and summaries; (b) the influence that prior knowledge, writing ability, and topic interest might have on free recalls and summaries written after reading; and, (c) similarities of summaries with other measures of comprehension, such as multiple-choice questions and standardized tests of reading comprehension. Terms associated with these concerns and others which are relevant to the study are defined in Appendix A.

Review of Related Literature

This section summarizes some of the previous studies which are pertinent to the present study. For a more complete review of the literature, see Appendix B.

The writing of summaries is a technique sometimes employed by classroom teachers to see how well their students have understood and remembered what they read in a textbook passage. Thus, summaries can be used to provide teachers with an indicator of what Rosenblatt (1978) has termed "efferent reading", or what remains after reading. Kintsch and van Dijk (1978) have postulated that readers formulate a gist of a passage while reading, and numerous researchers (e.g., Garner, 1985; Johnson, 1983; Taylor, 1984; Thomas & Bridge, 1980; Winograd, 1984) have implied that a summary written after reading reflects that gist. Research related to schema
theory (e.g., Rumelhart, 1980; Rumelhart & Norman, 1978) has supported this conclusion because readers, in order to fully comprehend a passage, must integrate the new knowledge the passage presents with their prior knowledge of the topic to form a new schema. Therefore, when a summary is written, it will contain both new and old information, reorganized and presented as a summary of the passage (Rumelhart, 1980) but representative of the reader's current schemata related to the topic.

The wide acceptance of a schema-theoretic view of reading has contributed to the use of written recalls to assess comprehension, not only in the classroom, but in various kinds of research studies. Taylor (1984), in a review of 15 recent studies which investigated comprehension, found that free recall was used as a comprehension measure in seven of the studies. More specifically, several studies have been identified which made use of written summaries in their methodology. For example, in their experiments relating vocabulary difficulty to reading comprehension, Freebody and Anderson (1983a, 1983b) had subjects write both a free recall and a short summary of the main ideas of the target passage. Similarly, Thomas and Bridge (1980) correlated summary scores with cloze comprehension scores and found a strong positive relationship, $r = .80$, thus lending credence to the validity of summaries as a comprehension measure.
Additionally, Smith (1985) investigated strategies and behaviors that proficient readers use while comprehending a difficult text and employed written summaries in an effort to reveal subjects' actual strategy use.

These four studies are representative of the growing trend of using summaries in comprehension research. This more widespread use of summaries, however, leads to the concerns listed previously, which will be more fully investigated in the present study.

**Need for the Study**

**Free recalls and summaries.** To date, only two studies have investigated children's ability to summarize as compared to their ability to tell everything they remember from reading a passage. Johnson (1983) directed children in grades one, three, and five, as well as adults, to listen to a story and then give orally both a free recall and a summary. While developmental differences were apparent, no significant differences were noted between the children's free recalls and summaries. Johnson concluded that, although the children made an attempt at summarization, their skills were still developing.

Taylor (1984), using expository text, had seventh-grade subjects write a summary after reading. The summaries were then scored both as a free recall (total
number of idea units recalled from the passage) and as a summary (taking into account the relative importance of the idea units recalled). In a qualitative comparison of the two sets of scores, Taylor concluded that the summary scoring provided more information regarding the subjects' understanding of the passage than did the free recall and, thus, was a superior measure of comprehension. However, this conclusion is suspect since the same written protocol was used to obtain scores for both a free recall and a summary.

The findings of these two studies make apparent the need for further research to clarify how free recalls and summaries are related. The present study will compare separate measures of free recall and summarizing in order to investigate this area of concern.

Prior knowledge. In light of schema theory, it would be difficult to read a summary written by a student and then conclude that all the information it contained came from reading the target passage. This is a problem prevalent in much reading comprehension research (Farr, 1969; Johnston, 1983). For example, it cannot be concluded that answering questions about a passage is dependent only upon information obtained from reading the passage itself. Preston (1962) found that 77% of subjects studied were able to correctly answer questions on a reading comprehension test even before reading the target
passage. Marks and Noll (1967) have suggested that when such results occur, it is highly likely that the items are measuring something other than reading comprehension; Farr (1969) concluded that, in fact, prior knowledge is what is being measured. Tuinman (1973-74) concurred, pointing out the limitations of questions on standardized tests for measuring comprehension; they may, in fact, be measuring prior knowledge. Johnston (1983) has even recommended purposely inserting some passage-independent questions into an instrument for the purpose of assessing the background knowledge that the reader is believed to possess.

Similar phenomena may occur when a summary is written. Rumelhart (1980) has commented that, once the reader has obtained meaning from a passage, we may not be able to differentiate between information obtained from the sensory input of the text and the information that is a product of the reader's interpretation based on prior knowledge. Such an integration of old and new is undoubtedly taking place when readers reconstruct the text in summary form. The assumption is made here, however, that the reader possesses the appropriate background knowledge to allow for a consistent interpretation of the target passage. This, in fact, may not be the case, and the lack of appropriate schemata has been shown to have a detrimental effect on comprehension (Bransford & Johnson,
1972; Rumelhart, 1980). Some readers, therefore, may not possess the relevant prior knowledge needed to understand the target passage, and their summaries reflect this lack of knowledge.

Not only could the appropriate schemata be lacking, but it may be that the author did not provide sufficient cues to the reader to activate that knowledge (Rumelhart, 1980). The well-known "Washing Clothes" text employed in a study by Bransford and Johnson (1972) presents a useful example. Most adult readers have all the background knowledge necessary for interpreting the passage, but the wording is so vague that readers fail to instantiate, or call forth, the appropriate schemata and thus are unable to understand the text.

Additionally, readers may possess schemata for the topic that are either inaccurate or incomplete and may use this information to interpret the passage in a way that is consistent for them but fails to grasp the author's intended message (Rumelhart, 1980). Such was the case in studies by Lipson (1983) and Alvermann, Smith, and Readence (1985) which found that readers' prior knowledge could have an inhibitive effect and could, in fact, interfere with their understanding of the text.

For the purposes of comprehension research in which a written summary is the dependent measure, it appears that the effect of readers' background knowledge should be
taken into account. This study will attempt to accomplish this by assessing the readers' prior knowledge of the target passage before reading takes place so that any such effect may be considered in a statistical analysis of summary scores.

**Topic interest.** It seems intuitive that students will understand better a topic that they like to read about, and the findings of several studies have supported this prediction (e.g., Baldwin, Peleg-Bruckner, & McClintock, 1985). Belloni and Jongsma (1978) found a facilitative effect on the comprehension of reluctant readers when they read stories about a topic for which they expressed interest. A later study by Asher (1980) had subjects use a picture-rating scale to indicate high or low interest in several topics. When subjects read stories related to the topics of pictures in which they had indicated high interest, their comprehension was higher than for the low-interest topics.

An important criticism of some of these studies concerns the problem of separating topic interest from prior knowledge. The question that arises concerns whether readers comprehend high-interest material better because of their desire to read about it or because they have greater prior knowledge about topics for which they express higher interest. Guthrie (1981), in a review of studies on interest and reading comprehension, concluded
that children probably comprehend high-interest material better simply because they know more about it. His assumption is that prior knowledge is highly correlated with interest and that prior knowledge is the causal factor. Pearson and Johnson (1978), however, have suggested that interest itself is a factor in reading comprehension.

A more recent study by Baldwin, Peleg-Bruckner, and McClintock (1985) sought to tease out these effects. Their findings indicated that both prior knowledge and interest were independent factors in reading comprehension and that, in fact, the two variables were virtually uncorrelated, $r = .09$. It seems to follow from these findings that when subjects indicate their comprehension of a passage by writing a summary about it, their interest in the topic will have an effect on the quality of the summary that they write. Thus, this study will examine the effect of topic interest as a predictor of summary writing ability.

**Writing ability.** Because answering questions and writing summaries place very different demands on readers (Johnston, 1983), it would seem illogical to qualitatively compare performance on these two very different types of measures and conclude that poor summary writers have faulty comprehension. It may be, in fact, that these subjects were simply poor writers. Research has shown
that, in order to become good writers, students need many opportunities to write (Tierney & Leys, 1984). Unfortunately, this is not often the case in many schools (Anderson, Hiebert, Scott, & Wilkinson, 1985). The fact remains that while students can learn to write when given time and instruction in developing this skill (Calkins, 1983; Graves, 1982), such conditions simply do not exist.

Johnston (1983) also pointed out that expressing ideas in writing and organizing information from memory are skills that are not normally taught to children and that the nature of such skills may be very different from reading comprehension. Students for whom writing is difficult may have very different criteria for deciding what is important to write and would probably write less information than would skillful writers. In addition, poor writers often have difficulty assuming the perspective of their future readers and, thus, may fail to include well-understood information that they believe is so "obvious" as to be unimportant. It may be that writing a well-organized summary presents such a difficult task that only a fraction of what readers obtained from the reading is actually being sampled. In order to make reasonable inferences about reading comprehension from a written summary, it seems necessary to assess the students' general writing ability.

**Summaries as a measure of reading comprehension.**
Studies which have used summaries as a comprehension measure have done so at various age levels, from grade one to graduate level. They have also used directions ranging in specificity from directing subjects simply to read a passage and write a short summary (e.g., Freebody & Anderson, 1983a, 1983b) to defining the term for the subjects and showing them an example (e.g., Taylor, 1984). The tacit assumption of these (and other) researchers is that the subjects know how to write a summary. There is a great deal of evidence to the contrary.

For example, Otto, Barrett, and Koenke (1969) found that only 29% of second graders could select the main idea from simple passages. Danner (1976), in comparing second and sixth graders, found that only sixth graders could select appropriate topic sentences for passages from a list of sentences provided. Brown and Smiley (1977) found a similar developmental effect for supplying topic sentences; younger students (ages 8, 10, and 12) could recognize topic sentences, but only older students (age 18) were able to supply their own. Similarly, Garner, Belcher, Winfield, and Smith (1985) found that fifth graders could recognize good summaries but were either unable to produce their own or to tell how they would write one. Johnson (1983) compared the summaries produced by children in first, third, and fifth grade with those produced by adults and found a strong developmental trend,
with adults being far more efficient at summarizing than the children. In addition, free recalls also given by the children did not differ appreciably from their summaries. A subsequent study by McGee and Tompkins (1982) reported similar findings and concluded that children perform less well because of differing language abilities, differing amounts of world knowledge, and lack of experience in manipulating texts. Winograd (1984) found that while eighth graders understand that main ideas should be included in a summary, these subjects in general, and poor readers in particular, lacked the strategies necessary for effective rule use when writing their summaries. As a result of subsequent research on strategy use, Garner (1985) recommended that rule-based instruction in summary writing be an integral part of the high-school curriculum.

The finding that summary-writing ability is late developing has prompted several training studies (e.g., Day, 1980; Taylor & Beach, 1984) in an attempt to teach summarizing strategies where they were lacking. These attempts have met with reasonable success and lend weight to the notion that the development of effective summary-writing ability should not be left to chance; it should be explicitly taught.

If students are unable to write effective summaries without being trained to do so, then it seems questionable to assess their reading comprehension by asking them to
summarize information. A portion of this study will attempt to examine the summarizing strategies that readers use in an effort to show that deficiencies in this area should lead to cautions when interpreting summaries as measures of their comprehension.

To date, only two studies have explored the relationship of summary scores to other widely-used measures of reading comprehension. In their study, Thomas and Bridge (1980) correlated summary scores with comprehension as measured by a cloze test and found a strong positive relationship, $r = .80$. Winograd (1984) examined the roles of effective summary rule use and sensitivity to importance as predictors of both a standardized reading comprehension test score and postreading comprehension questions. In neither case was rule use a significant predictor of the comprehension scores. However, sensitivity to importance accounted for a significant proportion of the variance in each of the comprehension measures.

Because both Thomas and Bridge and Winograd found a modest relationship between summary scores and other measures of comprehension, it would seem useful to examine more closely the relationship of summary scores to instruments that are typical of classroom comprehension measures, namely, multiple-choice questions and standardized reading comprehension tests. This study will
seek to gain information in these areas.

The following questions, therefore, will guide this study:

1. Is there a difference in the ability to write a free recall of a passage and the ability to summarize it?

2. Does the level of prior knowledge of the topic of a target passage influence the ability to write a summary of that passage?

3. Does interest in the topic of a target passage influence the ability to write a summary of that passage?

4. Does general writing ability influence the ability to write a summary of a target passage?

5. Is there a relationship between the ability to answer postquestions on a passage and the ability to write a free recall and a summary?

6. Is there a relationship between a reading comprehension score on a standardized test and the ability to write a free recall and a summary?
CHAPTER 2

METHOD

This chapter describes the subjects involved in the study, the materials that were used, and the procedures employed for data collection and analysis.

Subjects

The subjects were drawn from a large magnet middle school in a large southern city. Because the school is a magnet school, admission requirements stipulate that all students have a stanine of five or greater on the reading portion of the standardized achievement test used by the local district. Students whose reading stanines are in the range of five to seven are generally grouped into regular sections of their academic subjects, while those with stanines of eight or nine are placed into advanced sections. The subjects in this study were deliberately chosen from three regular sections of seventh-grade social studies so that the majority of them would be average or slightly above average in reading ability. This was confirmed by obtaining subjects' recent scores on the Comprehensive Assessment Program (Wick & Smith, 1980). Subjects' percentile rank scores on the standardized test ranged from 43 (stanine 5) to 97 (stanine 9), with a mean of 73.62 (stanine 6) and a standard deviation of 16.46.

The three classes consisted of 82 students, but
because of absences during the data collection period or missing test score data, the final sample for whom data analysis was undertaken was reduced to 53. Of these subjects, 30, or 57%, were female and 23, or 43%, were male. In addition, approximately 68% of the subjects were white, and approximately 32% were classified as black or another minority group.

**Materials**

Pilot testing was conducted in order to develop the materials and instruments used in the study. An additional purpose of the pilot studies was to field test the procedures for data collection. For a description of the pilot studies, see Appendix C.

**Target Passage**

The target passage was selected from the subjects' regular social studies textbook and was chosen so that it was novel to them. An additional consideration was that the topic of the passage would produce wide variability among subjects both in interest level and in prior knowledge. Such variability would be necessary for the variable to be a viable predictor of the summary scores. Pilot studies revealed the passage chosen to be adequate for these purposes.

The passage was entitled "What Have Labor and Business Done to Solve the Problems of the Machine Age?". The five running pages of text consisted of three main
sections which were labeled by headings and subdivided by boldface subheadings. The passage was 1,845 words in length and was written at a readability level of grade seven (Fry, 1977).

Multiple Choice Tests

In order to measure the subjects' knowledge of the target passage both before and after reading, a pool of 30 multiple-choice questions was developed. Following procedures outlined by Johnston (1984), the items were constructed by identifying essential vocabulary terms from the target passage and formulating multiple-choice questions based on these terms. Pilot testing was conducted to ensure that the questions were passage-dependent. Following an item analysis, the items were ranked in order of difficulty. Stratified random assignment was then used to assign 15 items to each of the prereading test and the postreading test. No item appeared on both tests. Samples of the two instruments appear in Appendix D.

Interest Inventory

To measure the subjects' level of interest in the topic of the target passage, an interest inventory was developed, adapted from an example offered by Readence, Bean, and Baldwin (1985). The inventory consisted of a list of 29 topics in American history and blanks for subjects to use in indicating their responses to the
topics. A sample of the instrument appears in Appendix E.

Subjects were asked to rate, on a scale of A to F (five levels, corresponding to school grades of A, B, C, D, and F), how much they would like to read about each of the topics. An A grade indicated a high level of interest, while an F grade indicated a low interest level. Pilot testing had indicated that subjects expressed a wide range of ratings for each of the 29 topics, but of particular relevance were the two items that corresponded to the topic of the target passage: the growth of industry and the development of labor unions in the United States. A suitable range of variability was obtained with the pilot sample for these items.

**Writing Sample**

General writing ability was assessed by having subjects write a short essay entitled "McKinley Middle School through the Eyes of a Seventh Grader". Pilot testing had shown this topic to be popular with the subjects, and none revealed any difficulty in carrying out the task. Essays were written on lined white paper provided by the researcher.

**Written Recalls**

After having read the target passage, subjects wrote from memory both a free recall and a summary. A sheet with written directions was provided to the subjects for writing each of their recalls. These directions
emphasized the difference in the nature of the two tasks in that a free recall includes everything the reader can remember from the passage read, while a summary should include only the main ideas and the details necessary to support them. A sample of each of the sheets appears in Appendix F.

**Procedures**

**Data Collection**

Data was collected by the investigator and two trained doctoral students during the subjects' regular social studies class periods. One researcher was randomly assigned to each class and remained with that group throughout the data collection period. All directions given to the subjects were scripted to maintain uniformity in the presentation. All directions to complete the data collection were read orally to the subjects to ensure that they understood what was asked of them. Copies of the scripts are presented in Appendix G.

**Day one.** The researchers introduced themselves to the subjects and briefly explained the nature of the study. Subjects were given a letter for their parents, explaining the study and seeking their cooperation. The parent letter is included as Appendix H. The final activity of the session was completion of the interest inventory. Subjects were able to complete the task, and no apparent difficulty was noted.
Day two. Subjects first answered questions on the multiple-choice prereading test. It was explained that although some of the information might be somewhat unfamiliar, they were to answer the questions using whatever they could remember from some other time. Additionally, they wrote the essay for the topic supplied by the researcher. Subjects were urged to do their best writing on the essay, much as they would do for a graded essay in an English class.

Day three. This session took place after a five-day buffer. First, subjects read the target passage. Then they completed one of two recall tasks from memory: either a free recall or a summary. When the first recall task was completed, the other one was undertaken. Task order was counterbalanced so that half the subjects were randomly-assigned to write the free recall first and the summary second. The other half of the subjects wrote the summary first, followed by the written free recall.

Day four. Subjects reread the target passage and answered the postreading test questions from memory. Following the completion of the task by all subjects within the group, a short discussion period took place in which subjects were allowed to ask questions about the study and the purpose of each of the tasks.

Scoring

Interest inventory. Subjects' responses on the
interest inventory (indicated as A through F) were converted to numerical values where a rating of A was assigned a score of 4 and a rating of F received a score of 0. Scores between the two extremes were ordered accordingly. Scores were tabulated, and means and standard deviations for each item were calculated. Although the range and size of scores on all the items on the inventory were of interest, only the scores on the two items related to the target passage were used in the data analysis. Mean scores on these two items were summed so that the total possible points was 8. Scores ranged from 0 to 8, with a mean of 3.45 and a standard deviation of 2.23.

Writing sample. Scoring criteria for the essays were derived from Diederich's (1974) holistic scoring system. In accord with this system, compositions were rated on the following aspects: (a) ideas (2, 4, 6, 8, or 10 points), (b) organization (2, 4, 6, 8, or 10 points), (c) wording (1, 2, 3, 4, or 5 points), and (d) flavor (1, 2, 3, 4, or 5 points). The design of Diederich's scale provides that compositions which rank high in the scoring criteria receive a score at or near the top of the scale, while poorly-written compositions receive correspondingly low scores. In addition, ideas and organization are rated twice as important as wording and flavor, so that the total points possible for any essay was 30. The score for
each essay then represented the subject's writing ability.

The essays were scored by both the researcher and a doctoral student in the field of English education. Rating was done independently, and a random sample of 20% of the protocols was examined in order to measure interrater agreement. This agreement, calculated as the ratio of agreements to the sum of agreements and disagreements, was .82. Raw scores ranged from 6 to 25, with a mean of 13.55 and a standard deviation of 4.77.

**Multiple choice tests.** These instruments were scored by the researcher using an answer key, with each correct answer receiving one point. A second rater scored the papers to check for accuracy, and no errors were found. Raw scores on the prereading measure ranged from 1 to 9, with a mean of 4.93 and a standard deviation of 2.13. The postreading test scores ranged from 3 to 13, with a mean of 8.26 and a standard deviation of 2.12.

**Written recalls.** The free recalls and summaries were rated on six different criteria in order to obtain scores for the six dependent variables. First, written recalls were scored for total number of words. Words were counted mechanically by the researcher, with a second rater checking the counting for accuracy. Words included in the free recalls ranged from 27 to 157, with a mean of 69.96 and a standard deviation of 31.12. The number of words in the summaries ranged from 24 to 157, with a mean of 61.38.
and a standard deviation of 30.79.

Second, the idea units were scored. A scoring template was prepared for the free recalls and summaries using a method described by Johnson (1970). First, the target passage was typed on plain white 8-1/2 by 11-inch paper, double-spaced. A group of 30 doctoral students in education were then asked to parse the text into idea units by reading it to themselves and marking a slash (/) every time they paused for emphasis, to catch a breath, or to enhance meaning. Johnson (1970) referred to idea units obtained in this way as pausal units and showed empirically that they represent a valid means of parsing the idea units of the text. The raters' markings were compared, and a slash was placed in the text wherever at least 50% of the raters had placed one. Agreement in most cases was much higher than the 50% criterion. As a result of this procedure, the text was divided into 195 idea units, with a mean unit length of 9.87 words.

The next step in preparing the template was retyping the text with the finalized slashes inserted. It was also divided into five sections to make the importance-rating task more manageable. A second group of raters was employed to rate the relative importance of the idea units. This group (N = 18) was made up of reading educators: students enrolled in either a masters degree program or a doctoral program, and graduate faculty in the
reading program.

Rating of the text for importance of ideas was accomplished in the following way. Raters were asked to read each section, one at a time, and to mark out the least-important idea units until approximately 25% of the idea units and words had been deleted. Next, they deleted another 25% of the units, and finally a third group of least-important idea units was eliminated. Additionally, the number of words deleted was required to be within 10 words of 25% of the total words for the section. This stipulation was based on the mean idea unit length (\( M = 9.87 \) words) and was made to ensure that approximately 25% of the text was deleted with each pass.

The idea units of text that remained after three sets of deletions (25%) represented the most important ideas of the passage. Each of the five sections of the passage was completed in the same manner. Pencils of different color were used for each level of deletions so that a standard color code could be used to assign an importance rating to each idea unit in the passage. Thus, the idea units could be assigned a score and ranked in level of importance, from four (highest) to one (lowest).

Once the rating task had been completed, the ratings for each idea unit were tabulated, and means were calculated. Next, the 195 idea units were ranked in order of importance. The top 25% of the idea units were
assigned an importance score of four, and the the next three quartiles received scores of three, two, and one, respectively. Finally, a scoring sheet template was derived, listing the idea units in the order in which they appeared in the target passage and indicating the importance score for each.

Idea units in subjects' free recalls and summaries were then counted using the template described above. A second rater counted the idea units in a random sample of 20% of the protocols, and interrater agreement was calculated to be .91. Idea units in the free recalls ranged from 3 to 24, with a mean of 9.49 and a standard deviation of 4.90. Idea units included in the summaries ranged from 1 to 20, with a mean of 7.85 and a standard deviation of 4.46.

The third type of scoring, number of important idea units, was accomplished using the template described above. Each time the protocol included an idea unit from the target passage, a point value of one to four was awarded, depending upon the relative importance of the idea unit itself. Idea units receiving a rating of 4 were considered important ideas. These idea units were then counted. Interrater agreement on the number of important idea units was computed to be .94.

Fourth, as a measure of the relative importance of the idea units included by the subjects in their recalls,
the proportion of the number of rated-important (level 4) idea units to the total number included was calculated. These scores, denoted as proportion of important idea units, ranged from .00 to 1.00 for both the free recalls and the summaries. The mean for the free recalls was .38, with a standard deviation of .19, and the means for the summaries was .46, with a standard deviation of .22.

Fifth, a score for efficiency of expression was calculated. First, points were totalled for each recall protocol to arrive at a total importance score. No points were awarded for information that was not found in the target passage. A second rater also scored the importance level of a randomly-selected 20% of the recalls, and interrater agreement was computed to be .90. Then, to arrive at a score for efficiency of expression (Garner, 1982), the total importance score of each recall was divided by the number of words in the written recall. The summarizing efficiency scores ranged from .08 to .68, with a mean of .40 and a standard deviation of 0.13. Efficiency scores for free recalls ranged from .12 to .71, with a mean of .38 and a standard deviation of .13.

Finally, written recalls were scored for the types of transformations of the original text which they displayed. Following criteria set forth by Winograd (1984), the recalls were compared to the original text and were scored in the following way. Subjects' recalls could have
transformed the target passage ideas in any of four possible ways: (a) reproductions, in which the text was essentially duplicated, were assigned 1 point; (b) run-on combinations, in which a single sentence combined two or more idea units, but in an unorganized way, received two points; (c) combinations, in which a sentence combined two or more idea units in a meaningful way, were assigned three points; and, (d) inventions, where individual sentences conveyed the meaning of several idea units or units of even a paragraph or more, received a score of four points. The researcher scored the recall protocols according to these criteria, and a second rater scored a random sample of 20% to establish interrater agreement. Agreement was calculated to be .90. Because adult summaries typically include a high proportion of combinations and inventions, the proportion of these transformations to the total number of sentences written was calculated. The scores for the summaries, denoted as summary rule use, ranged from 0.00 to 1.00, with a mean .45 and a standard deviation of .31. The corresponding scores for the free recalls also ranged from 0.00 to 1.00, with a mean of .45 and a standard deviation of .27.
CHAPTER 3

RESULTS

Three sets of analyses were undertaken with the data. First, a multivariate repeated measures analysis of variance was conducted in order to examine differences between the means for the free recalls and summaries. In addition, standard regression analyses were used to determine the predictive effects of the independent variables on the six criterion measures for the two tasks. Finally, correlations between the free recall and summary scores and scores on two other measures of comprehension, multiple-choice questions and a standardized reading comprehension test, were calculated. The raw data are presented in Appendix I.

MANOVA

There were two repeated independent measures: Task (free recalls and summaries obtained for each subject) and Order (immediate and delayed). The dependent variables were the six scores on the free recalls and summaries. For each of the effects, an effect size (ES) statistic $f$ was calculated, based on the eta-squared coefficient and according to procedures outlined by Cohen (1977). Cohen also enumerates criteria for assessing the relative size of the effects; small effects are those in the range of $f = .10$, moderate effects have $f$ values around .25, and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
large effects have an $f$ of at least .40.

The effect for Order was not significant, multivariate $F(6, 46) = 0.87, p < .52$. On the other hand, both the effects for Task, multivariate $F(6, 46) = 2.81, p < .02$, and Task X Order interaction, multivariate $F(6, 46) = 4.56, p < .001$, were significant. Individual univariate repeated measures ANOVAs indicated significant differences for several of the dependent variables. These results will be discussed in the sections that follow. Means and standard deviations for the dependent measures are presented in Table 1.

Insert Table 1 about here

**Words.** This dependent variable was a measure of the number of words that subjects included in their recall protocols. The effect for Task was not significant, $F(1, 51) = 2.33, p < .13$. The effect size (ES) was small, $f = .12$. This indicates that the number of words written did not differ by task; free recall and summary were similar in this regard. However, the Task X Order interaction was significant, $F(1, 51) = 20.27, p < .0001$. The effect size, $f = .37$, can be classified as moderate to large. A comparison of the number of words for the free recalls ($M = 76.18$) and summaries ($M = 77.64$) written in the immediate condition indicates that subjects did not differ
Table 1

Means and Standard Deviations for Dependent Variables

<table>
<thead>
<tr>
<th>Task</th>
<th>Free recall</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Immediate</td>
<td>Delayed</td>
</tr>
<tr>
<td>Words</td>
<td>76.18 (32.88)</td>
<td>63.00 (28.22)</td>
</tr>
<tr>
<td>Idea units</td>
<td>9.54 (5.19)</td>
<td>9.44 (4.66)</td>
</tr>
<tr>
<td>Important idea units</td>
<td>3.60 (2.75)</td>
<td>3.64 (2.45)</td>
</tr>
<tr>
<td>Proportion of important idea units</td>
<td>.37 (.19)</td>
<td>.40 (.16)</td>
</tr>
<tr>
<td>Efficiency</td>
<td>.35 (.13)</td>
<td>.42 (.13)</td>
</tr>
<tr>
<td>Rule use</td>
<td>.43 (.24)</td>
<td>.47 (.28)</td>
</tr>
</tbody>
</table>

Note. N = 53 overall
n = 25 for summary immediate, free recall delayed
n = 28 for summary delayed, free recall immediate
in the number of words they wrote in the first task. Both free recalls (M = 63.00) and summaries (M = 47.11) written in the delayed condition were shorter than their first-position counterparts, and the summaries were shorter than the free recalls. These results taken together indicate that when subjects first wrote down everything they could remember and then tried to summarize, they were successful in condensing their remarks, at least in terms of words. However, when they summarized first and then tried to recall everything, the outcome was the opposite; the free recall actually turned out to be shorter than the summary.

_Idea units_. This dependent variable measured the number of idea units, regardless of their relative importance, that subjects included in their recall protocols. Both the Task, F(1, 51) = 5.25, p < .03, and Task X Order interaction, F(1, 51) = 11.84, p < .001, were significant. Effect sizes were f = .20 for Task and f = .25 for the interaction. The differences observed for Task result from the fact that the means for all free recalls (M = 9.49) are significantly higher than for the summaries (M = 7.85). This would be expected, since a successful summary requires deletion of unimportant material. A look at the interaction, however, clarifies these results. It is noteworthy that free recalls written either in the immediate (M = 9.54) or delayed (M = 9.44)
condition remain nearly constant. But such was not the case with summaries. Those written in the delayed condition (M = 5.75) contained significantly fewer idea units than those in the immediate condition (M = 10.2).

**Number of important idea units.** This dependent variable measured the number of rated-important, or level 4, idea units included by subjects in their recall protocols. The main effect for Task, F(1, 51) = 0.00, p < .96, f = .00 was not significant. Only the Task X Order interaction, F(1, 51) = 8.87, p < .004, was significant. The effect size for the interaction, f = .20, was small to moderate. Again, free recalls written in either the immediate or delayed condition remained relatively stable (M = 3.60 and 3.64, respectively). Summaries written first, however, contain significantly more important idea units (M = 4.64) than those written second (M = 2.64). The situation becomes more apparent when the means for the task conditions are examined. Where free recalls were written first, followed by a summary, subjects wrote fewer important idea units in the summary (M = 2.64) than they did the free recall (M = 3.60). When the summary was written first, it contained more important idea units (M = 4.64) than the free recall (M = 3.64).

**Proportion of important idea units.** This dependent variable measured the proportion of rated-important (level 4) idea units to the total number of idea units included
in the recall protocols. The Task X Order interaction was not significant, $F(1, 51) = 1.23, p < .27$. Only the main effect for Task was significant, $F(1, 51) = 4.03, p < .05$. The effect size for Task, $f = .18$, was small to moderate. Subjects included a larger proportion of important information, as compared to total information included, in their summaries than they did in their free recalls, regardless of order. For example, summaries written in the immediate condition contained a higher proportion of important information ($M = .43$) than did the free recall which followed ($M = .40$), although this difference is not large. The difference was larger, however, for the free recall written in the immediate condition ($M = .37$) and for the summary written in the delayed condition ($M = .48$).

**Efficiency.** This dependent variable measured the proportion of important information that subjects were able to communicate relative to the number of words used. Neither the main effect for Task, $F(1, 51) = 0.23, p < .63, f = .00$, nor the Task X Order interaction, $F(1, 51) = 1.10, p < .30, f = .08$, was significant. The similarity of the means for the two tasks, regardless of order, is again apparent. Ironically, when the proportion of important information was highest ($M = .48$) for summaries written in the delayed condition, the efficiency rate was the second lowest for the four conditions ($M = .38$). This
condition was also the one in which the subjects wrote the smallest number of words ($M = 47.11$) and included the smallest number of idea units ($M = 5.75$) and important idea units ($M = 2.64$). Apparently, when the subjects wrote shorter summaries, the amount of important information they included and their efficiency in expressing it declined as a result of their efforts to be succinct.

**Rule use.** This dependent variable reflected the proportion of sentences in the recall protocol which were either combinations of idea units from the passage or inventions representing generalizations. As with efficiency, neither the main effect for Task, $F(1, 51) = 0.03, p < .87$, nor the Task X Order interaction, $F(1, 51) = 2.63, p < .11$, were significant. Both effect sizes, $f = .00$ and $.14$, respectively, were correspondingly small. These results suggest that subjects used about the same proportion of combinations and inventions in their free recalls and summaries, regardless of the order in which they were written.

Table 2 summarizes the results of the MANOVA and the follow-up univariate analyses. Specifically, the main effect for Task was significant for only two of the dependent variables: (a) number of idea units, and (b) proportion of important idea units. Task X Order interactions were significant for three of the dependent
variables: (a) number of words, (b) number of idea units, and (c) number of important idea units.

__________________________

Insert Table 2 about here

__________________________

Regression Analyses

There were three predictor variables for the analyses. These variables and the measures from which they were obtained are as follows: (a) prior knowledge (PK), from prereading multiple-choice questions about the passage; (b) writing ability (WA), from an essay-writing sample collected from each subject; and, (c) topic interest (TI), from an interest inventory administered prior to reading. Descriptive data for the three independent variables appears in Table 3.

__________________________

Insert Table 3 about here

__________________________

The six criterion variables were the six scores obtained for each subject's free recall and summary. Specifically, these were: (a) number of words, (b) number of idea units, (c) number of important idea units, (d) proportion of important idea units, (e) efficiency, and (f) rule use.

Effect sizes (ES) for the regression analyses were also calculated. The effect size statistic $f^2$
Table 2

Summary of MANOVA Results

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Task F</th>
<th>Task X Order F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words</td>
<td>2.33</td>
<td>20.27***</td>
</tr>
<tr>
<td>Idea units</td>
<td>5.25*</td>
<td>11.84***</td>
</tr>
<tr>
<td>Important idea units</td>
<td>0.00</td>
<td>8.87**</td>
</tr>
<tr>
<td>Proportion of important idea units</td>
<td>4.03*</td>
<td>1.23</td>
</tr>
<tr>
<td>Efficiency</td>
<td>0.23</td>
<td>1.10</td>
</tr>
<tr>
<td>Rule use</td>
<td>0.03</td>
<td>2.63</td>
</tr>
</tbody>
</table>

Note. N = 53.

* P < .05.  ** P < .01.  *** P < .001.
Table 3

Descriptive Statistics for Independent Variables

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior knowledge (PK)</td>
<td>4.93</td>
<td>2.13</td>
</tr>
<tr>
<td>Writing ability (WA)</td>
<td>13.55</td>
<td>4.77</td>
</tr>
<tr>
<td>Topic interest (TI)</td>
<td>3.45</td>
<td>2.23</td>
</tr>
</tbody>
</table>

Note. N = 53.
refers to the proportion of variance explained to the unexplained variance. Specific steps in calculation are described by Cohen (1977). For purposes of discussion, Cohen operationally defines small effect sizes as $\hat{f}^2 = .02$, moderate effect sizes as $\hat{f}^2 = .15$, and large effect sizes as $\hat{f}^2 = .35$. These labels will be applied where appropriate. Results of the regression analyses for free recalls and summaries are presented in Tables 4 and 5.

Words. For free recalls, the model explained 16% of the variance, $F(3, 49) = 3.22, p < .03$. PK was the strongest predictor, $F(1, 51) = 6.66, p < .01$, and by itself explained 11% of the variance. The other two predictors, WA and TI, accounted for only 4% and 2% of the variance, respectively, and failed to reach significance as predictors. For summaries, only 7% of the variance could be explained overall, and no predictors were significant.

Idea units. For free recalls, 18% of the variance was explained by the model, $F(3, 49) = 3.50, p < .03$. PK was the only significant predictor, $F(1, 51) = 4.99, p < .03$. PK alone accounted for 8% of the variance, and TI
Table 4
Regression Analysis for Free Recalls

<table>
<thead>
<tr>
<th>Criterion variable</th>
<th>Predictor variable</th>
<th>Beta weight (SE)</th>
<th>F</th>
<th>Pct. of variance</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words</td>
<td>PK</td>
<td>.34 (1.91)</td>
<td>6.66**</td>
<td>.11 (.12)</td>
<td></td>
</tr>
<tr>
<td>Idea units</td>
<td>PK</td>
<td>.29 (.30)</td>
<td>4.99*</td>
<td>.08 (.09)</td>
<td></td>
</tr>
<tr>
<td>Important idea units</td>
<td>WA</td>
<td>.27 (.07)</td>
<td>4.30*</td>
<td>.07 (.08)</td>
<td></td>
</tr>
<tr>
<td>Proportion of important idea units</td>
<td>PK</td>
<td>.15 (.01)</td>
<td>1.14</td>
<td>.02 (.02)</td>
<td></td>
</tr>
<tr>
<td>Efficiency</td>
<td>TI</td>
<td>.21 (.01)</td>
<td>2.30</td>
<td>.04 (.05)</td>
<td></td>
</tr>
<tr>
<td>Rule use</td>
<td>TI</td>
<td>-.14 (.01)</td>
<td>1.06</td>
<td>.02 (.02)</td>
<td></td>
</tr>
</tbody>
</table>

Note. N = 53.
* $p < .05$. ** $p < .01$.  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 5

Regression Analysis for Summaries

<table>
<thead>
<tr>
<th>Criterion variable</th>
<th>Predictor variable</th>
<th>Beta weight (SE)</th>
<th>F</th>
<th>Pct. of variance</th>
<th>f²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words</td>
<td>TI</td>
<td>.22 (.190)</td>
<td>2.46</td>
<td>.05 .05</td>
<td></td>
</tr>
<tr>
<td>Idea units</td>
<td>TI</td>
<td>.25 (.27)</td>
<td>3.49</td>
<td>.06 .07</td>
<td></td>
</tr>
<tr>
<td>Important idea units</td>
<td>TI</td>
<td>.28 (.15)</td>
<td>4.53*</td>
<td>.08 .08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WA</td>
<td>.27 (.07)</td>
<td>4.35*</td>
<td>.07 .08</td>
<td></td>
</tr>
<tr>
<td>Proportion of important</td>
<td>WA</td>
<td>.25 (.01)</td>
<td>3.29</td>
<td>.06 .07</td>
<td></td>
</tr>
<tr>
<td>idea units</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency</td>
<td>PK</td>
<td>.14 (.01)</td>
<td>.94</td>
<td>.02 .02</td>
<td></td>
</tr>
<tr>
<td>Rule use</td>
<td>WA</td>
<td>.17 (.01)</td>
<td>1.41</td>
<td>.03 .03</td>
<td></td>
</tr>
</tbody>
</table>

Note. N = 53.

* P < .05.
and WA explained 6% and 3%, respectively. For summaries, only 9% of the variance could be explained by the model, and no predictors were significant.

**Important idea units.** For free recalls, the model explained 17% of the variance, $F(3, 49) = 3.41, p < .02$. WA was the only significant predictor, $F(1, 51) = 4.30, p < .04$, and explained 7% of the variance. TI and PK accounted for 6% and 5% of the variance, respectively, but neither reached significance as a predictor.

For summaries, the model explained 16% of the variance, $F(3, 49) = 3.21, p < .03$. Two predictors were significant. TI alone explained 8% of the variance, $F(1, 51) = 4.53, p < .05$. WA explained an additional 7% of the variance after TI was accounted for, $F(1, 51) = 4.35, p < .05$.

**Proportion of important idea units.** No significant predictors were found for either the free recalls or the summaries. Only 4% of the variance in the free recalls could be explained by the model. For summaries, 8% of the variance was explained, but only WA even approached significance, $F(1, 51) = 1.14, p < .08$.

**Efficiency.** No predictors were significant for either the free recalls or the summaries. Percentages of variance explained were 6% and 2%, respectively.

**Rule use.** No significant predictors were found for either the free recalls or the summaries. In each case,
the model could account for only 3% of the variance.

**Correlational Analysis**

Pearson product-moment correlation coefficients were computed for each of the six dependent measures for free recalls and summaries with the postreading multiple-choice questions score (MC) and the standardized reading comprehension test score (STAN). These results are presented in Table 6.

Insert Table 6 about here

**Words.** The correlations for this dependent variable were significant in two instances. MC was significantly correlated with free recalls, $r = .37, p < .006$, and summaries, $r = .38, p < .005$.

**Idea units.** All correlations were significant for this dependent variable. MC was significantly correlated with free recalls, $r = .59, p < .0001$, and summaries, $r = .44, p < .0009$. Additionally, STAN correlated significantly with both free recalls, $r = .52, p < .0001$, and summaries, $r = .30, p < .03$.

**Important idea units.** All correlations for this dependent variable were significant. MC was significantly correlated with both free recalls, $r = .60, p < .0001$, and summaries, $r = .51, p < .0001$. In addition, STAN correlated significantly with both free recalls, $r = .52,$
Table 6

Pearson Product Moment Correlations for MC and STAN

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Free recall</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MC</td>
<td>STAN</td>
</tr>
<tr>
<td>Words</td>
<td>.37**</td>
<td>.21</td>
</tr>
<tr>
<td>Idea units</td>
<td>.59***</td>
<td>.52***</td>
</tr>
<tr>
<td>Important idea units</td>
<td>.60***</td>
<td>.52***</td>
</tr>
<tr>
<td>Proportion of important idea units</td>
<td>.19</td>
<td>.29*</td>
</tr>
<tr>
<td>Efficiency</td>
<td>.56***</td>
<td>.62***</td>
</tr>
<tr>
<td>Rule use</td>
<td>.32*</td>
<td>.25</td>
</tr>
</tbody>
</table>

Note. N = 53.

*p < .05.  **p < .01.  ***p < .001.
Proportion of important idea units. Only one correlation coefficient was significant for this variable. STAN was significantly correlated with free recalls, \( r = .29, p < .03 \).

Efficiency. All correlations for this dependent variable were significant. MC was significantly correlated with both free recalls, \( r = .56, p < .0001 \), and summaries, \( r = .32, p < .02 \). Additionally, STAN correlated significantly with both free recalls, \( r = .62, p < .0001 \), and summaries, \( r = .27, p < .05 \).

Rule use. For this dependent variable, MC had a significant correlation with both free recalls, \( r = .32, p < .02 \), and summaries, \( r = .34, p < .01 \). Finally, STAN correlated significantly with summaries, \( r = .44, p < .001 \).
CHAPTER 4

DISCUSSION

The purpose of this study was to investigate the use of written summaries as a measure of reading comprehension. One area of interest was the similarities and differences that exist between the ability to write a free recall and a summary after reading a passage. Additionally, the effects of several factors on the quality of written summaries were measured. These factors were prior knowledge, topic interest, and writing ability. Finally, the relationship between summaries and other measures of comprehension, namely, multiple-choice questions and a standardized test of reading comprehension, was examined.

First, comparisons of the means for the free recalls and summaries revealed significant differences on only two of the six dependent measures. One difference was in the number of idea units; free recalls contained more idea units than the summaries did. Given the nature of the directions that were given to the subjects, this result was not unexpected. For free recalls, subjects were told to write all that they could remember from the passage, regardless of the relative importance of the ideas that were recorded. Summaries, on the other hand, were to include only the most important information; it seems apparent that some idea units would be deleted in an

46
attempt to keep only that which was important.

The second difference between the summaries and the free recalls is related to the first. Subjects included a proportionally larger quantity of important information in their summaries than they did in their free recalls. This result indicates that the subjects, again, were making an attempt to summarize by including only the most important information.

A significant Task X Order interaction effect was also noted. However, because the main effect for Order was not significant, it appears that this difference was, in fact, an artifact of the methodology of the study. Hence, no conclusions can be drawn here.

Also revealing, however, were the measures in which the free recalls and summaries did not differ. Follow-up univariate analyses indicated no significant differences between the two tasks on number of words, number of important idea units, efficiency of expression, or rule use. These results were not surprising, given the similarities noted between the free recall and summary protocols themselves. Virtually the same style of writing was employed in both protocols, even though the directions indicated that this was not necessary. Hence, protocols did not differ in length, efficiency, or rule use. These observations, coupled with the lack of statistical differences, indicate that, although these seventh-grade
subjects were making an attempt to summarize, they were not entirely successful in their efforts. This conclusion confirms the findings of previous research (e.g., Garner, 1985) that seventh-grade students have not yet fully developed their summarizing abilities.

A comparison with adults' free recalls and summaries for the same passage was made in an attempt to reveal how the seventh graders' performance would compare with that of more proficient summarizers. Adults outperformed the seventh graders in three of six criteria in an almost three-to-one ratio. For the other two criteria, adults' efficiency of expression was nearly 20% better, and rule use was almost double that of the younger students. The one area in which seventh graders matched adult performance was in the proportion of important idea units; both groups had a score of .47 on this criterion. These findings parallel the developmental patterns noted by other researchers (e.g., Johnson, 1983; McGee and Tompkins, 1982; Winograd, 1984). It may be concluded, then, that although these seventh graders were attempting to summarize, their efforts were hampered by their lack of experience in this task.

Next, an attempt was made to examine other factors that might have influenced the subjects' summary writing performance. First, it was noted in the regression analyses that no significant predictors were found for
proportion of important idea units, efficiency of expression, or rule use. However, number of words, number of idea units, and number of important idea units are explainable in terms of the three predictor variables. For number of words, prior knowledge was found to explain a significant proportion of the variance for free recalls. It would seem that the more readers bring to the task of free recall, the greater the amount that they are able to write after reading. For summaries, however, number of words was not explained to a significant degree. It could be that, because of the succinctness required for summarizing, number of words is not necessarily a relevant measure for summaries.

For number of idea units, prior knowledge again predicted scores for free recalls but not for summaries. The differing natures of the two tasks again help to clarify the results. Since free recalls may include any idea unit, regardless of its relative importance, the amount that is already known may influence that amount recalled. On the other hand, the smaller number of idea units overall in the summaries may cause subjects to negate some of what they already knew in an attempt to be succinct.

For number of important idea units, writing ability was found to be a significant predictor for both tasks. This result was not unexpected, given Johnston's (1983)
observations that young writers have difficulty in deciding what is important to include in a composition.

Additionally, topic interest was found to explain a significant proportion of the variance in the number of important idea units in summaries. As had been found by Baldwin, Peleg-Bruckner, and McClintock (1985), readers' interest in a topic may influence their comprehension of the passage and, thus, the important information that they have available for inclusion in their summaries.

Free recalls and summaries were also compared with two other measures of comprehension, namely multiple-choice questions and a standardized test of reading comprehension, in an effort to determine how efficient free recalls and summaries were as comprehension measures. First, the multiple-choice scores were correlated with both the free recalls and the summaries on every dependent measure except for proportion of important idea units. The fact that the multiple-choice questions in this study were passage-dependent implies that an understanding of the target passage is related to the ability to write a free recall and a summary. The size of the correlations, however, was rather modest, indicating that there are other aspects of comprehension being measured by free recalls and summaries that are unrelated to passage comprehension. At best, it can be concluded that multiple-choice questions and written recalls are
measuring some of the same aspects of reading comprehension.

For standardized tests of reading comprehension, significant correlations were found for both free recalls and summaries on three of the dependent variables: (a) number of idea units, (b) number of important idea units, and (c) efficiency of expression. These results indicate that global comprehension ability, as measured by standardized tests, is only somewhat related to the ability to recall and record significant quantities of information from a target passage.

For two other dependent variables, however, the findings are not so equivocal. STAN was correlated with free recalls for proportion of important ideas but was uncorrelated with summaries on this measure. Similarly, STAN was significantly correlated with summaries for rule use, but no correlation was found with free recalls. It seems that these two variables, in particular, were affected by the counterbalancing of the two tasks. For this reason, these apparently confusing results are considered artifacts of the methodology employed rather than true relationships.

Based on these results, the following conclusions may be stated. First, the use of summaries seems to tell us little more than can be found out by the use of a free recall measure. Second, summarizing appears to be a task
for which seventh graders are not fully prepared. Third, some factors, other than reading comprehension, may also play a role in the ability to summarize. Finally, summaries appear to measure some, but not all, aspects of reading comprehension.

From these conclusions, the following implications for instruction which may be drawn. First, when summaries are used as a measure of reading comprehension, teachers should take into account the notion that some other factors may influence the quality of the summary. As shown in this study, prior knowledge, topic interest, and writing ability may, at some time, play a role in the writing of summaries. Second, because the ability to summarize has been shown in this and other studies (e.g., Garner, 1985; Winograd, 1984) to be one which develops slowly and over a period of time, some form of direct instruction in summarizing should be considered in order to help students develop this ability. Finally, because summaries appear to measure only some aspects of comprehension, and because students at the seventh-grade level are as yet immature in their development of summarizing ability, some additional measures of comprehension should be included to assess students' abilities in this area (Readence & Moore, 1983).

Additionally, several suggestions for future research can be made. First, because of the confounding effect
noted for the counterbalancing of the free recall and summary tasks, it would be advisable to consider having students first write a free recall and then a summary, a sequence more typical of classroom instruction. Second, the use of summarizing instruction should be considered in order to examine the value instruction has on the use of summaries as a comprehension measure. Third, target passages shorter in length should be used so that fatigue could be ruled out as a potentially confounding variable. Finally, some care should be taken to prevent the attrition of subjects that occurred in this study. Although attrition can be a problem in any study which involves more than one data collection session, careful design may be able to minimize these effects.

In conclusion, summary writing is an area which we have only begun to explore. The findings of this study have indicated that written summaries have a place in the assessment of reading comprehension. Clearly, however, more research is needed to determine just what that place is.
References


Verbal Behavior, 11, 717-726.


Fry, E. B. (1977). Fry's readability graph:


Johnson, N. S. (1983). What do you do if you can't tell


annual meeting of the National Reading Conference,
Clearwater Beach, FL.

of children's statements of the main idea in reading. In
J. A. Figurel (Ed.), Reading and realism: Proceedings of the
International Reading Association, 14, 692-697.


Preston, R. C. (1962). A new approach to judging the
validity of reading comprehension tests. In J. A.
Figurel (Ed.), Challenge and experiment in reading:
Proceedings of the International Reading Association, 7,
166-167.

Content area reading: An integrated approach (2nd ed.).
Dubuque, IA: Kendall/Hunt.

historical perspective on standardized reading

Rosenblatt, L. M. (1978). The reader, the text, the poem:
The transactional theory of literary work. Carbondale,
IL: Southern Illinois University Press.

cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer
(Eds.), Theoretical issues in reading comprehension:
Perspectives from cognitive psychology, linguistics,
artificial intelligence, and education (pp. 33-58). Hillsdale, NJ: Erlbaum.


Thomas, S., & Bridge, C. A. (1980). A comparison of subjects' cloze scores and their ability to employ macrostructure operations in the generation of
summaries. In M. L. Kamil & A. J. Moe (Eds.), Perspectives on reading research and instruction (pp. 69-77). Twenty-ninth Yearbook of the National Reading Conference. Washington, DC: National Reading Conference.


APPENDIX A

DEFINITION OF TERMS
DEFINITION OF TERMS

For the purposes of this study, the following terms are defined:

expository writing - text whose primary purpose is to describe or explain, rather than to tell a story.

free recall - a list, written without the text available, of all the ideas a reader can remember from a passage previously read.

general writing ability - the ability to express information in writing in a well-organized expository composition.

prior knowledge - the background knowledge about a topic that readers possess before reading a passage about that topic.

summarized recall - a measure of reading comprehension of a passage consisting of a written recitation of the reader's gist of the material read, usually organized around the structure of main idea(s) and supporting details.
summarizing ability - the ability to abstract the superordinate and subordinate ideas from a passage into a shortened, but well-organized, written composition.

topic interest - the degree to which students like to read about a given subject.
APPENDIX B

REVIEW OF THE LITERATURE
Review of the Literature

The writing of summaries is a technique frequently employed by classroom teachers to see how well their students have understood and remembered what they read in a textbook passage (Palincsar & Brown, 1984; Roe, Stoodt, & Burns, 1978). Summaries can be used to provide teachers with an indicator of what Rosenblatt (1978) has termed "efferent reading", or what remains after reading. Thus, the informational product that readers have constructed can be revealed when they write a summary.

Kintsch and van Dijk (1978; van Dijk & Kintsch, 1983) have postulated that readers formulate a gist of a passage while reading. This act is accomplished through a series of manipulations of the text designed to reduce the original body of information to only its essential components. Guided by their prevailing schemata for the topic and text type, readers apply such rules as deletion, generalization, and construction to the microstructure of a text to form a macrostructure, or mental gist, of the information they have deemed important enough to remember. Trivial or redundant propositions have been removed so that only the macropropositions remain.

Because the processes that readers use to form such mental gists resemble those that skilled readers use when writing a summary of text material, many researchers (e.g., Bean & Steenwyk, 1984; Taylor, 1984; Taylor & Beach, 1984; Thomas & Bridge, 1980) have inferred that a
summary written after reading is a reflection of that gist. While the gist itself may be quite brief (Cunningham, 1982), a reader seeking to write a summary can take this succinct statement and by applying the rules of addition, specification, and particularization, transform it into a more detailed written recall version of the text (Taylor & Beach, 1984). The resultant formulation is a new version of the text driven by the reader's unique schemata.

Recent research related to schema theory (e.g., Rumelhart, 1980; Rumelhart & Norman, 1978) has supported this conclusion, in that readers, in order to fully comprehend a passage, must integrate the new knowledge that passage presents with their prior knowledge of the topic to form a new schema. Therefore, when a summary is written, it will contain both new and old information, reorganized and presented as a summary of the passage (Rumelhart, 1980), but representative of the reader's current schemata related to the topic. It seems evident that the readers' prior knowledge must play a critical role in the manipulations that are necessary to produce a summary.

The wide acceptance of a schema theory of reading comprehension has contributed to the use of written recalls to assess comprehension, not only in the
classroom, but in various kinds of research studies (Kalmbach, 1986). Taylor (1984), in a review of 15 recent studies which investigated comprehension, found that free recall was used as a comprehension measure in seven of the studies. More specifically, several studies have been identified which made use of written summaries in their methodology. These will be described in the section to follow.

Written Summaries as a Measure of Comprehension

Freebody and Anderson (1983a), in their experiment relating vocabulary difficulty, text cohesion, and schema availability to reading comprehension, used 84 sixth-grade subjects who were classified as good readers. Subjects read one of six 250- to 300-word passages from a fifth-grade social studies text which had been altered in order to vary either text cohesion or vocabulary difficulty. After reading, subjects completed a distractor task, wrote both a free recall and a two- to-three sentence summary of the main ideas of the target passage, and completed a 13-item sentence verification task.

To assess the subjects' free recalls, the text was divided into propositions, where a clause or phrase expressed a new idea for the first time in the text. Summaries were scored by comparison of the subjects' summaries with those produced by a panel of adults. For the sentence verification task, subjects received a point
for each sentence correctly verified or rejected.

A multiple regression analysis was used to partition the variance. Reading ability, as measured by a standardized test of reading comprehension and language ability, was entered first into the regression equation in order to provide a more sensitive test of the other factors included in the design. However, no allowance was made for either the subjects' prior knowledge of the topic, writing ability, or summarizing ability. Interestingly, a significant main effect for passage appeared as a within-subjects variable, so that content of the passage read did account for some of the variance. Perhaps an assessment of the subjects' prior knowledge and interest in the topic would have been useful in order to control for such an effect. In fact, passages which contained ideas unfamiliar to sixth-graders proved to be most difficult for them to summarize. Subjects also tended to omit ideas that were "obvious", perhaps regarding them as too trivial for inclusion in a relatively short summary. Not unexpectedly, reading and language ability accounted for a significant portion of the variance in all three tasks, suggesting that these, too, must be included in a careful analysis of the data from the present study.

In a second experiment reported in their study, Freebody and Anderson (1983a) examined the effects of
prior knowledge and vocabulary difficulty on the recall of their sixth-grade subjects. Four passages were constructed: familiar and unfamiliar versions of both a game theme and a visiting theme. Subjects each read two passages: a familiar version of one theme and an unfamiliar version of the other. Vocabulary difficulty was a between-subjects variable. As in the first experiment, subjects read the passage and then wrote a free recall, a summary, and completed a sentence verification test. A regression analysis was performed as before.

Results showed that familiarity had the expected effect on recall. Freebody and Anderson described two possible ways in which this finding could have occurred. First, an unfamiliar topic can be the cause of encoding deficits; when ambiguous terms are encountered, bridging inferences cannot be successfully made. Alternatively, retrieval can be difficult because of a lack of structured prior knowledge. Connections among concepts are not obvious, so text cannot be recalled in an integrated form. Again, because of the demonstrated effect of topic familiarity on recall, it seems prudent to assess such knowledge before reading has taken place.

In a related study, Freebody and Anderson (1983b) examined the effects on text comprehension of differing proportions and locations of vocabulary difficulty.
Methodology was similar to the study reported above, with 79 sixth-grade subjects writing both a free recall and a summary and completing a sentence verification task. The effects of vocabulary difficulty on free recall were unclear and involved a number of interactions. However, the subjects' summaries more closely matched the adults' when difficult words were placed in unimportant locations in the text (subordinate propositions).

In a study that did much to validate summary scores as a measure of reading comprehension, Thomas and Bridge (1980) examined the relationship between subjects' cloze scores and their summarizing ability. Their assumption was that cloze performance is affected by some underlying factors at the macrostructure level which could be measured in a subject-structured recall task. They reasoned that the operations used by the subjects to manipulate the text for recall should shed light on the kinds of macrostructure procedures at work in cloze comprehension.

Eighteen eighth-grade subjects first completed a cloze test on a novel passage. Then they were asked to read the unmutilated text and write a summary of the ideas they recalled. The passage used was expository, about 300 words in length, and written at the eighth-grade level according to the Fry Graph (1977).

Summaries were scored according to a system based
upon rated importance of idea units included within the summaries. The passage was first parsed into idea units according to criteria set forth by Johnson (1970), each encompassing a complete thought. Interrater agreement was calculated and deemed to be satisfactory at a level of 70%. Next the raters reread and systematically deleted successive quartiles of the idea units (judged as the least important) until only the top quartile (composed of the most important idea units) remained. A score of four was assigned to the most important idea units, and the least important received a score of 1, with scores in between assigned according to their ranked importance. Summaries could then be scored on the basis of the relative importance of the ideas they contained.

As a measure of the operations the subjects employed when summarizing the text, Thomas and Bridge devised a scoring system to rate the degree to which the recalled unit preserved the meaning of the original unit. A score of three was assigned to those idea units which were verbatim recalls or good paraphrases of the unit. A zero was assigned if the unit could not be matched to any units within the original text, with scores of one or two assigned to variations in between. Finally, the two scores, rated importance and degree of match, were multiplied together to arrive at a total score; a score of 12 would be maximum for each unit in this type of scoring.
system.

A comparison of the subjects' cloze scores with their summarization scores was made using a Pearson product moment correlation. The analysis revealed a high ($r = .80$) correlation for the two measures, thus lending credence to the validity of summaries as a measure of reading comprehension. In addition, the high cloze-scorers tended to recall proportionally more of the highly-rated idea units than did the subjects who had low cloze-scores. This study added to the body of knowledge about summary as a comprehension measure, but again, the possible effects of prior knowledge, topic interest, reading ability, and writing ability were not partialed out in the analysis.

In a study with mature, experienced readers at the graduate school level, Smith (1985) looked at strategies and behaviors that readers use while comprehending a difficult text. Of particular interest was the way in which mature readers monitored their own reading behavior. Summaries were used as the dependent measure in an effort to reveal subjects' actual strategy use and their awareness of it.

An unreported number of graduate students (both masters and doctoral level) read a text from Scientific American entitled "Particles of Naked Beauty" (Mistry, 1983) which concerned theories of subatomic particles.
The article was 10 pages long, including illustrations. A prereading assessment determined that only one subject had any background in physics, and none of the subjects were in the habit of reading science journals.

Subjects were given three weeks to complete a series of tasks. They were to read the article and carefully keep a record of their own thoughts and behaviors while reading or preparing to read. In addition, they were asked not to confer with one another, but other resources could be used. Writing assignments, in addition to the log, included a summary and a schematic map.

After the three-week period, readers reported a variety of strategies for dealing with the text. A few sought additional help, but most seemed to "bootstrap" their way through the text alone. Smith felt that this technique was a result of the fact that subjects did not get to choose their own text. Their lack of interest in and prior knowledge of the topic interacted in such a way that they were unwilling to invest the time and energy necessary to employ "scaffolding" strategies to build understanding.

The summarizing task was criticized by some of the subjects who noted that they were able to write reasonable summaries of the article without any real understanding of it. This observation could probably be made in any study where the text is available for summarizing, but it would
be much more difficult to write a coherent summary without understanding if the text were not provided. The present study will make text unavailable during the summarizing task, which should help control for this situation as a possible limitation. Having the text available also seemed to prevent readers from deviating from the surface structure of the text. One subject expressed an unwillingness to try to restate ideas in her own words, given the poorly-developed state of her comprehension of the material. Nevertheless, based on pre- and post-reading statements from the subjects, Smith felt that measurable comprehension growth had taken place.

Smith concluded that perception of the ownership of a task can make noticeable differences in the amount of comprehension that takes place while reading. The subjects' expressed lack of knowledge or enthusiasm for the topic had a definite effect on their ability to read, understand, and summarize. This observation would seem to argue for the assessment of both prior knowledge and topic interest before reading takes place in order to account for their effects upon the quality of the summary that is subsequently written.

To summarize, several studies have been examined that used a summary or a free recall as a dependent measure. Increasingly, this seems to be a trend in reading comprehension research. However, the more widespread use
of summaries has led to some concerns which were investigated in the present study. Specifically, these are: (a) prior knowledge, (b) writing ability, (c) topic interest, and (d) summarizing ability. These will be discussed in the sections which follow.

Factors Affecting Comprehension

Prior Knowledge

In light of the implications of schema theory, it would be difficult to read a summary and conclude that all the information it contained came from reading the target passage. The effect of prior knowledge on recall has been well-documented and has always presented a problem in comprehension research (Farr, 1969; Johnston, 1983). For example, it cannot be concluded that answering questions about a passage is dependent only upon information obtained from reading the passage. Preston (1962) found that 77% of subjects were able to correctly answer questions on a reading comprehension test at greater than chance level even before reading the target passage. Marks and Noll (1967) have suggested that when such results occur, it is highly probable that the items are measuring something other than reading comprehension; Farr (1969) suggested that, in fact, prior knowledge is what is being measured. Tuinman (1973-74) concurred, pointing out the limitations of questions on standardized tests for measuring comprehension; they may, in fact, be measuring
prior knowledge. Johnston (1983) has even suggested that it might be pertinent to purposely insert some passage-independent questions in an instrument for the purposes of assessing prior knowledge. Such a move would shed light on just what the readers' background knowledge might be and how and when they make use of it.

Similar phenomena may occur when a summary is written. Rumelhart (1980) has commented that, once the reader has obtained meaning from a passage, we may not be able to differentiate between information obtained from the sensory input of the text and the information that is a product of the reader's interpretation based on prior knowledge. Such an integration of old and new is undoubtedly taking place when readers reconstruct the text in summary form. We are assuming here, however, that the reader possesses the appropriate background knowledge to allow for a consistent interpretation of the target passage. This assumption, in fact, may not be the case, and the lack of appropriate schemata has been shown to have a detrimental effect on comprehension and recall.

Bransford (1982), for example, gave subjects a short text in the form of a letter from one friend to another without sharing with them the topic which the writer was communicating. While the language in the letter was easy to read and contained no unfamiliar words, readers found that they could not understand it. When they were told
that the letter was about the character's efforts to convince an annoying neighbor to move away, readers then had no comprehension problems.

In a classroom study, Stevens (1982) provided students in social studies with background information on the Texan War. Another group received prereading information on the Civil War, a topic which was unrelated to the target passage. When both groups read the target passage about the Battle of the Alamo, the group that had the background knowledge on the Texan War significantly outperformed the other group on a posttest using questions. Stevens concluded that the appropriate background knowledge accounted for the comprehension and recall differences.

Some readers, therefore, may not possess the relevant prior knowledge needed to understand a particular target passage, and their summaries reflect this lack of knowledge. This is not to say, however, that these same readers could not comprehend at a more satisfactory level if given a passage for which their prior knowledge was sufficient.

Not only may the appropriate schemata be lacking, but it may be that the author did not provide sufficient cues to the reader to activate that knowledge (Rumelhart, 1980). The well-known "washing clothes" text employed in a study by Bransford and Johnson (1972) presents a useful
example. Fifty high-school subjects were given a passage to read either with or without the title "Washing Clothes". The wording of the passage was intentionally vague, and readers not supplied the title were left to their own devices for interpreting it. Passage recall was enhanced only for the condition in which the title was supplied. While most adult readers possess the "washing clothes" schema necessary for successful interpretation of the passage, they are unable to do so simply because the author has not provided enough clues for this schema to be instantiated, or called into play.

Additionally, readers may possess schemata for the topic that are either inaccurate or incomplete and may use this information to interpret the passage in a way that is consistent for them but fails to fully grasp the author's intended message (Rumelhart, 1980). Such was the case in studies by Lipson (1982; 1983) and Alvermann, Smith, and Readence (1985) which found that readers' prior knowledge could have an inhibitive effect and could, in fact, interfere with their understanding of the text.

Lipson (1982) pretested 28 third-grade subjects' knowledge of a topic. One week later subjects read the target passage and answered questions on a posttest. Lipson found that subjects were more likely to answer postquestions correctly when they had been correct on the corresponding pretest question and when they did not know
the pretest question. They were more likely to answer the
correctly if they had incorrectly
answered the related item on the pretest, suggesting that
their inappropriate prior knowledge interfered with their
ability to read and learn from the text.

In a related study, Lipson (1983) examined the
effects of religious affiliation as a source of cultural
background knowledge. Subjects were 32 fourth-, fifth-,
and sixth-graders who attended either Catholic or Hebrew
parochial schools. Subjects read either a culturally neutral passage or one of two other passages that were structured to be specific and familiar to members of the Catholic and Jewish faith. Performance on posttests was similar when children read either the neutral or specific passage. But when the information was only partially familiar, as when a Catholic subject read the Jewish passage or vice-versa, subjects tended to disregard the unfamiliar information, and comprehension was negatively impacted. Again, Lipson concluded that inappropriate prior knowledge could have a detrimental effect on comprehension and recall.

In a more recent study, Alvermann, Smith, and
Readence (1985) had 52 sixth-grade subjects either activate or not activate what they believed to be relevant background knowledge prior to reading a naturally-occurring science text. Two passages were used: one which
was compatible with the students' prior knowledge and one which contained information which was incompatible with their assessed background knowledge. Postreading measures of comprehension consisted of a free recall and multiple-choice questions.

Findings showed that subjects who activated prior knowledge before reading the incompatible text allowed this knowledge to interfere with their comprehension of the text. They actually relied more on their previous information on the topic than on the new information they were receiving and as a result, failed to learn from the new text. There was no difference between activators and non-activators on the compatible text. While the results seem to run counter to the generally-accepted practice of activating prior knowledge in order to enhance comprehension, the authors concluded that their findings lented weight to the idea of assessing students' prior knowledge so that misconceptions might be discussed and rectified prior to reading new information.

Thus, for the purposes of comprehension research in which a single summary is the dependent measure, it appears that we should take into account the effect that the readers' background knowledge might have. This study will attempt to accomplish this by assessing the readers' prior knowledge of the target passage before reading takes place so that any such effect may be considered in a
statistical analysis of summary scores.

**Topic Interest**

Second, it seems important to examine what effect topic interest might have on the ability to write a summary. Subjects in Smith's (1985) study, for example, expressed the opinion that the passage selected for their use in the study was boring and uninteresting; not surprisingly, their comprehension and recall was also very poor. It seems intuitive that students will understand better a topic that they like to read about, and the findings of several studies have borne out this prediction (e.g., Asher, 1980; Baldwin, Peleg-Bruckner, & McClintock, 1985). In addition, our notions about the usefulness of this concept have developed to the point where caution about the facilitative effects of high-interest materials on readers' comprehension has almost become a byword in the instructional literature (Niles, Baldwin, & Wood, 1984). A closer look at some of the studies which helped to bring about this trend seems in order.

Belloni and Jongsma (1978) used 12 stories that had been categorized, based on previous research, as either high- or low-interest topics for adolescents. Further categorization took place based on sex differences in topic interest. Subjects were 50 seventh-graders who were poor readers, and the stories used in the study were all two to four grade levels above their reading abilities.
They were asked to peruse the titles of the stories and choose the story which they would least like to read. In addition, they read a story whose title appealed to them. After reading each of the two stories, a cloze test, with 50 deletions per story, was administered. Results revealed a facilitative effect on the comprehension of reluctant readers when they read stories about a topic for which they expressed interest. In addition, subjects were able to successfully read stories that should have been above their frustrational level because of the intervention of topic interest.

A later study by Asher (1980) had 66 fifth-grade subjects use a picture-rating scale to indicate high or low interest in several topics. When subjects read stories related to the topics of pictures in which they had indicated high interest, their comprehension was higher than for the low-interest topics.

Asher's picture rating scale was employed in a subsequent study by Stevens (1981). Using 93 fifth- and sixth-grade students, Stevens added a Likert scale questionnaire presenting the same 30 topics represented by the Asher pictures. Means for each of the topics were calculated and ranged from a low of 5.57 (for skyscrapers) to high of 12.08 (outer space). Differences between rank levels of the topics were significant, and half of the 30 topics produced sex differences. Stevens concluded that
reading interests of children at this level varied widely and must be accounted for in planning instruction.

An important criticism of some of these studies concerns the problem of separating topic interest from prior knowledge. The question that arises concerns whether readers comprehend high-interest material better because of their desire to read about it or because they have greater prior knowledge about topics for which they express higher interest. Guthrie (1981), in a review of studies on interest and reading comprehension, concluded that children probably comprehend high-interest material better simply because they know more about it. His assumption is that prior knowledge is highly correlated with interest and that prior knowledge is the causal factor. Pearson and Johnson (1978), however, have suggested that interest itself is a factor in reading comprehension.

A more recent study by Baldwin, Peleg-Bruckner, and McClintock (1985) sought to tease out these effects. Subjects were 52 high-achieving seventh- and eighth-grade students. They were asked to complete an interest inventory, rating their interest in each of the 10 topics on a scale of one (lowest) to 10 (highest). The topics on the inventory corresponded to 10 passages selected from Britannica Junior. Next they took a 100-item multiple-choice pretest on the information contained in the
passages on the ten topics. Questions were passage-independent and factual in nature. Finally subjects read a target passage, taken from the same encyclopedia entries but located under different subheadings than the sections from which questions were derived. Additionally, the passage selected for each subject represented one of the following four conditions (L refers to low interest/prior knowledge and H refers to high interest/prior knowledge): HH, LL, HL, or LH. After reading, subjects answered 10 multiple-choice posttest questions which were passage-dependent.

Results of the repeated-measures ANOVA revealed that comprehension was significantly better for passages with high prior knowledge than for low prior knowledge, high topic interest was favored over low topic interest, and boys had higher comprehension than girls. Follow-up analyses indicated that both prior knowledge and interest were independent factors in reading comprehension and that, in fact, the two variables were virtually uncorrelated (Pearson r = .09). It seems to follow from these findings that when subjects indicate their comprehension of a passage by writing a summary about it, their interest in the topic will have an effect on the quality of the summary that they write. Thus, this study will examine the effect of topic interest as a separate predictor, over and above prior knowledge, of summary
Third, the role of general writing ability in writing summaries must also be taken into account. Because answering questions and writing summaries place very different demands on readers (Johnston, 1983), it would seem illogical to qualitatively compare performance on these two very different types of measures and conclude, as did Taylor (1984), that poor summary writers have faulty comprehension. It may be, in fact, that these subjects were simply poor writers. The existence of poor writers at the middle-school level is, unfortunately, borne out by the literature in this area. Opportunities to write have been shown to be an effective means of improving all the language arts, including reading comprehension (Tierney & Leys, 1984). Indeed, real language growth develops in an environment where children are encouraged to make connections between their lives and what they learn in school (Kirby & Kirby, 1985). Writing about their life experiences can have this kind of an effect, and regular writers become true authors as they develop the "eye of a writer", that is, a sense of audience necessary for communicating in written language (Calkins, 1983).

Unfortunately, however, all too often this opportunity to write does not take place. Every recent
survey of writing instruction in American classrooms has reached the same conclusion: children simply are not given the time to write (Anderson, Hiebert, Scott, & Wilkinson, 1985). For example, one recent study of writing instruction at grade levels one, three, and five revealed that less than 15% of the school's instructional time was devoted to writing (Bridge & Hiebert, 1985). Two-thirds of the writing that did occur consisted of word-for-word copying in printed workbooks. It appears that writing compositions of a paragraph or more in length is rare, even at the high school level (Applebee, 1981). There are those who equate writing instruction with more grammar instruction, and these persons assume that such teaching will produce better writers. Studies over the years, however, seem to have shown that this is not the case (e.g., Postman, 1967). The fact remains that while students in our schools will only learn to write when given time and instruction in developing their skills for communication (Calkins, 1983; Graves, 1982), such conditions simply do not exist.

Given the dismal state of students' opportunities to organize and present their thoughts on paper, it would seem logical that an inability in this area would have profound effect on summary writing. Johnston (1983) pointed out that since expressing ideas in writing and organizing information from memory are skills that are not
normally taught to children, the nature of such skills may be very different from reading comprehension. Subjects for whom writing is difficult may have very different criteria for deciding what is important to write and would probably write less information than would skillful writers. In addition, poor writers often have difficulty assuming the perspective of their future readers and thus may fail to include well-understood information that they believe is so "obvious" as to be unimportant. This should not necessarily be interpreted to mean that they had poor comprehension of the target passage, however. It may simply be that writing a well-organized summary presents such a difficult task that we are sampling only a fraction of what readers actually obtained from the reading. In order to make reasonable inferences about reading comprehension from a written summary, it seems necessary to assess the subjects' general writing ability and to statistically account for its effect on the summary writing score.

Summarizing Ability.

Overview. Finally, the importance of summarizing ability must be addressed. The studies that have used summaries as a comprehension measure have done so at various age levels, from grade one (e.g., Johnson, 1983) to graduate level (Smith, 1985). They have also used directions ranging in specificity from "read this passage
and write a summary" (e.g., Thomas & Bridge, 1980) to defining the term for the students and showing them an example (Taylor, 1984). The problem that ensues is that these researchers have assumed that the subjects know how to write a summary. There is a great deal of evidence to the contrary.

For example, even on the very rudimentary level of locating main ideas, Otto, Barrett, and Koenke (1969) found that only 29% of second-graders could select the main idea from simple passages. In a related study, Danner (1976) compared second and sixth graders and found that neither age group was able to make effective use of text structure in comprehension. When passages were intact or scrambled, subjects could recall appropriately only the intact passages. The younger subjects were able to recall in patterns similar to those of the older readers, but only the sixth graders could select appropriate topic sentences for passages from a list of sentences provided. Brown and Smiley (1977) found a similar developmental effect for supplying topic sentences; younger students (ages 8, 10, and 12) could recognize topic sentences, but only older students (age 18) were able to supply their own.

The overall finding of this group of studies is that summary-writing ability is late-developing and does not develop automatically. This conclusion has led to a
second group of studies, namely those that have sought to identify the strategies that readers use when they attempt to summarize. These will be outlined in the section that follows.

Studies examining strategy use. Johnson (1983) examined the development of children's ability to retell stories, particularly their ability to tell only the gist of the story. Of particular interest were the operations which children used for reducing the story to a "tellable" form, that is, a summary. Johnson distinguished between a plot summary and a gist because of observed differences in the amount of information deduction required. For the purposes of the study, Johnson defined a summary as a plot summary or a statement of the point or moral of the story.

Based on some of the extant theories of story grammar (Mandler, 1978; Mandler & Johnson, 1977; Stein & Glenn, 1979), Johnson reasoned that even young children possessed enough knowledge of story structure to guide their attempts at summarization. It was the deliberate manner in which they used this story schema that was of interest. In particular, Johnson was interested in the differences which might exist between children's ability to summarize as opposed to a request for a recall. Because of the developmental effects of the amount that children can recall (Glenn, 1978; Mandler, 1978; Stein & Glenn, 1979),
Johnson believed that a small number of propositions recalled could have an adverse effect upon the number of propositions available for reduction into a summary, and hence, on the quality of the summary itself. Additionally, children's lesser recall may indicate a less integrated representation of the information in the story. The limitations of working memory for retrieving and holding propositions while manipulating them into a summary may be responsible for these difficulties.

In the study under discussion, children in first, third, and fifth grades and adults were asked both to recall and summarize stories. Johnson sought to examine whether or not young children were able to summarize, the relationship between the story structure and the information selected for summaries, and whether or not the children and adults used similar strategies while summarizing.

The familiar "Three Little Pigs" story, as well as a story about a rabbit, were used in the study. In addition, "Goldilocks and the Three Bears" was used to demonstrate the recall and summarization tasks. Tasks were counterbalanced to control for order effects. Subjects were instructed to either "tell everything you remember, the whole story" (Johnson, 1983, p. 349) for the recall task, or to tell just what the story was about for the summary task. All retellings were tape recorded and
transcribed. After subjects had recalled and summarized each story, they were asked to produce a "shortest possible" summary by "telling the story the very shortest way that you can" (Johnson, 1983, p. 350).

Two kinds of approaches were used in scoring the recalls. First, the text was parsed into idea units based upon expert raters' dividing the text into verb-based statements. Two judges jointly classified idea units in the recalls using a two-level criterion as to whether a particular unit represented a "definite" or "ambiguous" representation of a macroproposition of the story.

In addition to this subject-determined scoring, a rule-based criterion was used. Four basic categories were employed: (a) representation of a single idea unit within the original story; (b) using a single idea unit to substitute a generalization for the original information; (c) representation of a single idea unit by noting that it was redundant to previous units; and (d) representation of more than one idea unit into a single unit by combination. Finally, each statement or group of related statements was given a quality score for how closely the original information was maintained.

Johnson found that the mean number of words decreased as subjects moved from recall to summary to the shortest summary task, and length of protocols increased significantly with the age of the subjects. Additional
findings were that length of summaries produced by adults was shorter than either the third- or fifth-graders but did not differ significantly from that of the first-graders. Overall, fewer idea units were represented in summaries than in recall, indicating that some reduction of information was taking place. A developmental trend in the number of idea units was also evident. Adults, then, were able to use fewer words while expressing more idea units, an indication of their superior ability to summarize efficiently. In terms of propositional transformations, adults used far more deletions, generalizations, and combinations than did any of the children.

Johnson's discussion stated that all groups had been able to summarize at least to some extent, although a clear developmental progression was apparent. Children were clearly relying more heavily on deletion strategies than were the adults, who were able to use higher-order strategies, such as generalization and combination, more successfully. Major differences between children and adults seemed to be explained by differential abilities in selecting important information for summary inclusion and formulating a concise representation of that information.

In a subsequent study, again using narrative text, McGee and Tompkins (1982) sought to further explore developmental trends when readers were processing
information into summaries. Thirty-six each of second- and fifth-graders and adults participated in the study. Subjects were tested individually. They first listened to an experimenter read a story aloud and then present a summary of that story. After some instructions as to what constitutes a good summary, each subject listened to a stimulus story and summarized it into a tape recorder. A second story was summarized in a written mode. Order of story presentation and summary mode were counterbalanced.

The stories were first parsed into story propositions using a story grammar proposed by Stein & Glenn (1979). Next, using methods described by Johnson (1983) and Winograd (1982), two scorers jointly parsed each summary into verb-based statements. Then, summary statements were classified into five summary operation categories: (a) reproductions, (b) inferences, (c) combinations, (d) generalizations, and (e) distortions. Reproductions, combinations, and generalizations were as described above in Johnson (1983). Inferences were statements that expressed ideas implied, but not explicitly stated, in story text. Distortions represented statements that were either incorrect or not implied by the story text. In addition, statements were classified into a story category according to Stein & Glenn (1979).

The proportion of statements classified into each of the summary operation categories was calculated, and an
analysis of variance was conducted with mode and operation as within-subjects variables and age as a between-subjects variable. Since distortions were not at all widespread in occurrence, they were dropped as an operation category.

A significant main effect for operation was found, with significant interactions between age and operation and between mode and operation. These results showed that adults used a greater proportion of reproductions, generalizations, and combinations, not unlike the findings in the Johnson (1983) study. Children tended to rely more heavily on reproductions than on any other operation. Mode tended to have an effect on number of statements reproduced, with oral summaries producing more reproductions than written summaries. In contrast, written summaries produced proportionally more generalizations than did oral summaries. Finally, subjects at all ages tended to use more higher-level transformations in their written summaries than in their oral summaries.

In their discussion of the findings, McGee and Tompkins noted that the earlier finding that children were able to use some of the same complex summarizing operations as adults was substantiated. In addition, developmental trends found by Johnson (1983) were again evident. These were explained using findings from studies by Brown and Day (1983) in that young children most often
use a copy-delete strategy. They first decide what they want to include in their summaries and then reproduce it from the text. Older readers, on the other hand, are more able to utilize transformations such as generalizations and combinations in order to successfully reduce the text.

In addition, children may perform less well than adults because of differing language abilities. Their vocabularies are not as well-developed as adults', nor do they possess the language maturity necessary for formulating well-stated complex sentences. Smaller amounts of world knowledge can also be a contributing factor, and children, because of their fewer experiences, may feel more bound by the surface features of the text than do adults.

In summary, McGee and Tompkins concluded that developmental trends in summary operations do exist, and that further, the mode of summarizing must be considered as a methodological factor when trying to draw conclusions about the cognitive processes that summary writers are using.

To further examine the developmental trends that appear to exist with regard to children's summarizing abilities, Garner, Belcher, Winfield, and Smith (1985) designed a study to find out what fifth-grade students can do. More specifically, they looked at the three
components of summarization: (a) recognition of good summaries, (b) production of good summaries, and (c) reflection upon adept production. Good and poor fifth-grade readers were asked to read a five-paragraph, 213-word text from a basal reader in use in the sampled schools. The passage topic was meteoroids, meteors, and meteorites, and the title supplied together with the content of the opening statements was deemed sufficient to activate whatever world knowledge on the topic that the readers possessed.

After reading the passage, subjects wrote a summary of the text material with the passage available. Five minutes were allowed for this task. Next the subjects were prompted to talk about how they went about writing their summaries and what text manipulations were made in order to do so. Both text and summary were available as retrieval clues. The final task was to watch a videotape of three fifth-grade students sharing their summaries and the strategies they used in writing them. Subjects were to rank the summary performance of the students in the tapes.

Summaries were scored by a method first described by Garner (1982) and known as efficiency of summarization. After having a panel of doctoral students rate the importance of the idea units within the text, the summaries received a score based on the sum of the points
earned for the idea units included and weighted by their relative importance. This sum served as the numerator in a fraction, while the total number of words in the summary served as the denominator. Thus, the resulting fraction represented efficiency, i.e., the ratio of the number of rated-important idea units to the total number of words employed. Verbal report protocols were scored by comparing them to summarization rules as outlined by Hare and Borchardt (1984). A point was awarded for each match with the recognized rules. Interrater agreement for both the summary task and the report task was .90.

Finally, the summary selection task was scored by noting how closely the subjects' rankings matched the actual quality of the summaries. Subjects were assigned a score of 0, 1, or 2 based on this criterion.

As hypothesized, both the good and poor readers were able to achieve some degree of success in recognizing the quality of the sample summaries. However, differences between good and poor readers clearly emerged when the production and strategy-report tasks were considered. While some of the good readers produced adequate summaries according to the rules they themselves had identified as important, most really could not perform the task. It seems that fifth-graders have a sense of what a summary requires, but not all are equally adept at providing one. Garner et al concluded that if students at the fifth-grade
level are to be asked to write summaries and other essay-type responses, then it will be necessary to provide explicit instruction in rule-driven summary production.

In a study whose purpose was to explore the use of summaries to assess reading comprehension, Taylor (1984) had sixth-grade students read and study a five-page section in their social studies textbook. First, they were asked to write a summary of the passage. It was explained that a summary includes the main ideas and important details of the passage, and an example was provided. Taylor did not specify in the paper whether or not the text was available for the summary-writing, but because of the "read and study" directions, it is assumed here that the summaries were written from memory. After the summary was completed, the subjects answered 15 short-answer questions on the passage. The assumption is made again here that these questions were answered from memory in order to keep the tasks parallel.

To score the summaries, Taylor first parsed the textbook passage into idea units and then rated the importance of each as either a main idea, an important detail, or an unimportant detail. When scoring the summaries, two points were assigned for each main idea included, one point for each important detail, and zero for unimportant details. Two other raters also assessed the summaries so that interrater agreement was
established. Taylor counted the total number of idea units as a free recall score, and the number of points assigned for relative importance constituted a summary score. Findings were that while two protocols might be quite different when scored for total idea units (as a free recall), they might be more accurately assessed and equitably compared when scored as summaries. Taylor concluded that the summary scoring employed in this study provided a better description of the subjects' protocols and was, thus, as a measure, superior to free recalls.

Taylor also made a qualitative comparison of the summary scores and question scores for each subject. What she found was that while two subjects might have the same score on the questions, thus leading to an assumption of equal comprehension, their summary scores might differ widely. This led Taylor to conclude that the summary was a more sensitive measure of comprehension than the questions. No statistical analysis of the data was offered by Taylor, so it is difficult to make inferences about what might occur with other samples of subjects. While Taylor acknowledged that some measure of writing ability might be included and used as a covariate in an analysis of the data, she did not do this, nor did she attempt to account for any prior knowledge or topic interest that the subjects might have had. It seems that these would have strengthened the study, but a subsequent
study (Taylor, 1985) attempted to address these problems.

In that study, Taylor (1985) sought further to identify the aspects of summary-writing that appeared to cause difficulty for sixth-grade subjects. The study compared written versus oral summaries to determine whether the writing task was too difficult for students at this grade level. Next, free recall and probed recall summarizing tasks were compared in order to determine whether students could recall in a free, unstructured way. Finally, students' summaries were compared with those of adults to detect the students' sensitivity to text structure and to the relative importance of ideas they chose to include.

Subjects were 45 sixth-grade students who were average or above-average readers. Two passages from the subjects' social studies text were used, comprising five pages and 133 idea units and three pages and 65 idea units, respectively. A panel of judges selected one-third of the idea units as the most important. Passages and tasks were counterbalanced for the study.

In the first session, subjects read one of the two passages and either wrote a free recall or a probed recall first. During the second session, subjects read and summarized the other passage and completed the tasks in the condition not yet encountered. Two other sessions were used in a similar way for the oral summaries.
In each session, subjects were told to read and study the passage. All recalls were given from memory, and oral recalls were tape recorded. Undergraduate college students were also asked to read and summarize either passage A or B to provide a template for comparison with the sixth-graders' summaries.

Summaries were scored both for total number of idea units and for number of important ideas recalled. A second rater scored the protocols, and interrater reliability was calculated at .94.

Results lended little support to the notion that writing as a production task caused difficulty for the students, since significant differences were not found between written and oral recall scores. However, it seems plausible that since writing an essay or a well-integrated summary requires organization of ideas not necessary in a free recall (either oral or written), valuable information could still be obtained from a pretest of the students' writing ability.

In comparing the sixth-graders' summaries with those of the college students, Taylor found that the younger students included fewer ideas and failed to follow the structure of the text when summarizing. In addition, they failed to include main idea statements to the extent that college students did. Based on these findings, Taylor recommended that since understanding and summarizing text
seem to be difficult, yet necessary, for children of this age, instruction in text structure and summarizing techniques seems warranted.

A further examination of strategic difficulties in summarizing was conducted by Winograd (1984) using good and poor eighth-grade readers. Winograd selected summarization as the dependent measure because of its value in revealing comprehension processes (Johnson, 1978; Johnston, 1981; Kintsch & van Dijk, 1978). Of particular interest were the dimensions of task awareness, sensitivity to importance, and use of summarization rules. Seventy-five eighth-graders and 40 adults participated in the study.

Eighth-grade subjects were first tested to determine their decoding accuracy and speed to make sure that these were not limiting factors in their overall comprehension. Next, they read an article from typical eighth-grade level materials and answered form memory five multiple-choice questions concerning it. After the questions were completed, they wrote a 60-word summary of the article with the text available. After a distractor task, subjects rated the relative importance of each sentence in the passage. Finally, given a vertical listing of all the sentences in the passage, subjects selected the five most important ones. The entire process was repeated with each of seven additional passages. Despite the fact that the
the multiple-choice comprehension questions could have cued the subjects for the summary, only the summary and sentence-rating tasks were counterbalanced.

Summary protocols were scored using a system from an earlier study (Winograd, 1982) designed to identify transformations of ideas from the original text. Four broad categories were used: (a) reproductions (copying or paraphrase of sentences); (b) combinations of two or more sentences; (c) run-on combinations, in which several sentences were combined but in a less-organized way, and (d) inventions, where sentences were produced that expressed implicitly-stated main ideas from the text. Interrater reliabilities ranged from .89 to .96. No decisions were made at this time as to the relative importance of the ideas included.

Analyses were performed on the following sets of data: (a) subjects' responses to interview questions, (b) importance ratings and selections, and (c) the summaries. With regard to subjects' task awareness, Winograd found that many of the subjects, both good and poor readers, made an explicit reference to the necessity of including main ideas in a summary. He concluded that awareness on the task goal was not a problem for students at this age.

In the area of sensitivity to importance, the good readers' responses correlated more closely with the adults' than did the poor readers'. While not correlated
with adults' judgments of importance, the poor readers' judgments were at least consistent within their group. Follow-up analyses revealed that poor readers tended to select ideas that, although perhaps trivial to the overall meaning of the passage, were interesting and colorful, and therefore, memorable.

A third major finding was that the relationship between information that poor readers considered important and that which they included in their summaries was not very strong. It appeared that position within the text had a large effect on the poor readers' decisions for inclusion. Both good readers and adults, however, displayed a strong correlation in their rated importance and inclusion of ideas into the summary.

In the use of summarization transformations, poor readers were again outperformed by the good readers and adults. Poor readers tended to use more reproductions and run-on combinations, while good readers and adults chose more combinations and inventions. In addition, good and poor readers deleted more ideas than did the adults, but adults used fewer words in their summaries. Winograd concluded that adults were more efficient summarizers, i.e., they were able to convey more ideas in fewer words.

Finally, Winograd looked at the relationship between ability to use appropriate strategies and comprehension of the passage (as measured by multiple-choice questions).
Using a hierarchical regression analysis (Cohen & Cohen, 1975), he entered the predictors into the equation in the following order: IQ score, speed and accuracy of decoding, orthogonal contrasts based upon responses to interview questions, sensitivity to importance, effective rule use, and two-way interactions between each of the main effects.

In the summary score, IQ, sensitivity to importance, and effective rule use accounted for a significant proportion of the variance. When passage comprehension was the dependent variable, only IQ and accuracy of decoding were significant main effects. Significant interactions occurred between IQ and decoding accuracy and between decoding accuracy and a response level contrast. Further examination of these results caused Winograd to conclude that they were due to ceiling effects, since most of the subjects did very well on the multiple choice questions (M = 4.02 out of a possible 5 points).

Variability on the instrument was also rather limited. On the Reading Comprehension Subtest of the Stanford Achievement Test, IQ and decoding speed accounted for a significant proportion of the variance (p < .0005 and p <.005, respectively). In addition, sensitivity to importance was significant at an alpha level of .05.

To summarize Winograd's findings, it is important to note that sensitivity to importance accounted for a significant proportion of the variance in every dependent
measure, even when IQ and decoding ability were taken into account. Second, rule use was a significant predictor only of the summary score, suggesting that strategy use appears to be task-specific rather than a global comprehension skill. Third, task awareness failed to account for production differences, which, given the large number of subjects revealing an awareness of what it means to summarize, is not surprising.

The finding that IQ was a significant predictor for every dependent measure is also not unexpected. While Winograd did not assess the subjects' prior knowledge of the topic outright, many researchers would argue that IQ is really a reflection of prior knowledge (Farr, 1969; Johnston, 1983). This finding lends support to an argument for the assessment of prior knowledge specific to a target topic, which the present study will do.

Winograd concluded that teachers should consider sensitivity to importance as an explanation for comprehension difficulties. In addition, he argued that difficulties in summarizing do not always indicate comprehension difficulties: "Although difficulties with the task of summarization may be symptomatic of comprehension difficulties, summarization difficulties are not necessarily confined to comprehension problems." (1984, p. 423). For the purposes of the present study, this conclusion is the most significant.
In a study with older readers, Garner (1985) explored further the discrepancy between task awareness and production ability. Critical of other studies that used verbal reports to assess task awareness (e.g., Winograd, 1984), Garner had ninth- and eleventh-grade subjects and college undergraduates write both a "good" and a "bad" short summary for the same piece of text. The relative acceptability of the two summaries was based on number of judged-important ideas included, number of words used, and integration of judged-important information.

The target passage was adapted from a Scientific American article entitled "Intuitive Physics" (McCloskey, 1983). The topic was widely-held misconceptions about objects in motion, which was a topic found to be uniformly unfamiliar to subjects at all age levels. Presumably, the low level of familiarity served as a control for any prior knowledge effects. While a measure of topic interest was not mentioned, it can be assumed from previous studies (e.g., Smith, 1985) that subjects probably also found the topic uniformly uninteresting.

A method described by Johnson (1970) was used to parse the text and rate the relative importance of idea units. Graduate students were employed to write optimal summaries so that a criterion for succinctness could be established. Integration level of judged-important ideas was scored on a five-point scale, categorized by the types
of manipulations of the text that were observed. Interrater reliability was calculated as .88.

After a MANOVA test was made to control the overall alpha level within the study, follow-up univariate F tests were made on each of the dependent variables. A pattern of significant differences between the youngest and oldest age groups was noted for idea awareness, idea production, succinctness production, integration awareness, and integration production.

Garner concluded that developmental differences even between these age groups do exist relative to sensitivity to importance. While all the subjects seemed to be aware of the need to include important information, only the oldest students were able to successfully do so. This finding points to production difficulties, rather than awareness, as the reason for poor summaries at the younger age levels. With regard to succinctness, it appeared that none of the subjects was aware of the need for this quality in a good summary, that is, efficient summarizing was not a frequent pattern. For informational integration, only 14% of all subjects were able to use any sort of transformations, and these represented combinations only. While college students outperformed the younger students to a significant degree, even their performance was much less than desirable.

Garner recommended that, since deficiencies in
summarizing abilities are not uniform across ages, neither should our instructional strategies be. She believes that rule-driven instruction would be the most beneficial for high school students, with specific training in locating important ideas and reproducing them as a gist. Learners should also be made aware of the value of integration for achieving succinctness and efficient presentation.

In a study that spanned a wider range of ages, Brown and Day (1983) sought to further clarify developmental trends noted by other researchers. Three studies were undertaken. The first examined summarizing strategies employed by writers in fifth, seventh, and tenth grades, as well as college undergraduates. Two expository passages from seventh-grade geography texts were specially adapted so as to provide easily-discernible opportunities for rule use. Specifically, Brown and Day hoped to locate instances where the following five summarizing rules were employed: (a) deletion of trivial information, (b) deletion of redundant information, (c) superordination of lists, (d) selection of topic sentences where they existed, and (e) invention of topic sentences where they were not present.

Subjects were asked to read the text three times. Next, they were told to write what they thought was a good summary of the text. Finally, they were to put this summary aside and write a second summary, constrained
to a length of 60 words. Both summaries were written with the text available.

Results showed that the deletion rules were applied with equal proficiency at all grade levels. Only the tenth-grade and college students were able to make reasonable use of the superordination and selection rules, and these two groups used the two rules only about half the time. The invention rule was used to an appreciable extent only by the college students and again, it was used only on half the occasions where it was appropriate to do so. Because even the college students demonstrated a need for improvement in their summarization rule use, Brown and Day undertook a second experiment with expert summary writers.

The expert subjects for this study were fourth-year graduate students in the English department at the University of Illinois. Two subjects were selected, based on their interest in the study and their willingness to comply with a talk-aloud procedure in an effort to reveal their actual strategy use. Data collection procedures were identical to those of the first experiment, except that the subjects attempted to talk aloud while writing their summaries. These protocols were tape-recorded and transcribed.

As expected, the expert subjects exhibited almost perfect use of the deletion and superordination rules.
Their use of the selection rule appeared to match that of the college students in Experiment 1, partly because of a limitation of Brown and Day's scoring system. In actuality, the experts tended to invent their own topic sentences by combining information across paragraphs, an event which made use of the selection rule superfluous. Consequently, instances of utilization of the invention rule were quite high.

A more surprising finding appeared as a result of the verbal protocols. The experts exhibited very little metacognitive knowledge of what a summary should be or what their own strategies were for constructing one. Their actual strategy use, however, revealed that they possessed the requisite knowledge and were able to make efficient, although unconscious, use of it. Brown and Day found that the on-line introspections were much more consistent with the subjects' real strategy use, while their retrospections appeared contradictory. These findings confirm other conclusions about verbal report data (e.g., Ericsson & Simon, 1980).

Finally, Brown and Day examined the performance of novice summary writers. This population consisted of junior college students, whose academic abilities were judged to be somewhat less than those of regular undergraduates, based on criteria for admission to their academic programs. Materials and procedures were the same
as for Experiment 1, but subjects were allowed to use scratch paper for notetaking, revision, or any other way in which they found to be helpful.

An analysis of the summaries showed that these subjects were able to use the deletion rules effectively, but utilization of the higher-level rules was much less satisfactory. In fact, their performance lay somewhere between that of the seventh- to tenth-graders but considerably lower than that of the college undergraduates. In the case of the superordination and invention rules, the junior college subjects most closely matched the level of the seventh-graders. Brown and Day explained these results by pointing out how an academically-disadvantaged background can affect subsequent performance where text manipulation is required.

To summarize Brown and Day's findings, it can be said that use of summarization rules presents a very discernible developmental trend. Additionally, students whose academic experiences are impoverished may fail to keep pace with their peers in what would seem to be a normal progression of summarizing skill development. Much like other researchers (e.g., Garner, 1985; Garner et al., 1985; Taylor, 1985), Brown and Day called for specific instruction in summary writing with particular attention given to the rules whose use is found to be deficient.
The last study in this group to be reviewed was conducted with graduate students in a masters program in education. Hidi's (1984) study differs somewhat from the others in this group in that she sought to determine the strategies employed by expert readers when they summarized long, complex texts.

Hidi used as a target passage a 2500-word article which had three distinguishing characteristics. First, the text was extremely complex in its content and structure. Second, its discourse structure differed markedly from the usual formats familiar to these subjects. Finally, the nature of the author's style necessitated frequent inferencing on the part of the readers.

Hidi asked 28 subjects to read the article and write a summary of it in about 150 words. Subjects had the text available for the summarizing task. The average number of words used by the subjects was 172, with summaries ranging from one to seven paragraphs in length.

To evaluate the summaries, two independent researchers examined the 107 sentences of the original text and established 33 of the sentences (roughly one-third) as main ideas. Then, they coded the recall protocols for agreement with the judged-important sentences. Results showed that 90% of all sentences in the summaries represented one or more of the 33 main
ideas, and only 10% of the information represented unimportant ideas. Hidi concluded, therefore, that locating and restating important information was not a problem for these mature readers.

A study of the structure of the subjects' summaries showed that most subjects tended to follow the order of the ideas in the text quite closely, with little integration or combination of ideas. Combinations that did occur tended to be of neighboring ideas, with little effort given to a true synthesis. Nor were the subjects successful at inventing topic sentences where none existed, similar to the findings of Brown and Day (1983). Hidi referred to this type of summary as a precis, rather than a synthesized summary. She reasoned that subjects were afraid to interpret, for fear of making mistakes, and relied instead upon sentence-by-sentence translation of the material.

To summarize, this group of studies has presented a convincing case that middle-grade students, and to some extent, high school and college students, have difficulties with summary-writing. It seems that these students are aware of what the task involves but are simply not able to execute a successful strategy for producing an efficient summary. This has led to speculation by several of the researchers (notably, Garner, 1985; Garner, et al., 1985; Taylor, 1985) that
direct instruction in summary skills is necessary before these students can become adept at the strategies necessary to produce one. While the present study does not include an instructional component, the next section will focus on several of the representative training studies.

Training studies in summarization. Building on the work of studies which had shown that students in the middle grades have difficulty both with text organization (e.g., Meyer, Brandt, & Bluth, 1980; Taylor, 1980) and with expository writing (Bossone & Troyka, 1976; Cooper, Cherry, Gerber, Fleischer, Copley, & Sarlisky, 1979), Taylor (1982) reasoned that such problems could result from a lack of awareness of the text structure. Therefore, Taylor set out to devise a strategy that would work both to improve reading comprehension and expository writing.

The strategy developed was the writing of a hierarchical outline form of the text which could then be used to write a summary of the material. The outline was based on the headings and subheadings typically found within chapters in expository textbooks. The procedure was guided at first by the teacher and consisted of the following steps. First, the students previewed a passage, three to five pages in length, and located the headings and subheadings. These were then written as the Roman
numerals and the capital letters of a skeletal outline, leaving spaces between to fill in with further notes. Next, students read the passage section by section and filled in the skeletal outline. A main idea statement was generated for each subsection and written next to the corresponding capital letter, and two to four supporting details were listed and numbered beneath it. After each subsection was completed, they moved on to the next subsection until an entire section was complete. At that point they generated a topic sentence for the entire section. After all sections had been completed in this manner, students went back over the outline and looked for ideas that seemed to belong together. These were joined by drawing lines between the ideas and labeling with a key phrase the relationship between the ideas. Finally students studied the outlines and practiced retelling them with a partner.

In field testing of the strategy, Taylor found that students using the strategy improved both in their recall of expository material and in the quality of their written compositions (Taylor, 1980). Thus, a strategy designed to improve familiarity with text organization had the additional effect of improving reading comprehension and writing ability.

In an extension of the procedure used by Taylor, Taylor and Beach (1984) taught the hierarchical strategy
to a group of seventh-graders. A conventional group also received instruction and practice in reading and answering questions, while a control received no special instruction. All groups used passages from a junior-high social studies textbook.

Before instruction began, all subjects were given a pretest of writing ability and reading ability as measured by recall and short-answer tests. They were also asked, after reading, to rate their familiarity with the topic before the reading had taken place. After seven weeks of instruction at one hour per week, students were tested again in both reading and writing ability. For the material judged relatively familiar, the summary instruction proved effective over both the question-and-practice treatment and the control. However, for the relatively unfamiliar material, the summary training was not effective over the question practice treatment.

Bean and Steenwyk (1984) compared a rule-governed approach to summarization with a more intuitive strategy, GIST, and a trial and error control group. Both the rule-governed and GIST strategies significantly improved sixth graders' summary writing and reading comprehension. The authors reasoned that both approaches were effective because they both embody features of direct instruction associated with high student achievement (e.g., modeling, application, and feedback).
In a replication and extension of the earlier study, Bean and Steenwyk (1985) added a hierarchical mapping strategy in place of the trial and error group. They predicted that all three groups would experience growth in summarization and reading comprehension. Based on the previous study, they further predicted that there would be no significant differences between the three groups in summarization and reading comprehension. Contrary to their hypotheses, however, Bean and Steenwyk found that the mapping group significantly outperformed both the rule-governed and the GIST groups on the comprehension measure. The authors concluded that the visual effects of the mapping structure might contribute to its effectiveness. Additionally, they observed a noticeable enthusiasm for the strategy from the students, which they reasoned could also have contributed to its success.

Two studies have examined the effects of training studies with college students. In the first of these, Day (1980), trained low-ability community college students using a rule-driven approach. The treatments differed from one another in the way that they were integrated with various self-management or monitoring strategies. The first group received self-management alone, simply utilizing a self-checking procedure to see if their summaries were adequate. The second group received rules alone. These were adapted from Kintsch and van Dijk's
(1978) five summarizing rules: (a) delete redundancy, (b) delete irrelevancies, (c) subordinate subtopics, (d) select topic sentences, and (e) create topic sentences where none exist. The third group received treatments one and two in sequence, while the fourth group integrated the rules and self-management procedures. An instructional design which featured modeling, feedback, and practice was used.

The dependent measure was the proportion of the time that subjects used the summarization rules when constructing summaries. All subjects could use the two deletion rules with no difficulty. On the other three rules, the integration group significantly outperformed the self-management alone group in every case. Additionally, the integration group was significantly better than any of the other three groups on both the subordination rule and the creation rule. Day concluded that, especially for low-ability students, a combination of summary training with usable self-management strategies was the most effective approach that could be taken.

Finally, Hare and Borchardt (1984) sought to extend Day's (1980) findings to low-income, minority high school students. They added two other rules, combining paragraphs, which they felt would make the summaries more nearly like those produced by skilled readers/writers, and a rewriting rule, which should help to produce more
"polished" summaries. Subjects received their training in one of two ways, either inductively or deductively. Four training classes, each lasting two hours, were held.

Pretesting consisted of writing two summaries of 80 words or less, listing rules that a person might use when summarizing, a cloze comprehension test, and a measure to assess sensitivity to importance. Posttest tasks were identical but administered both immediately and after a delay of two weeks. Summaries were scored for efficiency, adapted from Garner (1982), and defined as number of ideas included divided by total number of words. In addition, rule use was rated on a scale of 1 to 3 (no, inconsistent, or consistent rule use).

Results showed that the experimental group significantly outperformed the control for both the summary efficiency and summary process (rule use) scores. Moreover, these effects were maintained over the two-week delay period. Thus, the effects of summarization instruction were found not only to be positive but persistent as well.

To summarize, the studies in this section have looked at how summarizing ability can be improved through training. The researchers responsible for these studies have realized that summary ability is late-developing at best and does not always develop at all without some intervention. Their assumption is that if we expect
students to summarize well, then we must teach them to do so.

Implications for the Present Study

This review has traced the development of our knowledge of summary-writing and the kinds of strategies that readers use to construct them. In particular, an emphasis has been on how summaries are used for assessment of reading comprehension, the major concern for the present study. It would be useful, at this point, to summarize the findings with respect to the factors identified for this study.

With regard to prior knowledge, only Garner (1985) and Winograd (1984) have even hinted at its effects. Garner chose a text that was judged to be uniformly unfamiliar to all subjects, presumably so that prior knowledge effects would be cancelled. Winograd, on the other hand, did not address prior knowledge per se, but used an IQ score in his analyses. According to Johnston (1983), many IQ tests are, in fact, measuring prior knowledge rather than ability, so Winograd's use of IQ may be taken as an indirect acknowledgement of the effect that prior knowledge might have. Winograd's results revealed a highly significant main effect for IQ on every measure, thus lending credence to the argument for partialing out prior knowledge from the summary score in the present study.
Topic interest was not addressed at all by any of the studies, although Smith (1985) included comments from her subjects as to the "boring" nature of the piece she chose for them to read. She acknowledged the fact that self-selection can make a wide difference in readers' willingness to read about a topic, a finding borne out by studies by Asher (1980) and Baldwin, Peleg-Bruckner, and McClintock (1985). These findings, while not related to summary-writing itself, nevertheless cannot be ignored when considering the use of summaries to assess reading comprehension. For this reason, topic interest is viewed as an important predictor of an individual's summary score.

Writing ability has been overlooked to an even greater extent. To this researcher's knowledge, only Taylor and Beach (1984) have looked at subjects' writing ability prior to summary training. This measure was in response to Taylor's own recommendations from her previous study (1984), in which she had suggested that writing ability be assessed and used as a covariate. It seems intuitively clear that, if a student has difficulty expressing his thoughts clearly on paper, then writing a well-integrated summary will be a nearly-impossible task. Therefore, the present study will examine writing ability as a predictor of the summary-writing score.

Finally, summarizing ability itself must be
addressed. While a number of studies described above have sought to explain the strategic difficulties observed in writers of this age, none has looked at all the factors in concert. Both awareness and production difficulties will be examined in comparing free recall and summary scores, and efficiency of rule use will be assessed by scoring the summaries themselves qualitatively. In addition, the relationship between summary score and comprehension questions will be examined in order to present an assessment of the validity of summaries as a comprehension measure.

In sum, the present study will attempt to look at several factors that should be considered when attempting to use summaries to measure comprehension. To repeat, as Winograd (1984) has so succinctly pointed out, "although difficulties with the task of summarization may be symptomatic of comprehension difficulties, summarization difficulties are not necessarily confined to comprehension problems" (1984, p. 423).
References


Journal of Verbal Learning and Verbal Behavior, 17, 229-247.


Journal of Verbal Learning and Verbal Behavior, 9, 12-20.


Mandler, J. M. (1978). A code in the node: The use of story schema in retrieval. Discourse Processes, 1, 14-


International Reading Association, 14, 692-697.


Tuinman, J. J. (1973-74). Determining the passage-dependency of comprehension questions in five major


APPENDIX C

PILOT STUDY
PILOT STUDY

The instruments used in this research, namely the multiple-choice pretest-posttest, interest inventory, writing ability pretest, and summary-writing measure, were pilot-tested with two groups of seventh-grade students from the same school as the subjects in the study. These students, however, were not subjects in the main study. Additionally, four passages were pilot-tested for possible use as the target passage. Related goals were to determine the amount of time that would be required for various facets of the study, reliable scoring criteria for each of the instruments, and what directions would be appropriate for use with the subjects in the main study.

The first pilot study was conducted in a manner similar to the procedures used in the main study. A group of 15 subjects first completed the interest inventory form. Next, they wrote a short essay entitled "McKinley Middle School through the Eyes of a Seventh Grader". After a short break, subjects read one of two passages (A or B) from their regular social studies textbook. In order to test the passage dependency of questions that had been developed for passage A, all subjects answered a set of 26 questions based on that passage. Finally, they read passage C and wrote a summary of it from memory. It should be noted that all passages used in the pilot studies were unrelated to each other in content and were
novel to the subjects.

In a second pilot study, another group of 18 subjects read one of two passages (C or D) and studied them briefly. Next, they wrote a free recall from memory of the passage they had read. Again, for testing passage-dependency of questions, all subjects answered a set of 33 multiple-choice questions based on passage D.

The results of the pilot studies may be summarized as follows:

a) Based on reading times, recall scores, and postquestion results, Passage D was deemed the most appropriate for use as the target passage;

b) Questions proposed for the pretest-posttest proved to be passage-dependent and presented a difficulty range suitable for producing wide variability;

c) The topic chosen for the writing sample was well-received by the subjects and produced essays that fit the scoring criteria;

d) The interest inventory provided suitable variability;

e) Proposed scoring protocols were reliable and easy to use; and

f) Maximum times recorded (ranging from 6 minutes for the interest inventory to 17 minutes for reading the target passage) would allow for data collection in the time sequence originally proposed.
DIRECTIONS: Select the best answer for each question and circle the letter of that answer.

1. When individual workers tried to work for better conditions, they found that
   A) factory owners were usually willing to listen.
   B) they needed special training in order to be heard.
   C) they might be fired and replaced by someone else.
   D) improvement was slow but definite.

2. The Knights of Labor was founded by
   A) Uriah S. Stephens.
   B) John L. Lewis.
   C) Henry Ford.
   D) John D. Rockefeller.

3. Early attempts to form labor unions (in the 1830's) were not successful because
   A) the employers did not want them.
   B) these national unions were too large to be manageable.
   C) their dues were more than many workers could pay.
   D) most workers were farmers who didn't need unions.

4. An old saying became the slogan for the workers' efforts to win better conditions. Choose the best word to complete this slogan:
   "In ______ there is strength."
   A) union
   B) teamwork
   C) variety
   D) work

5. When representatives from both a labor union and the employer sit down to try to resolve a dispute themselves, the process is known as
   A) binding arbitration.
   B) mutual accord.
   C) strike-breaking.
   D) collective bargaining.

6. After some initial success, the Knights of Labor eventually failed because
   A) it was not able to recruit enough members to support its cause.
   B) of strong competition from the American Federation of Labor.
   C) its leadership was unable to handle so many different kinds of workers.
   D) the workers were not really ready for a union of its size and complexity.
7. Which of the following was the chief labor union in this country until the 1930's?
   A) the Knights of Labor
   B) the Congress of Industrial Organizations
   C) the Grange (the National Grange for the Patrons of Husbandry)
   D) the American Federation of Labor

8. Because of increased use of machinery in manufacturing in the late 1800's, the owners of the factory
   A) cared more about their workers.
   B) didn't see as much of the workers as they used to.
   C) increased the wages of the workers.
   D) tried to make working conditions safer for their workers.

9. A factory which hires only union members is called a
   A) closed shop.
   B) union shop.
   C) open shop.
   D) double shop.

10. The American Federation of Labor consisted of
    A) local unions of skilled workers.
    B) large groups of unskilled workers.
    C) workers from the railroad industry.
    D) one large group of skilled and unskilled workers.

11. The leader in starting the CIO was
    A) Cyrus McCormick.
    B) William H. Taft.
    C) John L. Lewis.
    D) Oliver H. Kelley.

12. Which of the following were called in to break up the strike of 1877?
    A) the FBI
    B) the local police
    C) the Knights of Labor
    D) federal troops

13. Strikes are often made more effective by the use of
    A) violent threats.
    B) picket lines.
    C) boycotts of the employers' product.
    D) police protection.
14. Employers in recent years have tried to improve conditions for their workers because
A) the unions have forced them to.
B) the workers have demanded it.
C) the employers realize that contented workers are better workers.
D) the government has passed laws requiring these improvements.

15. Factory workers in the late 1800's were discontented because
A) industry was not expanding rapidly enough to suit them.
B) inflation made their money worth less.
C) working conditions were in need of improvement.
D) they were afraid of the new machinery in the factories.
(POSTTEST)

NAME__________________________________________

DIRECTIONS: Select the best answer for each question and circle the letter of that answer.

1. The use of machinery in the late 1800's changed the lives of factory workers by
   A) requiring them to learn new skills.
   B) motivating them to work harder.
   C) making their jobs less secure.
   D) making their work more interesting.

2. When workers banded together to demand better conditions, they found that
   A) the employers would not listen to their pleas.
   B) the employers couldn't fire and replace them all.
   C) they had more respect for themselves.
   D) they got immediate and successful results.

3. The Haymarket Riot took place in
   A) New York City
   B) Philadelphia
   C) Minneapolis
   D) Chicago

4. The Knights of Labor was a union that
   A) focused on particular crafts.
   B) was open to workers from all fields.
   C) kept out black workers.
   D) recruited members from other unions.

5. The steel industry strike of 1892 took place in
   A) Pittsburgh, Pennsylvania
   B) Homestead, Pennsylvania
   C) Erie, Pennsylvania
   D) Bethlehem, Pennsylvania

6. The American Federation of Labor was founded by
   A) John L. Lewis
   B) Samuel Gompers.
   C) Andrew Carnegie.
   D) William Jennings Bryan.

7. Labor unions in the 1830's were
   A) made up of farmers.
   B) national in their scope.
   C) made up of steel workers.
   D) local groups.
8. Members of the CIO were mainly from the
   A) steel industry.
   B) mass-production industries.
   C) skilled craft unions.
   D) chemical industry.

9. The Strike of 1877 involved workers in which of the following industries?
   A) steel industry.
   B) rubber industry.
   C) railroad industry.
   D) meat-packing industry.

10. A factory in which non-union workers must join the union within a certain period of time after they are hired is called a
    A) closed shop.
    B) union shop.
    C) open shop.
    D) double shop.

11. After the Panic of 1873,
    A) union membership soared.
    B) jobs were hard to get.
    C) inflation used up much of the workers' wages.
    D) the Knights of Labor merged with the AFL.

12. The CIO differed from the AFL in that
    A) it was made up of workers from the auto industry.
    B) it was founded earlier and lasted longer.
    C) its leaders were better-prepared and more responsible.
    D) its members were both skilled and unskilled.

13. The strongest weapon that workingmen can use is
    A) the strike.
    B) arbitration.
    C) employee involvement in decision-making.
    D) complaints to the union management.

14. When a board of fair-minded persons is employed to settle a dispute between two parties, the process is known as
    A) compromising.
    B) arbitration.
    C) peaceful negotiations.
    D) laissez-faire.
15. A factory which employs both union and non-union workers is known as a
A) closed shop.
B) union shop.
C) open shop.
D) double shop.
APPENDIX E

INTEREST INVENTORY
NAME__________________________________

INTEREST INVENTORY

Suppose that you were going to select the topics you would read about in American History. Please rate the following topics according to how much you would enjoy reading about them. On each space give a grade of A, B, C, D, or F based on your opinion. An A means, "It's wonderful; I love it!" An F means "It's terrible! Take it away and bury it, quick!"

THE COLONIAL PERIOD (1600-1760)
  Biography _____
  Human Drama _____
  Family Life _____
  Freedom & Justice _____

THE REVOLUTIONARY WAR ERA (1760-1785)
  Biography _____
  Politics _____
  War _____

THE CIVIL WAR ERA (1850-1876)
  Slavery _____
  Famous Battles _____
  Politics _____
  Biography _____
  Human Drama _____

POST CIVIL WAR ERA (1865-1900)
  Pioneers _____
  American Indians _____
  Tales of the Wild West _____
  Growth of Industry _____
  Development of Labor Unions _____

WORLD WAR II (1939-1945)
  Battles in the Pacific _____
  Military Planes _____
  The War in Europe _____
  Atomic Weapons _____
  Freedom & Justice _____

POST WORLD WAR II AMERICA (1945 TO PRESENT)
  Civil Rights _____
  Women's Liberation Movement _____
  Increased Opportunities in Education & Recreation _____
  Changes in American Foreign Policy _____
  Advances in Science, Literature, Art, & Music _____
  The Presidency _____
  Biographies of Black Leaders _____
APPENDIX F

FREE RECALL AND SUMMARY INSTRUMENTS
On this sheet, write down everything you can remember about the textbook pages you just read. Don't worry about whether the things you remember are important or not; just write them down anyway. What you write doesn't have to be in sentence/paragraph form, but you should express each idea clearly enough that anyone who reads it will be able to understand what you mean. You probably will need about 10-15 minutes for this activity, and it is OK to write on the back if you need to. When you have finished, put the paper back into the envelope and follow any other directions your teacher has given you. Do your best and good luck!
On this sheet, write down a summary of the textbook pages that you just read. A summary includes just the main ideas and important parts that the author was trying to tell you. This means that you shouldn't write everything you remember; just write down the things that you think were the most important. You should write your ideas in complete sentences and in paragraph form. You will probably need about 5-10 minutes for this activity, and it is OK to write on the back if you need to. When you have finished, put the sheet back into the envelope and follow any other directions your teacher has given you. Do your best and good luck!
APPENDIX G

DATA COLLECTION SCRIPTS

150
DIRECTIONS FOR INTEREST INVENTORY - DAY 1

Hello, my name is ______________________________ and I am a student at LSU. I am working on a research project and I'll be working with you for the next few days in your social studies class. Today I am going to ask you to fill out a questionnaire for me about topics you might read about in a social studies class. I'll hand these out and then we can look over the directions together.

(HAND OUT THE PAPERS)

Please write your name on the paper in the space provided. Now follow along while I read the directions aloud.

(READ THE DIRECTIONS)

So you see, you have the opportunity to give grades to some social studies topics. Your teacher will not see your papers, so you don't have to worry about your grade. Just think about each one and do your best. Are there any questions?

(ALLOW FOR QUESTIONS)

When you are finished, just turn your paper face down on your desk and wait quietly. OK, you may begin.

(WHEN ALL HAVE FINISHED, COLLECT PAPERS)

Now I have for you a letter to your parents. It explains all about the research project. Please take it home to them today. It is not necessary for them to sign it.

(HAND OUT THE LETTERS)

Finally, I have some instructions for the next three days. First, you must bring your social studies textbook to class every day. We will be using it. Second, tomorrow it is very important that you bring an ink pen to write with. This is because your paper will be xeroxed, and pencil doesn't show up.

Any other questions?

(ALLOW FOR QUESTIONS)

Then a quick reminder: bring your book every day and be sure to bring a pen tomorrow. See you tomorrow!
DIRECTIONS FOR PRETEST - DAY 2

Hello again. Today we will be doing two activities together. The first will be answering some multiple-choice questions, and the second will be a writing activity. First, let me ask: did you bring your textbook today?

(ASK FOR SHOW OF HANDS)

Good. This was a trial run. We will not actually use the textbook today, but we must have it on Monday and Tuesday. Be sure that you bring it on those days. Did anyone forget to bring a pen?

(ASK FOR SHOW OF HANDS)

Good, because you will need it today. First we will work on the questions. These will be about a topic that you may not know too much about. But we are interested in seeing how much you do know about it right now. Just do your best to answer every question.

(PASS OUT PRETEST)

Write your name on your paper. Now read the directions silently while I read them aloud.

(READ DIRECTIONS. STRESS THE UNDERLINED WORDS.)

It is very important that you answer every question. If you don't know the answer, just try to choose the one you think is best. Remember, don't skip any questions. Please notice that there are questions on the back of the sheet, so don't forget them. Please note that you are to circle the letter of the answer you choose. Are there any questions?

(ALLOW FOR QUESTIONS)

When you have finished, turn your paper face down on your desk and wait quietly. You may begin.

(COLLECT PAPERS WHEN ALL HAVE FINISHED. BE SURE THAT NAMES ARE ON PAPERS)
DIRECTIONS FOR WRITING SAMPLE - DAY 2

Now we will do the writing activity. First I will give you a sheet of paper. You will probably need only this one sheet, but I have extra sheets. If you find that you need it, please just come up quietly and take one.

(HAND OUT PAPER, ONE SHEET PER PERSON)

Please write your name at the top of your paper. (PAUSE) Today you are going to write a short essay about a topic that I will give to you. Although this is not actually for a grade (Ms. ______ and your English teacher will not see it), I want you to write as if you were going to hand it in to your English teacher. In other words, do your very best writing. This means that you will try to organize your thoughts clearly, perhaps making an outline before you begin writing. Paper is available if you wish to do this. As you write, you will want to be aware of the person who will read your essay and how clearly you are getting the message across to that person. Once you have written, you should reread and make any changes that you feel will make your paper better. Before I give you the topic, are there any questions?

(ALLOW FOR QUESTIONS. IF THEY ASK ABOUT LENGTH, TELL THEM THAT THIS IS AN INDIVIDUAL MATTER, BUT THAT ONE PAGE IS ABOUT AVERAGE.)

The topic for your essay is "McKinley Middle School through the Eyes of a Seventh Grader". This is your chance to write what you really think. Remember, only the research team will see your paper - none of your teachers, principals, or parents.

(WRITE TOPIC ON BOARD)

Please remember to do your very best writing. If there are no further questions, then you may begin. Remember, you may come quietly to get extra paper if you need it.

(COLLECT PAPERS WHEN ALL HAVE FINISHED. BE SURE THAT NAMES ARE ON PAPERS.)

That is all we will do for today. I will return on Monday and Tuesday. Again, let me remind you to bring your book and pen on those two days. Thank you for your help, and have a good weekend!
DIRECTIONS FOR READING
AND RECALL/SUMMARY ACTIVITIES - DAY 3

Well, I'm back again. I hope you are all rested from your holidays and ready to help me again. I want to thank you again for your help; so far the results have been very good! Today we will be doing some reading and writing, so please get out your textbook and pen and clear everything else off your desk. While you are doing this, I will be passing out a packet of materials to you. Your name is on the envelope, so please listen for it to be called. When you receive your packet, do not open it until you are told to do so.

(CALL OUT NAMES AND HAND OUT PACKETS. YOU HAVE TWO BLANK ENVELOPES IF YOU NEED THEM FOR ANY NEW STUDENTS.)

Now open your textbook to page 486. You will be reading silently pp. 486-491. This is a section about the development of business and labor in this country. You will notice that it is section 2 in the chapter: just a word about what went on in section 1 might help set the stage for you. A character named Mr. Charles Jackson was presented and described as a typical factory owner in the late 1890's. His "factory" was used to describe some of the problems that industry faced as it developed. So when you read about Mr. Jackson, remember that he is a fictitious character who was introduced earlier.

You are to read all of these six pages except for p.489 and the bottom of p. 491. (Stop when you get to the "Checkup" questions on p. 491) Read carefully because you will be doing some other activities using this material later. You will probably need about 10-15 minutes for your reading. Are there any questions so far?

(ALLOW FOR QUESTIONS)

When you have finished your reading, close your book and put it away. Then open your packet. Find the sheet on top labeled ACTIVITY #1. Take it out of the packet and read the directions carefully. Do what it says to the best of your ability. When you have finished, put it back into the packet. Then take out the sheet labeled ACTIVITY #2 and do it in the same manner. Pay careful attention to the time listed for both activities so that you will be sure of finishing both of them. When you have finished ACTIVITY #2, put it back into the envelope and sit quietly. You will have finished your work for the day. I will write these instructions on the board for you to refer to. Do you have any questions?
(ALLOW FOR QUESTIONS. AFTER THESE ARE COMPLETED, WRITE THE FOLLOWING DIRECTIONS ON THE BOARD:

STEP 2: Do ACTIVITY #1 from your packet and return it to the packet when finished. Watch your time! 
STEP 3: Do ACTIVITY #2 from your packet and return it to your packet when finished. Watch your time! 
STEP 4: Be sure that you have written your name on both sheets. Then sit quietly until the packets are collected. 

WATCH YOUR TIME CAREFULLY ON EACH ACTIVITY!!"
DIRECTIONS FOR RE-READING
AND POSTQUESTIONS - DAY 4

Well, we're almost there! This is your last day to help me - then you can get back to work with Ms. _______. Today we are going to read and answer some questions about what we read. Then I'll give you a short summary of what this project is all about and what you have been doing.

First, the reading. I want you to open your textbook to the same pages as yesterday, pp. 486-491. You are going to re-read these pages, much the same as you would do if you were studying a regular classwork assignment. Again, you may skip p. 489. When you have finished reading, close your book and put it away. I will bring you the questions that you are to answer. They are multiple choice, similar to the ones we had the other day. When you have finished these, raise your hand and I will pick them up. Then please wait quietly until the others have finished. Do you have any questions?

(ALLOW FOR QUESTIONS)

Remember that you will be answering questions, so read carefully. You may begin.

(WATCH FOR BOOKS TO BE PUT AWAY. HAND POSTTEST TO THOSE WHO HAVE FINISHED. MAKE SURE THEY PUT THEIR NAMES ON THEIR PAPERS. PICK UP PAPERS WHEN THEY HAVE FINISHED. WHEN ALL HAVE FINISHED, SAY...)

Now I'll tell you a little bit about this research project. Our team is interested in finding out about how students your age read and learn from their textbooks. In particular, we would like to know about ways that teachers can find out what their students have learned. One way that teachers sometimes use is to ask questions. That is why we asked questions before you read (as a pretest, to see what you knew already) and after you read (as a posttest, to see what you had learned). Sometimes students are asked to tell or write down what they remember from their reading. That is why we asked you to do that. Sometimes teachers ask students to write a summary of what they learned, so we had you do that, too. How interested you are in what you read makes a difference in how well you understand, so we measured that, too. What we will do now is to put all this information together to try to find out what we want to know: the best way for teachers to measure what their students have learned.

Incidentally, if you want to know why we had the
envelopes yesterday, it was because we scrambled the order of the activities: some of you wrote the recall first while others wrote the summary first. The reason for this was that we thought that writing one or the other first could make a difference in how well you did on the other activity. So we scrambled them in order to find out. We sealed them up so that knowing what the other activity was wouldn't have an effect on how well you did on the first task. Are there any questions you would like us to answer?

(IF THEY ASK ABOUT RESULTS, TELL THEM I'LL LET THEIR TEACHERS KNOW.)
January 15, 1986

Dear Parents:

I am a doctoral student in Education at LSU presently conducting research for my dissertation. My study involves seventh grade students and the manner in which they read and study their textbook material. I have selected McKinley Middle Magnet School as the site for my investigation, and your child is a member of one of the social studies classes that will be participating in the study. The study will use their regular social studies textbook and will occur in the classroom setting without undue disturbance of the regular classroom routine. Cooperation and approval have been received from the teachers involved, as well as Mrs. Taylor and Mrs. Williams, Principals.

I am asking your permission for your child to participate in the study. I assure you that the instruction he or she receives will be beneficial and in no way detrimental to the usual learning environment. Please contact the school if you have any objection to your child's participation.

Thank you for your support and cooperation. Classroom research is essential if we are to continue to improve the quality of the education we provide to our children.

Sincerely,

Martha H. Head

Principal

Mildred H. Taylor

Assistant Principal for Instruction:

Josie M. Williams
KEY FOR RAW DATA VARIABLE NAMES

REC..................................Record number
ORDER..................................Summary order condition
   (First = 1)
   (Second = 2)
FREE RECALLS:
    FR/WD2................................Number of words
    FR/IU.................................Number of idea units
    FR/FOURS.............................Number of important idea units
    FR/4/PRO.............................Proportion of important idea units
    FR/EFF...............................Efficiency
    FR/RULE.............................Rule use
SUMMARIES:
    SUM/WD2.............................Number of words
    IU2..................................Number of idea units
    FOURS2..............................Number of important idea units
    SUM4/PRO............................Proportion of important idea units
    EFF2.................................Efficiency
    RULE2...............................Rule use
ADDITIONAL VARIABLES:
    PK2..................................Prior knowledge
    WA2..................................Writing ability
    INT2.................................Topic interest
    CAP..................................Standardized reading comprehension test
    COMP.................................Multiple-choice questions

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>REC</th>
<th>PK2</th>
<th>WA2</th>
<th>INT2</th>
<th>CAP</th>
<th>COMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>62</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>15</td>
<td>5</td>
<td>82</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>15</td>
<td>5</td>
<td>72</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>13</td>
<td>4</td>
<td>94</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>13</td>
<td>2</td>
<td>49</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>9</td>
<td>2</td>
<td>88</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>14</td>
<td>4</td>
<td>97</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>15</td>
<td>5</td>
<td>97</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>72</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>43</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>14</td>
<td>4</td>
<td>58</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>8</td>
<td>7</td>
<td>67</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>19</td>
<td>5</td>
<td>70</td>
<td>9</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>25</td>
<td>3</td>
<td>82</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>21</td>
<td>1</td>
<td>54</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>60</td>
<td>7</td>
</tr>
<tr>
<td>17</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>49</td>
<td>5</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>14</td>
<td>4</td>
<td>92</td>
<td>12</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>13</td>
<td>1</td>
<td>78</td>
<td>10</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>86</td>
<td>8</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>8</td>
<td>2</td>
<td>80</td>
<td>6</td>
</tr>
<tr>
<td>22</td>
<td>5</td>
<td>17</td>
<td>2</td>
<td>47</td>
<td>8</td>
</tr>
<tr>
<td>23</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>88</td>
<td>8</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>15</td>
<td>4</td>
<td>78</td>
<td>9</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>16</td>
<td>8</td>
<td>97</td>
<td>12</td>
</tr>
<tr>
<td>26</td>
<td>6</td>
<td>19</td>
<td>2</td>
<td>69</td>
<td>9</td>
</tr>
<tr>
<td>27</td>
<td>8</td>
<td>21</td>
<td>4</td>
<td>94</td>
<td>12</td>
</tr>
<tr>
<td>28</td>
<td>3</td>
<td>11</td>
<td>4</td>
<td>86</td>
<td>8</td>
</tr>
<tr>
<td>29</td>
<td>3</td>
<td>10</td>
<td>0</td>
<td>53</td>
<td>6</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>15</td>
<td>5</td>
<td>45</td>
<td>6</td>
</tr>
<tr>
<td>31</td>
<td>5</td>
<td>9</td>
<td>6</td>
<td>57</td>
<td>6</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>19</td>
<td>2</td>
<td>67</td>
<td>11</td>
</tr>
<tr>
<td>33</td>
<td>7</td>
<td>19</td>
<td>1</td>
<td>65</td>
<td>7</td>
</tr>
<tr>
<td>34</td>
<td>6</td>
<td>14</td>
<td>3</td>
<td>97</td>
<td>10</td>
</tr>
<tr>
<td>35</td>
<td>6</td>
<td>9</td>
<td>3</td>
<td>82</td>
<td>5</td>
</tr>
<tr>
<td>36</td>
<td>6</td>
<td>14</td>
<td>7</td>
<td>97</td>
<td>11</td>
</tr>
<tr>
<td>37</td>
<td>9</td>
<td>22</td>
<td>0</td>
<td>70</td>
<td>6</td>
</tr>
<tr>
<td>38</td>
<td>4</td>
<td>18</td>
<td>2</td>
<td>86</td>
<td>8</td>
</tr>
<tr>
<td>39</td>
<td>5</td>
<td>14</td>
<td>3</td>
<td>74</td>
<td>8</td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>20</td>
<td>3</td>
<td>41</td>
<td>8</td>
</tr>
<tr>
<td>41</td>
<td>6</td>
<td>17</td>
<td>7</td>
<td>92</td>
<td>12</td>
</tr>
<tr>
<td>42</td>
<td>5</td>
<td>15</td>
<td>6</td>
<td>86</td>
<td>8</td>
</tr>
<tr>
<td>43</td>
<td>1</td>
<td>15</td>
<td>8</td>
<td>58</td>
<td>13</td>
</tr>
<tr>
<td>44</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>82</td>
<td>8</td>
</tr>
<tr>
<td>45</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>74</td>
<td>8</td>
</tr>
<tr>
<td>46</td>
<td>4</td>
<td>14</td>
<td>2</td>
<td>80</td>
<td>8</td>
</tr>
<tr>
<td>47</td>
<td>5</td>
<td>9</td>
<td>3</td>
<td>90</td>
<td>9</td>
</tr>
<tr>
<td>48</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>84</td>
<td>8</td>
</tr>
<tr>
<td>49</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>82</td>
<td>11</td>
</tr>
<tr>
<td>50</td>
<td>3</td>
<td>11</td>
<td>6</td>
<td>43</td>
<td>3</td>
</tr>
<tr>
<td>51</td>
<td>2</td>
<td>12</td>
<td>6</td>
<td>60</td>
<td>8</td>
</tr>
<tr>
<td>52</td>
<td>6</td>
<td>18</td>
<td>0</td>
<td>74</td>
<td>7</td>
</tr>
<tr>
<td>53</td>
<td>2</td>
<td>22</td>
<td>3</td>
<td>72</td>
<td>7</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>75</td>
<td>9</td>
<td>3</td>
<td>.375</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>38</td>
<td>8</td>
<td>5</td>
<td>.625</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>143</td>
<td>17</td>
<td>6</td>
<td>.353</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>33</td>
<td>9</td>
<td>4</td>
<td>.444</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>59</td>
<td>8</td>
<td>3</td>
<td>.375</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>105</td>
<td>19</td>
<td>7</td>
<td>.368</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>90</td>
<td>14</td>
<td>4</td>
<td>.286</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>15</td>
<td>6</td>
<td>3</td>
<td>.500</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>56</td>
<td>6</td>
<td>4</td>
<td>.667</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>66</td>
<td>5</td>
<td>2</td>
<td>.400</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>57</td>
<td>4</td>
<td>1</td>
<td>.250</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>48</td>
<td>4</td>
<td>0</td>
<td>.000</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>66</td>
<td>4</td>
<td>2</td>
<td>.500</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>76</td>
<td>10</td>
<td>5</td>
<td>.500</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>69</td>
<td>7</td>
<td>5</td>
<td>.714</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>50</td>
<td>3</td>
<td>1</td>
<td>.333</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>85</td>
<td>7</td>
<td>1</td>
<td>.143</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>76</td>
<td>14</td>
<td>6</td>
<td>.429</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>62</td>
<td>8</td>
<td>3</td>
<td>.375</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>63</td>
<td>10</td>
<td>1</td>
<td>.100</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>56</td>
<td>8</td>
<td>4</td>
<td>.500</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>74</td>
<td>9</td>
<td>1</td>
<td>.111</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>49</td>
<td>5</td>
<td>2</td>
<td>.400</td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>157</td>
<td>16</td>
<td>7</td>
<td>.438</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>89</td>
<td>17</td>
<td>4</td>
<td>.211</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>97</td>
<td>6</td>
<td>3</td>
<td>.500</td>
</tr>
<tr>
<td>27</td>
<td>1</td>
<td>84</td>
<td>17</td>
<td>11</td>
<td>.647</td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>53</td>
<td>9</td>
<td>6</td>
<td>.667</td>
</tr>
<tr>
<td>29</td>
<td>2</td>
<td>34</td>
<td>4</td>
<td>1</td>
<td>.250</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>29</td>
<td>5</td>
<td>0</td>
<td>.000</td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>89</td>
<td>10</td>
<td>4</td>
<td>.400</td>
</tr>
<tr>
<td>32</td>
<td>2</td>
<td>111</td>
<td>16</td>
<td>8</td>
<td>.500</td>
</tr>
<tr>
<td>33</td>
<td>2</td>
<td>83</td>
<td>6</td>
<td>2</td>
<td>.333</td>
</tr>
<tr>
<td>34</td>
<td>2</td>
<td>150</td>
<td>20</td>
<td>6</td>
<td>.300</td>
</tr>
<tr>
<td>35</td>
<td>1</td>
<td>67</td>
<td>11</td>
<td>4</td>
<td>.364</td>
</tr>
<tr>
<td>36</td>
<td>2</td>
<td>81</td>
<td>12</td>
<td>6</td>
<td>.500</td>
</tr>
<tr>
<td>37</td>
<td>2</td>
<td>103</td>
<td>14</td>
<td>2</td>
<td>.143</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
<td>45</td>
<td>7</td>
<td>3</td>
<td>.429</td>
</tr>
<tr>
<td>39</td>
<td>2</td>
<td>39</td>
<td>6</td>
<td>4</td>
<td>.667</td>
</tr>
<tr>
<td>40</td>
<td>2</td>
<td>38</td>
<td>5</td>
<td>1</td>
<td>.200</td>
</tr>
<tr>
<td>41</td>
<td>2</td>
<td>121</td>
<td>24</td>
<td>12</td>
<td>.500</td>
</tr>
<tr>
<td>42</td>
<td>2</td>
<td>83</td>
<td>10</td>
<td>6</td>
<td>.600</td>
</tr>
<tr>
<td>43</td>
<td>1</td>
<td>67</td>
<td>11</td>
<td>6</td>
<td>.545</td>
</tr>
<tr>
<td>44</td>
<td>2</td>
<td>120</td>
<td>17</td>
<td>5</td>
<td>.294</td>
</tr>
<tr>
<td>45</td>
<td>1</td>
<td>62</td>
<td>9</td>
<td>2</td>
<td>.222</td>
</tr>
<tr>
<td>46</td>
<td>1</td>
<td>27</td>
<td>7</td>
<td>2</td>
<td>.286</td>
</tr>
<tr>
<td>47</td>
<td>2</td>
<td>79</td>
<td>12</td>
<td>3</td>
<td>.250</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>36</td>
<td>3</td>
<td>3</td>
<td>1.000</td>
</tr>
<tr>
<td>49</td>
<td>2</td>
<td>44</td>
<td>10</td>
<td>2</td>
<td>.200</td>
</tr>
<tr>
<td>50</td>
<td>2</td>
<td>43</td>
<td>3</td>
<td>1</td>
<td>.333</td>
</tr>
<tr>
<td>51</td>
<td>1</td>
<td>38</td>
<td>9</td>
<td>1</td>
<td>.111</td>
</tr>
<tr>
<td>52</td>
<td>2</td>
<td>64</td>
<td>8</td>
<td>2</td>
<td>.250</td>
</tr>
<tr>
<td>53</td>
<td>1</td>
<td>64</td>
<td>5</td>
<td>2</td>
<td>.400</td>
</tr>
<tr>
<td>REC</td>
<td>SUM/WD2</td>
<td>IU2</td>
<td>FOURS2</td>
<td>SUM4/PRO</td>
<td>EFF2</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>-----</td>
<td>--------</td>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>75</td>
<td>8</td>
<td>4</td>
<td>.500</td>
<td>.333</td>
</tr>
<tr>
<td>2</td>
<td>53</td>
<td>7</td>
<td>5</td>
<td>.714</td>
<td>.434</td>
</tr>
<tr>
<td>3</td>
<td>80</td>
<td>17</td>
<td>8</td>
<td>.471</td>
<td>.650</td>
</tr>
<tr>
<td>4</td>
<td>53</td>
<td>7</td>
<td>3</td>
<td>.429</td>
<td>.415</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
<td>6</td>
<td>2</td>
<td>.333</td>
<td>.531</td>
</tr>
<tr>
<td>6</td>
<td>92</td>
<td>13</td>
<td>5</td>
<td>.385</td>
<td>.391</td>
</tr>
<tr>
<td>7</td>
<td>65</td>
<td>8</td>
<td>3</td>
<td>.375</td>
<td>.338</td>
</tr>
<tr>
<td>8</td>
<td>109</td>
<td>20</td>
<td>10</td>
<td>.500</td>
<td>.605</td>
</tr>
<tr>
<td>9</td>
<td>132</td>
<td>12</td>
<td>4</td>
<td>.333</td>
<td>.280</td>
</tr>
<tr>
<td>10</td>
<td>24</td>
<td>4</td>
<td>1</td>
<td>.250</td>
<td>.500</td>
</tr>
<tr>
<td>11</td>
<td>57</td>
<td>6</td>
<td>3</td>
<td>.500</td>
<td>.351</td>
</tr>
<tr>
<td>12</td>
<td>42</td>
<td>4</td>
<td>1</td>
<td>.250</td>
<td>.286</td>
</tr>
<tr>
<td>13</td>
<td>149</td>
<td>10</td>
<td>4</td>
<td>.400</td>
<td>.188</td>
</tr>
<tr>
<td>14</td>
<td>83</td>
<td>11</td>
<td>4</td>
<td>.364</td>
<td>.386</td>
</tr>
<tr>
<td>15</td>
<td>39</td>
<td>7</td>
<td>3</td>
<td>.429</td>
<td>.538</td>
</tr>
<tr>
<td>16</td>
<td>38</td>
<td>5</td>
<td>2</td>
<td>.400</td>
<td>.395</td>
</tr>
<tr>
<td>17</td>
<td>71</td>
<td>9</td>
<td>2</td>
<td>.222</td>
<td>.310</td>
</tr>
<tr>
<td>18</td>
<td>109</td>
<td>13</td>
<td>7</td>
<td>.538</td>
<td>.389</td>
</tr>
<tr>
<td>19</td>
<td>32</td>
<td>4</td>
<td>3</td>
<td>.750</td>
<td>.469</td>
</tr>
<tr>
<td>20</td>
<td>58</td>
<td>6</td>
<td>1</td>
<td>.167</td>
<td>.224</td>
</tr>
<tr>
<td>21</td>
<td>57</td>
<td>5</td>
<td>2</td>
<td>.400</td>
<td>.263</td>
</tr>
<tr>
<td>22</td>
<td>55</td>
<td>7</td>
<td>1</td>
<td>.143</td>
<td>.255</td>
</tr>
<tr>
<td>23</td>
<td>34</td>
<td>4</td>
<td>2</td>
<td>.500</td>
<td>.412</td>
</tr>
<tr>
<td>24</td>
<td>63</td>
<td>6</td>
<td>4</td>
<td>.667</td>
<td>.317</td>
</tr>
<tr>
<td>25</td>
<td>126</td>
<td>15</td>
<td>8</td>
<td>.533</td>
<td>.389</td>
</tr>
<tr>
<td>26</td>
<td>29</td>
<td>3</td>
<td>3</td>
<td>1.000</td>
<td>.414</td>
</tr>
<tr>
<td>27</td>
<td>105</td>
<td>17</td>
<td>10</td>
<td>.588</td>
<td>.533</td>
</tr>
<tr>
<td>28</td>
<td>37</td>
<td>6</td>
<td>1</td>
<td>.167</td>
<td>.432</td>
</tr>
<tr>
<td>29</td>
<td>47</td>
<td>5</td>
<td>2</td>
<td>.400</td>
<td>.319</td>
</tr>
<tr>
<td>30</td>
<td>56</td>
<td>6</td>
<td>1</td>
<td>.167</td>
<td>.232</td>
</tr>
<tr>
<td>31</td>
<td>38</td>
<td>5</td>
<td>2</td>
<td>.400</td>
<td>.342</td>
</tr>
<tr>
<td>32</td>
<td>101</td>
<td>14</td>
<td>8</td>
<td>.571</td>
<td>.475</td>
</tr>
<tr>
<td>33</td>
<td>45</td>
<td>3</td>
<td>2</td>
<td>.667</td>
<td>.244</td>
</tr>
<tr>
<td>34</td>
<td>27</td>
<td>4</td>
<td>1</td>
<td>.250</td>
<td>.481</td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td>5</td>
<td>2</td>
<td>.400</td>
<td>.375</td>
</tr>
<tr>
<td>36</td>
<td>38</td>
<td>1</td>
<td>0</td>
<td>.000</td>
<td>.079</td>
</tr>
<tr>
<td>37</td>
<td>58</td>
<td>4</td>
<td>2</td>
<td>.500</td>
<td>.241</td>
</tr>
<tr>
<td>38</td>
<td>64</td>
<td>9</td>
<td>4</td>
<td>.444</td>
<td>.469</td>
</tr>
<tr>
<td>39</td>
<td>33</td>
<td>5</td>
<td>4</td>
<td>.800</td>
<td>.545</td>
</tr>
<tr>
<td>40</td>
<td>26</td>
<td>2</td>
<td>0</td>
<td>.000</td>
<td>.154</td>
</tr>
<tr>
<td>41</td>
<td>37</td>
<td>7</td>
<td>5</td>
<td>.714</td>
<td>.676</td>
</tr>
<tr>
<td>42</td>
<td>39</td>
<td>8</td>
<td>2</td>
<td>.250</td>
<td>.564</td>
</tr>
<tr>
<td>43</td>
<td>105</td>
<td>14</td>
<td>9</td>
<td>.643</td>
<td>.476</td>
</tr>
<tr>
<td>44</td>
<td>100</td>
<td>15</td>
<td>6</td>
<td>.400</td>
<td>.430</td>
</tr>
<tr>
<td>45</td>
<td>69</td>
<td>11</td>
<td>4</td>
<td>.333</td>
<td>.478</td>
</tr>
<tr>
<td>46</td>
<td>80</td>
<td>15</td>
<td>7</td>
<td>.467</td>
<td>.600</td>
</tr>
<tr>
<td>47</td>
<td>77</td>
<td>8</td>
<td>4</td>
<td>.500</td>
<td>.338</td>
</tr>
<tr>
<td>48</td>
<td>37</td>
<td>7</td>
<td>3</td>
<td>.429</td>
<td>.595</td>
</tr>
<tr>
<td>49</td>
<td>37</td>
<td>5</td>
<td>2</td>
<td>.400</td>
<td>.432</td>
</tr>
<tr>
<td>50</td>
<td>20</td>
<td>1</td>
<td>1</td>
<td>1.000</td>
<td>.200</td>
</tr>
<tr>
<td>51</td>
<td>73</td>
<td>11</td>
<td>5</td>
<td>.455</td>
<td>.438</td>
</tr>
<tr>
<td>52</td>
<td>25</td>
<td>2</td>
<td>2</td>
<td>1.000</td>
<td>.320</td>
</tr>
<tr>
<td>53</td>
<td>82</td>
<td>9</td>
<td>6</td>
<td>.667</td>
<td>.378</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
VITA

Martha H. Head received a Bachelor of Science degree in Mathematics and Science Education from Louisiana State University in May 1971. She subsequently taught science and mathematics for 12 years in the public schools of Louisiana and Kentucky. She received a Master of Arts degree from Louisiana State University in Reading Education in December 1982 and was awarded the Doctor of Philosophy degree in Reading Education by Louisiana State University in May 1986. She is married and the mother of a daughter and a son.
DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Martha H. Head

Major Field: Education

Title of Dissertation: Factors Affecting Summary Writing and Their Impact on Reading Comprehension Assessment

Approved:

[Signatures]

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination:

May 5, 1986