

To the Graduate Council:

I am submitting herewith a thesis written by Jo Ellen Backus entitled "An Evaluation of the Readability of Five Science Series." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Special Education.

W. J. Schindler

W. J. Schindler, Major Professor

We have read this thesis
and recommend its acceptance:

Phyllis Huff

Susan M. Benner

Accepted for the Council:

Carmichael

Vice Provost
and Dean of The Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in her absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature

Jo Ellen Backus

Date

July 20, 1988

AN EVALUATION OF THE READABILITY OF FIVE
SCIENCE SERIES

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Jo Ellen Backus

August 1988

ACKNOWLEDGMENTS

I wish to thank my family, friends and professors for all their help and support during the time I have been in graduate school. Without their love and support I would not have been able to complete my graduate program.

I want to thank my husband, William R. Backus, for being my biggest supporter and my friend and sounding board. My five children have had to make many adjustments in their schedules during my years in graduate school and I am very appreciative of their efforts on my behalf. I also have to say a big thank you to my grandchildren, Brad, Ben, Matthew and Anne Marie, for the love they have shown and the little things they have done to make me laugh at times when I thought there were only tears.

I also want to say how much I appreciate the time and effort given on my behalf by Dr. W. J. Schindler, Dr. Susan Benner, and Dr. Phyllis Huff. Without their professional guidance I would not have been able to realize the dream of attaining a master's degree. I have the utmost respect for these three women. I admire them for their dedication to education and to the students with whom they work.

Thank all of you for the role you are playing in my life.

ABSTRACT

This study examined the readability of fourth, fifth, and sixth grade science textbooks of five publishers of science textbooks. The Dale-Chall readability formula was used to determine the actual readability level of the textbooks and the Worksheet for Analysis of Instructional Materials Beyond Readability Scales (Schindler, 1980) was used in examining new words introduced in a lesson and the number of times each was reinforced. Three samples were taken from each textbook. Samples were randomly selected from the front, middle, and end sections of each text. The actual readability of the textbooks was found to be higher than the publishers' estimated readability. The fourth grade textbooks were the most difficult to read and the sixth grade textbooks were the least difficult. There did not appear to be a correlation between the number of times new words were reinforced and actual readability across the textbooks examined.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Background.	1
Statement of the Problem.	3
Definition of Terms	4
Assumptions	6
Summary	6
II. RELATED LITERATURE	7
III. DESIGN FOR EVALUATION OF SCIENCE SERIES.	14
Purpose	14
Data Collection Procedures.	14
Instrumentation	15
Data Analysis	16
IV. RESULTS.	17
Introduction.	17
Individual Analysis of Each Science Series. . .	17
Comparative Analysis of the Five Science Series.	25
Summary	29
V. CONCLUSIONS.	31
Discussion and Conclusions.	31
Limitations	33
Suggestions for Further Research.	34
BIBLIOGRAPHY	36
APPENDICES	39
APPENDIX A. DALE-CHALL READABILITY FORMULA- TABLE II	40
APPENDIX B. WORKSHEET FOR ANALYSIS OF INSTRUCTIONAL MATERIALS BEYOND READABILITY SCALES	41
VITA	43

LIST OF TABLES

TABLE	PAGE
1. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Holt Science Series.	18
2. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Silver Burdette Science Series	20
3. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Scott Foresman Science Series.	21
4. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Heath Science Series.	23
5. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Harcourt, Brace, Jovanovich Science Series.	24
6. Mean Index and Percentage of New Words Reinforced.	26
7. Actual Calculated Grade Levels.	28

CHAPTER I

INTRODUCTION

1. Background

The ability of a student to read and understand or comprehend a textbook in a specific content area in the grade for which the book was written is frequently overlooked as a problem in most school systems in the United States. This is often the result of having a textbook that is above the readability level of the average student in the grade for which the book is written. The introduction of new words and the number of times they are reinforced, by being repeated five or more times, can also be an area of difficulty.

Spache (1964) states the following:

The selection of books of difficulty appropriate for the reader is a recurring problem of parents, teachers, and librarians. The readability of the material, or in other words, the degree of comprehensibility, must constantly be matched to the reading abilities of the child or adult for whom a book is apparently intended (p. 21).

A major concern in the science field is whether or not the books of a series are in fact appropriate for the grade level a publisher indicated. Betts (1957) said "Readiness is a problem at all levels of instruction" (p. 5). This statement is probably still true today. Educators must look at student needs as well as curricula in selecting a series and make selections that will meet those needs. Betts also stated that a wide range of language abilities exists at any one

"grade level." This is a need that educators should consider in their selection of a series.

Publishers do not always conduct comprehensive evaluations of the text materials they publish. Therefore, it is up to the educators to comprehensively evaluate the readability of the material and then make decisions as to the suitability of the books.

The teaching of science involves being able to introduce new materials to students in a way that they can comprehend. Science is an area where the student is involved in new discovery as well as explanations of many of the things going on in the world around him. Teachers may use a variety of ways to present this material to the students. However, it remains that one of the primary tools in teaching science, as well as other subjects, is the textbook.

The purchasing of textbooks by a school system is usually an extensive financial commitment. Careful evaluation of these books needs to be made in order to make a selection that will be the most suitable for the students using them. Teachers and others involved in planning curriculum must be involved in the investigation of both cost and appropriateness of the science curricula, so that the highest quality of program is attained.

The validity of the publisher's estimated grade placement of the book should be investigated. One should not rely entirely on the word of the publisher. It is often assumed that the grade level designated by the publishers is correct and that the book will be appropriate. Spache (1978) says that publishers customarily designate

their textbooks as being suited for specific grade levels and these grade level designations are undoubtedly made in good faith, but are frequently erroneous. The student becomes frustrated when using textbooks that are too difficult or deviate from one to several readability levels from the beginning to the end of the book. The utilization of comprehensive evaluation techniques can eliminate the use of the inappropriate materials that cause some of these frustrations, because these techniques can determine the actual grade level of the book. This is especially important when a student's reading ability is lower than the grade placement.

Every day educators are confronted with the child who had trouble reading for one reason or another. Betts (1957) stated that children vary in their rates of learning. This is certainly one problem area. The child's inability to read might be attributed to several things, the materials, the instruction, or a reading difficulty. Certainly, being exposed to inappropriate material will decrease this child's chance of success. The educator should be willing to try all avenues open to him/her in order to give the student the best possible chance of success.

2. Statement of the Problem

This study is being done to comprehensively evaluate the following, as related to science textbooks in the curricula:

1. The relationship between publishers' stated grade level and actual readability levels of five science series.

2. Reinforcement of new words by their appearance at least five times in the lesson.

3. The consistency of grade level and readability within each text.

The comprehensive evaluation techniques include: (1) computing the actual readability of the textbooks; (2) measuring the number of reinforcements of new words introduced in a lesson; and (3) checking the consistency of the actual readability and grade level within each text. These have a direct effect on the actual grade level of the texts in a science series.

The hypotheses for this study were:

1. The grade level estimates of the publishers are accurate.
2. The new words introduced in a lesson are reinforced sufficiently in the lesson at a rate of five or more repetitions.

3. Definition of Terms

For the use of this study the following terms will be defined as:

1. Comprehensive evaluation is defined as an evaluation technique for science textbook series encompassing the use of the Dale-Chall readability formula and the Worksheet for Analysis of Instructional Materials Beyond Readability Scales (Schindler, 1980).

2. Comprehensive readability is defined as the grade level of textbooks as computed by the readability formula, new words introduced

and the number of times each is reinforced in a lesson. Other factors may also be a part of comprehensive readability.

A. Actual readability is the calculated grade level according to readability formulas.

B. Estimated readability is the grade level designation that publishers have placed on published textbooks.

3. Factors are the characteristics of published materials that have an effect on the actual readability of that material. Some readability factors would be vocabulary words, frequency of word reinforcement, interest-value, style of writing, typography, and sentence length.

4. New words are the new vocabulary words introduced in a lesson.

5. Reinforcement is the frequency of repetitions of new vocabulary words within a lesson.

6. Science series is defined as a published sequence of textbooks, teacher manuals, and student workbooks. Each textbook is assigned a grade level by the publisher.

7. Word lists accompany readability formulas to check percent words in the calculations. They were chosen by the authors of the readability formula to validate the calculated results.

8. Worksheet for Analysis of Instructional Materials Beyond Readability Scales--using this worksheet, the new words in a lesson can be listed and the frequency of reinforcement can be tabulated (Schindler, 1980).

4. Assumptions

There are some assumptions commonly made about science texts, as well as other programs in the curricula. Some of those assumptions are:

1. The program is written with the known oral language of the student taken into consideration.
2. It is not the responsibility of the teacher to evaluate all aspects of science, or other, materials before using them.
3. Reading is transferring one's oral language to the printed page and all children are able to do this and to do it at the same rate of learning.
4. Science vocabulary words do not need to be given special attention.

These assumptions will be addressed in the comprehensive evaluation of the five science series.

5. Summary

Reading and understanding or comprehending a textbook in a specific content area in the grade for which it was written is a problem that is seen in many school systems. Many factors need to be evaluated in determining the appropriateness of a textbook, such as readability, new word reinforcement, sentence structure, and sentence length. Educators need to be aware of these factors and know how to use them in selecting textbooks for the classroom.

CHAPTER II

RELATED LITERATURE

Looking at today's science curricula, the appropriateness of a chosen textbook would seem to be very important. Most science teachers today rely heavily on the textbook as the primary source of information. When a child first starts to read, he understands spoken materials at some level but does not understand these same materials in print at all (Gleitman and Rozin, 1977). Gleitman and Rozin state that reading is transferring one's oral language to the printed page, and that it is sometimes assumed that all children have been able to do this, and that they do it at the same rate of learning. The content of the material in any textbook must be written so it is easily understood by the student.

Reading and understanding or comprehending what is read is important in any curriculum. Readability can be described as an attempt to measure or predict a reader's ability to read textual materials. Watson (1981) states that the understanding of reading materials is related to the comprehensive readability of the text. Watson notes that this comprehensibility means not only the grade levels of each book in a series, but it takes into consideration the grade levels of the new words introduced and the number of times the new words are reinforced. Klare (1963) suggests that:

Readability is . . . predictive in the sense that it provides an estimate of difficulty of writing without requiring the reader to read it and undergo tests on it. In other words, it provides the kind of information about readability that a writer or teacher would have to judge through experience, or measure through a reading test (p. 34).

There are many readability formulas available for use in checking the readability of printed materials. The Dale-Chall formula was chosen for this study. This readability formula is easy to use to determine the grade level of material the teacher is using. Any educator can use a readability formula to check the reading level of printed materials.

The results of an extensive computer search of existing literature, indicates that very few studies have been made concerning readability of science textbooks. Not only have there not been many studies, but also most of those that were done were prior to 1970. The few conducted since 1970 are presented below, along with other older studies.

Herrington and Mallison (1958) found that there was a difference between the consistency with which readability formulas evaluated the grade level of reading difficulty of science textbooks and the consistency with which reading experts evaluated the grade level of the same textbooks. They found that the readability formulas were more consistent and suggest that a teacher should use readability formulas for checking grade levels of science textbooks. In addition to checking readability level, Herrington and Mallinson suggest that the teacher analyze the content of the textbook and evaluate it with

respect to its suitability. There should be as much concern about what the child reads as there is about the ability to read it.

Mallinson (1958) reports that extensive studies have resulted in these findings:

1. The reading levels of many textbooks in science are too advanced for the students for whom they are written.
2. The difference between the levels of reading difficulty of the easiest and the most difficult textbook in any area of science are significant.
3. In some textbooks of science whose average level of reading difficulty seems satisfactory, there are passages that would be difficult for some students in grades well above those for whom the book is intended.
4. Many textbooks of science contain non-technical words that could be replaced with easier synonyms (p. 475).

Mallinson, Sturm, and Mallinson (1954), in a study on the reading difficulty of general physical science and earth science textbooks, report that the levels of reading difficulty within the textbooks vary greatly. They say that in any textbook some passages are not likely to cause difficulty for the students for whom they are intended, while other passages will be difficult for even the most superior students. The average grade levels of difficulty of the textbooks also varied greatly.

Steinberg (1981) reports that when the Dale-Chall readability formula was used to measure the readability of intermediate grade science textbooks, it consistently gave grade equivalents that were higher than the publishers' designations. Gilbert (1973) compared readability levels and ranges of major elementary science texts with the approximate reading levels of each of the respective grades. He also wanted to determine to what extent a relationship exists among

the readability formulas. The data indicated that when various reading formulas were applied to the texts of selected publishers the averages and ranges tended to be slightly higher than should be expected of elementary students. Terminology and sentence structure used by the writers were thought to be factors that contributed to the scores. When different reading formulas were applied to the same passages the readability levels were compatible. The Spache formula and the Fry Graph were used for Grades 1, 2, and 3. The Dale-Chall formula, the Lorge formula, and the Fry Graph were used for Grades 4, 5, and 6.

Newport (1965) conducted a study designed to determine the readability level of nine continuous series of elementary school science textbooks. By using the Spache formula for Grades 1, 2, and 3, he found no considerable differences. In using the Yoakam formula for Grades 4, 5, and 6 the readability levels varied widely from series to series. Newport notes that there might be inadequacies in the Yoakam formula for measuring science textbooks.

Another concern regarding readability is vocabulary by grade level. Introduction of new words and how often they are reinforced are factors that need to be considered. New words need to be reinforced at least five times in the initial lesson context (Kingsley, 1965). The amount of reinforcement of new vocabulary words enhances one's actual retention of the vocabulary.

O'Toole and Bedford (1969) found the lack of science terms appearing on readability word lists to be a contributing factor in

the inadequacy of readability of science materials. They compiled a list of science words that commonly occurred in science textbooks and that could be correctly identified at least 60% of the time by the population of fourth graders they tested. These words were then used to develop a supplemental word list of 36 words. These 36 words when used with the Dale word list and the Dale-Chall reading formula affected reading difficulty levels of 4 of 5 fifth grade elementary science textbooks. Reading levels of three textbooks dropped one grade level and the other one dropped two grade levels.

Steinberg (1981), in summarizing readability research on science materials, reports that common vocabulary in science textbooks is likely as big an influence on readability as is scientific terminology. Science textbook publishers might need to look at the terminology used and strive to have some uniformity in the vocabulary used.

Gilbert (1973) determined that it is justifiable to say that use of readability formulas should be encouraged as they can serve as good indicators of the general ranges of reading difficulty. He also points out that publishers of elementary science textbooks should make every attempt to adjust reading levels of their books to the reading level of the children who would use them.

Doran and Sheard (1974), in discussing evaluation of science textbooks, remind us that even in classes where grouping occurs, all children within a group may not be at the same reading level. They say that considering that individuals have three reading levels: the independent level, the instructional level, and the frustration

level, it becomes crucial that an objective assessment be made of the materials that are to be used. The frustration level is when the reader is unable to recognize many of the words and thus unable to comprehend the material. The instructional level is when the reader recognizes 95% of the words and comprehends 75% of the material. At the independent level the reader knows almost all of the words and has a command of the material presented. Science texts designed to cover content need not be at the frustration level. If a child is able to read assignments at the independent, or even at the instructional reading level, both classroom performance and classroom behavior will benefit.

Jorgenson (1977) reports from his study that as materials become more difficult for students, the students comprehend less and become more dependent on others for assistance. He also found that the material difficulty-ability level difference scores are related to student behavior. Behavior improved as the materials became easier for the students.

Shymansky and Yore (1979) state the following:

Reading skills are no less critical in reading science topics than they are in reading any other materials. In fact, the vocabulary peculiar to science, content loading, sentence structure, the use of symbols, graphics, directions, and if-then statements, make science reading skills more difficult and more critical to overall student success. If a textbook is a student's only source of input and he is unable to read at an appropriate level, it is unlikely that he can even begin to understand the science concepts involved (p. 670).

Once a decision has been made to use textbook material in a science curricula, readability should be a major concern of the teacher.

The research reviewed indicates that the readability levels of science textbooks have varied from the publishers' estimated grade levels. Since there is little evidence of readability studies being conducted in relation to science texts in the past 10 years, it is important that this study be made. There is a need to have a current comprehensive evaluation of the readability of publishers' science series, including an analysis of the new words and the frequency of repetitions of them.

CHAPTER III

DESIGN FOR EVALUATION OF SCIENCE SERIES

1. Purpose

The purposes of this study were:

1. To determine if there is a difference between the publishers' estimated grade levels of five science series and the actual readability of the series, as determined by the Dale-Chall readability scale.
2. To determine if there is a difference between the number of times new words are reinforced and the recommended five times.
3. To determine if there is a consistency of grade level and readability within each text.

2. Data Collection Procedures

Five science series used in various school systems have been chosen for this study. The series chosen were as follows: (1) Holt, Rinehard and Winston (1984); (2) Silver Burdette (1985); (3) Scott, Foresman (1984); (4) D. C. Heath (1985); and (5) Harcourt, Brace, Jovanovich (1985). The grades to be evaluated are fourth, fifth and sixth grades. Having taught science at the sixth grade level, the researcher was interested in the readability of textbooks at this level and decided to evaluate the intermediate grades. Data were gathered by the researcher. The books were provided by the researcher.

3. Instrumentation

The following instruments were used to collect data for this study: (1) The Dale-Chall Readability Formula (see Appendix A) was chosen because of its validity, word list and structure. Klare (1963) reports that formulas for readability are compared with similar formulas to establish validity. The Dale-Chall proved valid when scores were compared in several samples. Harris and Spivey (1985) report that "Among most widely used formulas are those by Spache (1974) for the primary grades and by Dale-Chall (1948) for grades four through college" (p. 598). The Dale-Chall Readability Formula uses three 100 word samplings. Along with the word samplings there is a word list, the Dale List, which is used in finding the hard words. Samples were taken from three sections of each textbook, the beginning, the middle and the end of each text. (2) Worksheet for Analysis of Instructional Materials Beyond Readability Scales (see Appendix B) was developed by Schindler (1980) to analyze new words in a lesson or story. It analyzes attributes other than the readability. Sampling, once again, is taken from three sections of each textbook, beginning, middle, and end. The new words introduced in each sample are recorded in the far left column. The second and third columns are for recording how many times each new word was reinforced in that lesson and the next three appropriate lessons. This same form includes recording the skills to be taught as listed in the teacher's manual of each series and whether or not they are actually taught.

4. Data Analysis

Data were analyzed by charting each science series according to the readability scores and frequency of new word reinforcements.

To compare the data a ratio was computed by placing publishers' estimated grade level over the actual readability level and then subtracting this from one. The number derived from this calculation will be an index number for each book. A 20% confidence level was used as the adequate tolerance for index numbers close to 0, +.1 or -.1. A positive index value indicated that the actual readability, according to the readability formula, is higher than the estimated readability. A negative index indicated that the actual readability is lower than the estimated readability. This method is consistent with other readability studies done by Watson (1981) and O'Connor (1988).

The frequency of new words reinforced were charted in three columns; less than five, on five, and more than five. Percentages were calculated by number of reinforcements over total number of new words introduced.

CHAPTER IV

RESULTS

1. Introduction

The purpose of this study was to evaluate the readability of five different science series adapted for use in Tennessee. The questions to be explored were the relationship between publisher's designated readability and the readability as determined using the Dale-Chall Readability Formula (see Appendix A), and the relationship between new words introduced and the reinforcement of new words according to the recommended five times or more.

The five science series used in this study were: (1) Harcourt, Brace, Jovanovich (1985); (2) Holt, Rinehart and Winston (1984); (3) D. C. Heath (1985); (4) Silver Burdett (1985); and (5) Scott, Foresman (1984). The grades that were evaluated were the fourth, fifth, and the sixth grades.

2. Individual Analysis of Each Science Series

The observations of each science series according to the two questions are shown below.

HOLT, RINEHART AND WINSTON

1. Does the publisher estimate of readability match the actual computed readability score?

The actual computed readability of the three Holt texts used in the study was higher than the publisher's estimated grade level,

according to the Dale-Chall formula. The fourth grade was the furthestmost from the publisher's estimated readability according to the calculated index numbers (see Table 1). The fourth grade text showed a .26 index, meaning that the actual readability was higher than the publisher's estimated grade level. The fifth grade text with an index of .22 was also higher than the publisher's estimated grade level. The sixth grade text was closest to the publisher's estimated readability. The calculated index for the sixth grade was .06, which means that the actual readability was within the 20% confidence level.

The hypothesis that the grade level estimates of the publisher are accurate is not supported in the fourth and fifth grade text in the Holt series but is supported in the sixth grade text.

Table 1. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Holt Science Series

Estimated Readability	Actual Readability	Index	New Words Introduced Percent Reinforced		
			Less	Five	More
4	5.4	.26	36	11	53
5	6.4	.22	60	7	33
6	6.4	.06	0	0	100

2. What percentage of new words are reinforced at the recommended five times or more than five times in the text material?

In the Holt series 64% overall of the new words were reinforced at five times or more than five times. When each grade level text is examined, the fourth grade had 64% of the new words reinforced five times or more than five times in the lesson. The fifth grade had 40% and the sixth grade showed 100% of the new words reinforced five times or more than five times in the lesson (see Table 1).

The hypothesis that the new words introduced are reinforced sufficiently in the lesson at a rate of five or more repetitions is supported in the Holt science series.

SILVER BURDETTE

1. Does the publisher estimate of readability match the actual computed readability score?

According to the Dale-Chall readability calculations the Silver Burdette series has an overall higher readability than the publisher's estimated readability. The text labeled grade four had the highest calculated readability with an index of .35. The fifth grade text had a calculated index of .28 and the index of the sixth grade text was .17. These index computations indicated that while all of the texts have higher calculated readability scores, the fourth grade text seems to be the more difficult to read and the sixth grade text is the least difficult of the three books.

The hypothesis that the grade level estimates of the publisher are accurate is not supported in the Silver Burdette series.

2. What percentage of new words are reinforced at the recommended five times or more than five times in the text material?

Less than one half of the new words in the Silver Burdette science series are reinforced at the recommended five times or more than five times. The text labeled grade five had 52% of the new words reinforced at the recommended five times or more, while the text labeled grade four had 38% and grade six had 40% of the new words reinforced five or more times (see Table 2).

The hypothesis that the new words introduced are reinforced sufficiently in the lesson at a rate of five or more repetitions is not supported in the Silver Burdette series.

Table 2. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Silver Burdette Science Series

Estimated Readability	Actual Readability	Index	New Words Introduced Percent Reinforced		
			Less	Five	More
4	6.2	.35	62	0	38
5	6.9	.28	48	4	48
6	7.2	.17	60	30	10

SCOTT FORESMAN

1. Does the publisher estimate of readability match the actual computed readability score?

When the readability of the Scott Foresman series was calculated, it was found that the readability, using the Dale-Chall formula, for grades four, five and six was higher than the publisher's estimated grade level. The calculated index for the fourth grade text was higher than the calculated index for the texts labeled grades five and six (see Table 3). The index score for the sixth grade text was .02 which indicates that it was very close to the publisher's estimated grade, while the index for the fifth grade text was .15, which is higher than would be acceptable.

The hypothesis that the grade level estimates of the publisher are accurate is supported in the grade six text of the Scott Foresman series, but is not supported in the fourth and fifth grade texts.

Table 3. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Scott Foresman Science Series

Estimated Readability	Actual Readability	Index	New Words Introduced Percent Reinforced		
			Less	Five	More
4	5.2	.23	25	25	50
5	5.9	.15	50	12	38
6	6.1	.02	33	0	67

2. What percentage of new words are reinforced at the recommended five times or more than five times in the text material?

All of the science texts evaluated in the Scott Foresman series reinforce more than 50% of the new words the recommended five times or more in the lesson. The text labeled for grade four had 75% of the new words reinforced the recommended five times or more than five times. The text labeled grade five had 50% reinforced at the recommended five times or more. Sixty-seven percent of the new words in the text labeled grade six were reinforced five or more times (see Table 3).

The hypothesis that the new words introduced are reinforced sufficiently in the lesson at a rate of five or more repetitions is supported by this data.

HEATH

1. Does the publisher estimate of readability match the actual computed readability score?

The text labeled grade six in the Heath science series has a lower actual calculated readability than the estimated readability. The calculated index for this text was a negative .05 which indicates the readability was lower than the publisher's estimated readability. The actual readability of the text labeled grade five was in the acceptable range with an index of .04. The fourth grade text was quite a bit higher than the estimated readability. The calculated index of the actual readability of the text labeled grade four was .31 (see Table 4).

Table 4. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Heath Science Series

Estimated Readability	Actual Readability	Index	New Words Introduced Percent Reinforced		
			Less	Five	More
4	5.8	.31	67	8	25
5	5.2	.04	64	7	29
6	5.7	-.05	70	6	24

The hypothesis that the grade level estimates of the publisher are accurate is supported for grades five and six in the Heath series. The hypothesis is not supported for grade four.

2. What percentage of new words are reinforced at the recommended five times or more than five times in the text material?

The new words in a lesson fall substantially below the recommended five or more times in reinforcements. Those new words reinforced five or more times were 36% for the fourth grade, 30% for the fifth grade and 33% for the sixth grade (see Table 4).

The hypothesis that the new words introduced are reinforced sufficiently in the lesson at a rate of five or more repetitions is not supported in the evaluation of the Heath series.

HARCOURT, BRACE, JOVANOVIĆH

1. Does the publisher estimate of readability match the actual computed readability?

Two of the texts in the Harcourt, Brace, Jovanovich series were higher in readability than the publisher's designated grade level. The text labeled grade four had a calculated index of .33 which is much higher than the estimated readability. The text labeled for grade five also had a readability, according to the Dale-Chall formula, that was higher than the publisher's estimated readability. The index for grade five was .19. The text for grade six actually had a readability that was lower than the publisher's designated grade. The calculated index was a negative .03, which indicates that the readability of the sixth grade text was slightly lower than the designated grade (see Table 5).

The hypothesis that the grade level estimates of the publisher are accurate is supported for grade six text in the Harcourt, Brace, Jovanovich series. It is not supported for grades four and five.

Table 5. Estimated and Actual Readability, Index and Percentage New Words Reinforced for Harcourt, Brace, Jovanovich Science Series

Estimated Readability	Actual Readability	Index	New Words Introduced Percent Reinforced		
			Less	Five	More
4	6.0	.33	50	17	33
5	6.2	.19	20	20	60
6	5.8	-.03	25	12	63

3. Comparative Analysis of the Five Science Series

The five science series were analyzed across texts in order to answer the question:

1. Does the use of comprehensive evaluation process suggest reasons for discrepancies of the publishers' estimated readability and the actual readability scores based on the Dale-Chall Readability Formula?

Table 6 shows mean values for mean index numbers and mean percentages of reinforcement of new words for all five science series. These means were used to compare the five series to one another.

The Holt, Rinehart and Winston, the Harcourt, Brace, Jovanovich, and the Scott Foresman series all show a higher actual readability than the publisher's estimated readability. With mean indexes of .18, .16, and .13 respectively, it appears that these three series have approximately the same readability. The reinforcement of new words for these three series are comparable also, ranging from 60% to 72% of new words reinforced five or more times (see Table 6).

The Silver Burdette science series had the highest readability of the five science series evaluated. The mean index numbers for actual computed readability indicate that this series is quite a bit higher than the publisher's estimated readability. The number of new words reinforced five or more times was 45%, with 36% being reinforced more than five times. These percentages could be a factor in the readability being so high in the Silver Burdette series. There are other factors that affect readability also.

Table 6. Mean Index and Percentage of New Words Reinforced

Series	Mean Index	Mean Percent Reinforced		
		Less	Five	More
Holt	.18	36	11	53
Silver Burdette	.26	55	9	36
Scott Foresman	.13	40	13	47
Heath	.10	67	7	26
Harcourt, Brace, Jovanovich	.16	28	16	56

The mean index for the Heath Series indicates the actual readability is only slightly higher than the publisher's estimate (see Table 6). The readability is acceptable at .1 index score. The Heath series was the only one acceptable across all of the texts revised, although the text for grade four was higher with an index of .31 readability than were the texts for grades five and six. Sixty-seven percent of the new words in the Heath series were reinforced less than five times. This would not support the theory that new words need to be reinforced at five times or more per lesson. However, there are other factors that influence readability and these would have to be considered also, such as sentence structure, style of writing, and length of sentences. These could be relevant in determining a reason for readability differences.

2. Is there consistency of grade level and readability within each text?

There was no consistency of grade level and readability within each text. The fourth grade text of Scott Foresman was easiest in the beginning and hardest at the end of the text, as one would expect a text to be. The fifth grade text was close to the same readability and grade level with a grade of 5.1 at the beginning and 5.9 at the end. The middle of the text was high at 6.8 grade level. The sixth grade text was high at the beginning with 7.3 and the middle and end were low with grade levels of 5.3 and 5.7 respectively (see Table 7).

In the Heath texts the middle of the text was higher than the beginning and the end of the fifth and sixth grade texts. The actual calculated readability levels for the beginning and the end were lower than the designated grade levels (see Table 7).

In the Silver Burdette series the readability was lowest at the end of the fourth grade text with a 4.9 actual grade level and the highest in the middle with 7.8. The fifth grade text was higher at the beginning with a 7.7 actual grade level and the end, 7.3, than it was in the middle, 5.6. The sixth grade text was the same at the beginning and the end with a 6.7 actual grade level and very high in the middle with 8.1 (see Table 7).

The sixth grade text in the Holt series was only slightly higher at the beginning with a 6.8 grade level. The middle and the end had the same calculated grade level of 6.2. The sixth grade Holt text was the only one analyzed that had an actual readability grade

Table 7. Actual Calculated Grade Levels

Series	Fourth Grade			Fifth Grade			Sixth Grade		
	Beg.	Mid.	End	Beg.	Mid.	End	Beg.	Mid.	End
Silver Burdette	6.0	7.8	4.9	7.7	5.6	7.3	6.8	8.1	6.7
Holt	5.6	4.7	6.0	5.7	7.4	5.9	6.8	6.2	6.2
Scott Foresman	4.3	5.0	6.2	5.1	6.8	5.9	7.3	5.3	5.7
Heath	6.1	5.3	6.0	4.7	5.9	4.9	5.3	6.5	5.1
Harcourt, Brace, Jovanovich	7.1	6.0	4.8	5.0	7.5	6.0	5.2	5.3	6.9

that was the same as the designated grade at the beginning, middle and end. The fourth grade text was high at the beginning and end and lower in the middle. The fifth grade was higher in the middle than it was at the beginning and end (see Table 7).

The fourth grade Harcourt, Brace, Jovanovich text was harder at the beginning and easiest at the end. The fifth grade was easiest at the beginning with a readability grade of 5.0 and hardest in the middle with a 7.5 actual grade level. The sixth grade text was easiest in the beginning with a 5.2 actual grade level, with the middle having a 5.3 grade level. The end had the hardest readability at a grade level of 6.9 (see Table 7).

There was no real consistency of grade level and readability within each text. Overall there was a great deal of variation in actual readability from the beginning to the end within each text (see Table 7).

4. Summary

The hypothesis that the grade level estimates of the publisher are accurate is supported in the sixth grade text of the Holt series, but not for the fourth and fifth grades. It is not supported at all in the Silver Burdette series and is only supported in the sixth grade text of the Scott, Foresman series. In the Heath series the hypothesis is supported for the fifth and sixth grades, but not for the fourth grade. The hypothesis that the grade level estimates of the publishers are accurate is supported only by the sixth grade text in the Harcourt, Brace, Jovanovich series.

The hypothesis that the new words introduced are reinforced sufficiently in the lesson at a rate of five or more repetitions is supported in the Holt, Scott Foresman and Harcourt, Brace, Jovanovich series, but not in the Silver Burdette and Heath series.

In analyzing the five science series across the text it was found that the fourth grade texts were the most difficult to read and the sixth grade texts were the least difficult.

CHAPTER V

CONCLUSIONS

1. Discussion and Conclusions

This chapter is to discuss the results of the comprehensive evaluation process in relation to the five science series. The information given in Chapter IV indicates inaccurate readability of all five of the science series.

The methods used to evaluate comprehensive readability of these science series enabled the researcher to explore the texts and to draw conclusions about them. The across-text analysis showed Silver Burdette to be much higher in actual readability than the publisher's estimated readability. The Holt, Harcourt, Brace, Jovanovich, and Scott Foresman series were all higher in actual readability than the publisher's estimated readability. However, these three series were within the same range, or all had about the same actual readability scores. The Heath series, while still high in actual readability, had the lowest actual computed readability of the science series evaluated. This series was within the acceptable range of readability.

The texts labeled for grade four in all five series had the greatest discrepancies between actual grade level and the publishers' estimated grade level. All of the fourth grade texts were substantially higher in actual computed readability than the publishers' designated

grade. This researcher believes that other factors have influenced the difficulty of these texts. One factor that could have influenced the actual readability might be that it is at the fourth grade level that content begins to appear in textbooks. There is also a science vocabulary list of words available and these could have an influence on the actual readability. If these science words were removed when computing readability, it might lower the actual readability significantly. Other factors that might influence the actual readability would be sentence structure, style of writing or the length of the sentences.

The sixth grade textbooks were the lowest in actual readability, overall. All of the sixth grade texts evaluated were within an acceptable range of actual readability, with the exception of the Silver Burdette series, which was much higher in actual readability than the publisher's estimated readability. Two of the series, Harcourt, Brace, Jovanovich and Heath, had actual readabilities that were slightly lower than the publisher's estimated readability. The Holt and Scott Foresman series were well within the acceptable range in actual readability. One observation made by the researcher was that in all five series the fourth grade texts were the highest in actual readability, the fifth grade texts, while high in actual readability, were slightly lower than the fourth grade texts, and the sixth grade texts were lower than the fifth grade texts.

There did not appear to be a correlation between number of times new words were reinforced and actual readability across the texts.

The Heath series had the lowest actual readability and the lowest percentage, 33%, of words reinforced five or more times in a lesson. Silver Burdette had the highest actual readability and the new words were reinforced five or more times, 45% of the time. The other three series, which had about the same actual readability scores, had the highest percentage of new words reinforced five or more times, ranging from 50 to 72%. One factor that might cause this type of observation would be the new words that are introduced and their importance in the lesson. In science materials there are many vocabulary words that are unique to the science content and might only need to be used once in a given lesson.

2. Limitations

There were some limitations in this study. The number or samples used to compute actual readability was an area of limitation. Newer readability scales utilize five samples. Second, other factors that influence actual readability, such as sentence structure, style of writing and sentence length, could not be evaluated by the comprehensive evaluation.

Only taking three samples from each text in a series might have an influence on the actual readability scores. The samples were taken from front, middle, and back of each text and probably give a good overall actual readability score. However, the overall actual readability score might change somewhat if more samples were used.

This would not, however, affect the fact that there could be a lot of variation within a text. Length of text might determine the number of samples needed.

There was a variation in the relationship between the actual readability and the percentage of new words reinforced five times or more than five times. Not all of the series showed the kind of relationship where readability is lower when new words are reinforced more than 50% of the time. From this, the researcher concludes that there are other factors that affect the differences in actual readability and the publishers' estimated readability. These factors could be sentence length, content, style of writing, or sentence structure. The factors that were evaluated in this study and these other factors are all related to the actual readability of the text materials.

3. Suggestions for Further Research

Observations made in this study reveal that there are additional possibilities for research. The first observation would be to evaluate the five science series evaluated here and other science series using the science vocabulary lists that are available. This could have a marked effect on actual readability. Another option would be to evaluate the science series, as well as other content texts, such as social studies, reading, and math, according to the other factors thought to influence readability, for example, sentence length or sentence structure.

A third option might be to evaluate the advantages of using one particular series for all grades in a school system. This might give more continuity in content and give the student a more consistent base to build on each year. Each text should build on the previous year's text in a series published by one publisher.

Another option for further research would be to evaluate the science series by combining the factors used in this study with other factors that are thought to influence comprehensive readability. This type of study would give a more complete analysis of actual readability. Before this could be done, however, techniques would have to be developed for evaluating such things as sentence structure and style of writing.

There are probably other options for additional studies and further research into the factors that affect readability. Educators need to be aware of the research being done in regard to the textbooks they might choose to use. This is especially true when choosing texts that they will be using for a period of time. Not only do they need to be aware of research that has been done, they need to conduct their own readability tests, check content, and new word reinforcement in evaluating textbooks for use in their school system. Our children's education is a valuable investment and they deserve the best we have to offer.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Betts, E. A. Foundation of Reading Instruction. New York: American Book Co., 1957.
- Dale, E., and J. S. Chall. A Formula for Predicting Readability. Bureau of Educational Research. Columbus: Ohio State University. Reprinted from the Educational Research Bulletin, 1958, 27, 5-6.
- Dale List of 3000 Familiar Words. Dale, E., and J. S. Chall. A Formula for Predicting Readability. Bureau of Educational Research. Columbus: Ohio State University. Reprinted from the Educational Research Bulletin, 1948, 27.
- Doran, Rodney L., and Doris M. Sheard. Analyzing Science Textbooks. School Science and Mathematics, 1974.
- Gilbert, C. D. An Examination of Readability Levels for Selected Basic Science Texts. School Science and Mathematics, 1973.
- Gleitman, L. R. and P. Rozin. Structure and Acquisition of Reading: Relations Between Orthographies and the Structure of Language. In: A. S. Reber and D. L. Scarborough (eds.), Toward a Psychology of Reading. New Jersey: Lawrence Erlbaum Associates, 1977.
- HBJ Science Series. Orlando Fla.: Harcourt Brace Jovanovich, 1985.
- Harris, Albert J. and Edward R. Sipay. How to Increase Reading Ability. Longman, Inc., 1985.
- Heath Science Series. Lexington, Mass.: D. C. Heath and Company, 1985.
- Herrington, R. L. and G. G. Mallinson. An Investigation of Two Methods of Measuring the Reading Difficulty of Materials for Elementary Science. Science Education. December 1958, 42 385-90.
- Holt Science Series. New York, N.Y.: Holt, Rinehart and Winston, 1984.
- Jorgenson, Gerald W. Relationship of Classroom Behavior to the Accuracy of the Match Between Material Difficulty and Student Ability. Journal of Educational Psychology. 1977. 69 1.

- Kingsley, R. F. Adjustment of Curriculum for Handicapped Students. Presentation at Kent State University, Kent, Ohio. October, 1965.
- Klare, G. P. The Measurement of Readability. Ames: Iowa State University Press, 1963.
- Mallinson, G. G. Textbook and Reading Difficulty in Science Teaching. The Science Teacher. December, 1958, 25, 474-475.
- Mallinson, G. G., H. E. Sturm, and L. M. Mallinson. The Reading Difficulty of Textbooks in Junior High School Science. School Review, December, 1950, 58, 536-40.
- Newport, J. F. The Readability of Science Textbooks for Elementary School. The Elementary School Journal, 1965, 66, 40-43.
- O'Toole, R. J. and J. P. Bedford. Science Vocabulary and Readability Level. Journal of Research in Science Teaching, 1969, 6, 161-62.
- Schindler, W. J. Worksheet for Analysis of Instructional Materials Beyond Readability Scales. Knoxville, Tenn.: The University of Tennessee, Knoxville, 1980.
- Scott, Foresman Science Series. Glenview, Ill.: Scott Foresman and Co., 1984.
- Shymansky, J. A. and L. D. Yore. Assessing and Using Readability of Elementary Science Textbooks. School Science and Mathematics, 1979, 79, 670-676.
- Silver Burdette Science Program. Morristown, N.J.: Silver Burdette Co., 1985.
- Spache, G. D. Good Reading for Poor Readers. Champaign, Ill.: Garrard Publishing Co., 1964.
- Spache, G. D. Good Reading for Poor Readers. Champaign, Ill.: Garrard Publishing Co., 1978.
- Steinberg, Joan Emily, Ed.D. The Relationship Between Vocabulary and Readability of Science Textbooks for Grades Four, Five and Six. University of San Francisco, 1981.
- Watson, Melissa E. An Evaluation of the Readability of Five Developmental Reading Series. Masters of Science Thesis. The University of Tennessee, Knoxville, 1981.

APPENDICES

APPENDIX A

DALE-CHALL READABILITY FORMULA-TABLE II

Article: _____	No. _____	Page No. _____	Page No. _____
Author: _____	From _____	From _____	From _____
Publisher: _____	To _____	To _____	To _____
1. Number of words in the sample	_____	_____	_____
2. Number of sentences in the sample	_____	_____	_____
3. Number of words not on Dale List	_____	_____	_____
4. Average sentence length (divide 1 by 2)	_____	_____	_____
5. Dale score (divide 3 by 1, multiply by 100)	_____	_____	_____
6. Multiply average sentence length (4) by .0496	_____	_____	_____
7. Multiply Dale score (5) by .1579	_____	_____	_____
8. Constant	3.6365	3.6365	3.6365
9. Formula Raw score (add 6, 7, and 8)	_____	_____	_____
Average raw score of _____ samples	Analyzed by _____	Date _____	Date _____
Average corrected grade-level	Checked by _____	Date _____	Date _____

APPENDIX B

WORKSHEET FOR ANALYSIS OF INSTRUCTIONAL MATERIALS

BEYOND READABILITY SCALES

Name of Book _____

Workbook _____

Publisher _____

Proposed Grade Level _____

Other Information _____

Pick a story at the beginning, middle and end of each level and analyze according to information requested below:

[illegible]

<u>List New Words</u>	<u>Number of times reinforced</u>	<u>Reinforcements in next 3 lessons</u>

Looking at teachers manual, list skills to be taught in this lesson:

<u>Skills</u>	<u>Are they taught</u>

Skills to be taught or reinforced in workbook

Check appropriate observation
taught reinforced neither

VITA

Jo Ellen Backus was born in Falls Mills, Virginia on April 13, 1935, the second child of Harry and Ethel Burton Harris. She is one of five children, having two brothers and two sisters. She attended elementary school in Bluefield, Virginia. She attended high school in Branwell, West Virginia and Bluefield, Virginia, graduating from Graham High School, Bluefield, Virginia in 1952.

In 1954 Jo Ellen married her high school sweetheart, William R. Backus, presently Associate Professor of Animal Science at The University of Tennessee, Knoxville. She worked as a bookkeeper and office manager while her husband earned his bachelor's, master's and PhD degrees. She and her husband also reared five children, all of whom attended college.

In 1979 Jo Ellen decided to fulfill her dream of getting a college degree. She graduated from Roane State Community College in 1981 and received her bachelor's degree in elementary education from The University of Tennessee, Knoxville in 1983. She began teaching in the Alcoa City School System and started working toward a Master's degree in Special Education at The University of Tennessee, Knoxville in the fall of 1983. She will receive a Master of Science degree in August 1988.