

To the Graduate Council:

I am submitting herewith a thesis written by Melissa E. Watson entitled "An Evaluation of the Readability of Five Developmental Reading 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:

Carolynn A. Washburn
Susan M. Benn

Accepted for the Council:

L. Evans Cobb
Vice Chancellor
Graduate Studies and Research

AN EVALUATION OF THE READABILITY OF
FIVE DEVELOPMENTAL READING SERIES

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

Melissa E. Watson

August 1981

3062951

ACKNOWLEDGMENTS

I wish to express my gratitude to my family, friends, and professors for all the love and support they gave me throughout my four years of graduate school. Without their commitments, I could not have successfully completed my graduate program.

I would like to thank my parents, Clyde and Elaine Watson, for giving me life 26 years ago and for raising me not to settle for second best. They also presented me with the tools to make it in life, a flashlight and a knife. Cheryl, my sister, I especially want to thank for sharing her unique sense of humor and for her tremendous influence on my outlook on life.

Secondly, I want to thank all my friends who lifted me from the depths of this thesis with their love, humor, and encouragement. I especially want to thank Teresa, Kara, Jim, Gordy, Greg, and Phil.

Thirdly, for their professional and personal guidance, I wish to express my appreciation to Dr. W. J. Schindler, Dr. Carolyn Washbon, and Dr. Susan Benner. These professionals have my utmost respect for their dedication to education and for their gift of love to children. Our educational system has been enriched by these ladies through their devotion to children and to their students.

Thank you all for giving a part of you to me.

Melissa E. Watson

ABSTRACT

The purpose of this study was to evaluate five developmental reading series as to the readability scores, grade levels of new words introduced, and the reinforcement of the new words. The readability scores were calculated by the Spache Readability Formula and the Dale-Chall Formula. The new words introduced in the series were listed on the Worksheet for Analysis of Instructional Materials Beyond Readability Scales (Schindler, 1980) and then matched with A Functional Basic Word List for Special Pupils (Tudyman and Groelle, 1963) to find the actual grade levels. Reinforcements of new words were tabulated also on the worksheet by Schindler.

From these instruments of comprehensive evaluation, one developmental reading series showed a direct relationship between graded new words and the frequency of reinforcements with the actual readability versus the publisher's estimated readability. The other four developmental series showed a relationship influenced by some other factor of readability. From these findings, the other factors that are thought to influence readability have limited the relationship between readability scores and the new words introduced with frequency of reinforcement.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Background	2
Statement of the Problem	4
Definitions of Terms	5
Assumptions	7
Questions to be Explored	7
II. RELATED LITERATURE	9
Readability	9
The aspects of readability	9
Readability in relation to teaching reading	11
Comprehensive Evaluation of Readability	15
Readability scales	16
New words	18
A worksheet for further analysis	19
Summary	22
III. DESIGN FOR EVALUATION OF READING SERIES	23
Purpose	23
Data Collection Procedures	23
Instrumentation	24
Data Analysis	25
IV. RESULTS	27
Introduction	27

CHAPTER	PAGE
III. (Continued)	
Individual Analysis of Each Developmental Reading	
Program	28
Comparative Analysis of the Five Developmental Reading	
Programs	45
IV. CONCLUSIONS	48
Discussion and Conclusions	48
Limitations	51
Suggestions for Further Research	52
BIBLIOGRAPHY	54
APPENDICES	57
APPENDIX A. SPACHE READING FORMULA	58
APPENDIX B. DALE-CHALL READABILITY FORMULA-TABLE II	60
APPENDIX C. WORKSHEET FOR ANALYSIS OF INSTRUCTIONAL MATERIALS	
BEYOND READABILITY SCALES	61
VITA	63

LIST OF TABLES

TABLE		PAGE
1.	Estimated and Actual Readability, Index, Frequency and Percentage New Words, Percentage New Words Reinforced for SRA Corrective Reading Text (1974)	29
2.	Estimated and Actual Readability, Index, Frequency and Percentage New Words, Percentage New Words Reinforced for Ginn Texts (1961 and 1976)	30
3.	Estimated and Actual Readability, Index, Frequency and Percentage New Words, Percentage New Words Reinforced for Macmillan Text (1970)	34
4.	Estimated and Actual Readability, Index, Frequency and Percentage New Words, Percentage New Words Reinforced for Distar Text (SRA 1974)	37
5.	Estimated and Actual Readability, Index, Frequency and Percentage New Words, Percentage New Words Reinforced for Houghton-Mifflin Text (1978)	41
6.	Mean Index, Frequency, Percentage New Words, Percentage of New Words Reinforced	47

CHAPTER I

INTRODUCTION

In virtually every school system in America today, curricula are being influenced by various developmental reading programs. These sequenced programs of reading materials include teachers' manuals, readers for children, and workbooks to reinforce skills. In most systems, the developmental reader has become the dominating factor in the reading curricula. Because of the prominence of developmental reading programs, curriculum planners should be selective in adopting the appropriate text materials for reading.

A major concern in choosing a series is whether or not the individual books within the series are in fact appropriate for the grade level the publisher indicates. Educators must then make selections based on student needs as well as curricular desires.

Research shows that little effort is made to conduct a comprehensive evaluation of the text material, so the responsibility falls on the teacher selecting the series. Spache (1978) states the following:

Publishers customarily designate their readers and other textbooks as suited for specific grade levels. These grade levels designations are undoubtedly made in good faith according to the opinions of the author and the publishing staff. But like other pooled judgements they are frequently erroneous (p. 31).

This is why teachers must use the instruments available to

comprehensively evaluate a developmental reading series and then make decisions as to its usefulness.

The comprehensive evaluation processes are relatively simple and may prove to be a valuable use of the educator's time. "The use of procedures such as, readability scales, graded word lists, and worksheets for further analysis can save time and eliminate frustrations for both the teacher and student" (Schindler, 1981).

Background

The process of teaching reading involves teaching a sequence of many skills including comprehension and vocabulary. Teachers may use a variety of materials to teach such skills yet the primary materials used to teach reading are the developmental reading series. The need for sequence in teaching reading creates the dependency on a guide for skills and vocabulary development. Teachers have little time to write stories, therefore the reliance on these series becomes inevitable.

In addition, the cost of these series is so great that careful evaluation needs to be made in order to select the most suitable developmental reading program. This investment, by a school system, becomes an extensive financial commitment. Curriculum planners must investigate the cost and appropriateness of developmental reading programs so that the highest quality of reading is attainable.

The reliance on developmental reading programs encourages curiosity about the validity of the publisher's estimated grade placement of each book in the series. When a publisher places a grade

level designation on any book, it is often assumed that the designation is correct and the instruction will be appropriate. Spache (1978) suggests that, "The question of accuracy of publishers' level designations is perhaps most acute in the area of basal readers" (p. 31). He found that many basal texts are far too difficult for the child using them. If the text materials used are too difficult or deviate from the actual grade level the student becomes frustrated. Comprehensive evaluation techniques can eliminate the use of inappropriate materials that cause possible frustrations, because the actual grade level of the book can be determined. From the evaluation, the appropriate material can be matched with the child.

The greatest injustice to children, in their early reading experiences, is to introduce material too difficult for their grade level. Children recognize and build their vocabulary from words that are familiar to them. They then transfer from oral language to the written page. Naturally, the words that a child utilizes, when speaking, will be the building blocks of his/her reading vocabulary. Because of the child's transfer from oral to written, developmental readers must be based on guidelines for grade level designations and control of vocabulary. Reed (1970) reports the following:

The authors of basal readers have not recognized this need, because they do not view reading as a two stage process: First, acquiring the skill to recognize written language that is already familiar to the child in oral form and second, applying this skill to expanding one's control of the language, including its vocabulary (p. 98).

This two stage process of reading is the guide for writing a well developed reading series. When these stages are ignored, the

student may develop a negative attitude toward reading.

Educators are confronted every day with students who cannot read for one reason or another. Whether or not the problem lies within the student, within the instruction, or within the material careful evaluation must be implemented. If the child is exposed to inappropriate material, his/her chance of becoming an experienced and willing reader is decreased.

Frustration and lack of interest are becoming increasing problems in our classrooms today. When the text is not readable the reader soon loses interest in not only the reading book, but also in other reading material. Teachers must learn to utilize comprehensive evaluation techniques to test development reading series in order to make appropriate selections.

Statement of the Problem

The purpose of this study was to use comprehensive evaluation to determine the following:

1. Is there a difference between the publishers' stated grade levels of five developmental reading series and the actual readability according to readability scales?
2. Is there a difference between the grade levels of the new words introduced and the grade level designations from a graded word list?
3. Are the new words introduced, reinforced at least five times within the story of introduction?

The comprehensive evaluation techniques involve: (1) computing the actual readability of text books; (2) determining the grade levels of the new words introduced, and (3) measuring the number of reinforcements of these new words. These aspects of comprehensive evaluation have a direct effect on the actual grade level of the books in each developmental reading series.

Definitions of Terms

1. Comprehensive evaluation is defined as an evaluation technique for developmental reading programs encompassing the use of readability formulas, graded word lists, and the Worksheet for Analysis of Instructional Materials Beyond Readability Scales (Schindler, 1980).

2. Comprehensive readability is defined as the grade level of books computed by readability formulas, the grade levels of the new words introduced, and the number of times the new words are reinforced in the published material. Other factors of printed material may also be a part of comprehensive readability.

3. Developmental reading program is defined as a published sequence of text books, teacher manuals, and student workbooks. Each book is assigned a grade level by the publisher. These programs are designed to teach specific skills such as word attack, vocabulary, and comprehension in reading.

4. Factors are the characteristics of published materials that have an effect on actual readability of that material. The readability factors are: vocabulary words, frequency of word reinforcement,

typography, interest-value, style of writing, and sentence length.

5. Graded word lists are published word lists with separate grade levels attached to each group of words. Some lists are separated into categories with individual grades attached and others are recommended for a set number of grades.

6. New words are the new vocabulary words introduced before each new story in a developmental reading series.

7. Readability is defined as the actual grade level calculation derived from readability formulas. These formulas were developed as a tool to calculate the actual grade levels of published text materials. It is the term used for how readable the material is in printed form.

A. Actual readability is the calculated grade level by readability formulas.

B. Estimated readability is the grade level designation publishers place on the published text books.

8. Reading can be defined as the act of extracting meaning from the printed page. Some skills involved in reading are: word attack, vocabulary recognition, and comprehension.

✓ 9. Reinforcement is the frequency of new vocabulary words in a basal text story after introduction. "The recommended number of reinforcements is five repeated citations of the word within the story of introduction" (Kingsley, 1965).

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

11. Worksheet for Analysis of Instructional Materials Beyond Readability Scales. From this worksheet, the new words in the series can be listed and the frequency reinforcement can be tabulated. Also, skills that are printed in the teachers' manual can be recorded and subjectively evaluated (Schindler, 1980).

Assumptions

The following is a list of assumptions that are commonly made about developmental reading programs:

1. The developmental reading program is written with the oral language of all students taken into consideration.
2. The publishers' estimates of readability or grade levels for each book in the series are accurate.
3. The grade levels of the new words introduced in each text book of a developmental series are appropriate for the grade level of each book.
4. The new words introduced are reinforced sufficiently in the published stories according to Kingsley's (1965) recommendations of five repetitions.
5. It is not the responsibility of the teacher to evaluate all aspects of reading materials before using them with children.

Questions to be Explored

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

2. Does the new word analysis match the actual computed readability?
3. What percentage of new words introduced in the text are on or below the publisher's estimated readability?
4. What percentage of new words are reinforced at the recommended five times or more than five times in the text material?
5. Does the use of the comprehensive evaluation process suggest reasons for discrepancies of the publisher's estimated readability and the actual readability scores based on readability formulas?

CHAPTER II

RELATED LITERATURE

Readability

The Aspects of Readability

Looking at today's reading curricula, the developmental reading series seems to have become the basis for reading instruction. From the beginning of a child's reading career, exposure to different kinds of material is required. Since the developmental reading series is the starting point for most children, close evaluation of these materials is essential. "At the beginning of reading instruction, the pupil understands spoken materials at some level, while he understands the same materials in print not at all" (Gleitman and Rozin, 1977, p. 6). According to this statement, reading is transferring from one's own oral language to the printed page. Teachers assume that developmental reading programs are written with this fact in mind, and that the beginning reader is first priority. The content of material must be written so that it is easily read and understood by the young child.

The understanding of reading material is related to the comprehensive readability of the text. Comprehensive readability means not only the grade levels of each book in a series, but also the grade levels of the words introduced, and the number of times these 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).

Rather than relying on experience alone, educators can use readability formulas to verify the actual grade levels.

Klare (1963) also defines readability in three ways:

1. An indication of legibility of either handwriting or typography.
2. An indication of the ease of reading due to either the interest-value or the pleasantness of writing.
3. An indication of the ease of understanding or comprehension due to the style of writing (p. 1).

All of these factors of readability, as defined by Klare (1963), affect the actual readability of published materials. The factors of comprehensive readability addressed in this study were: grade levels of books, grade levels of new words, and the frequency of new word reinforcements.

A second concern regarding comprehensive readability is the vocabulary used and the correct placement of vocabulary by grade level. Naturally the vocabulary has to be controlled with consideration of how extensive the beginning reader's oral vocabulary is. One way that publishers have attempted to solve this problem is through studies involving children in various grades. A group of words given to the children were used to determine what words were in their sight vocabulary. Then the words were placed in the developmental reading series according to the grade levels of the children. The difficulty with this type of selection procedure is that only a small population

of children is tested. The readability of text material is often suited for only a small population of children, and ignores all other types of beginning readers not sampled.

Along with this aspect of comprehensive readability is the need for reinforcement of new vocabulary words. The new words should be reinforced in context at least five times (Kingsley, 1965). Reinforcement of vocabulary does enhance actual vocabulary retention.

Readability in Relation to Teaching Reading

Developmental reading series have become the backbone of many reading programs. If there were no series, many children would receive even more inadequate instruction in relation to vocabulary, sequenced skills, and basic comprehension. Many teachers need a guide for reading instruction and therefore will continue to be dependent on developmental programs.

Since there is so much reliance placed on developmental reading programs, careful judgements must be made on their readabilities. All components of a developmental series must be developed in accordance with several restrictions which include vocabulary restrictions, language restrictions, and concept restrictions. Reed (1970) believes that:

The principal characteristics of the basal series that serve to distinguish them from other types of beginning reading textbooks are, first, rigid control of the introduction of new vocabulary items, and second, the view that the process of reading, and the use of reading are essentially one and the same thing, so that reading becomes a developmental skill to be taught from first grade through high school. Both of these characteristics deserve careful scrutiny (p. 98).

Here, the key phrase is "developmental skill" in reference to the actual teaching of reading. Indeed the series must have rigid control of vocabulary, as well as organized levels in which it is written. The readability and the vocabulary both influence the actual writing of a series so that it reflects a developmental sequence. Any concept learned is done so by building from one skill to developmentally harder skills. This same sequence is involved in learning to read.

When a developmental reading program is published it has three conjunctive parts. These three parts should reflect Reed's characteristic analysis of the basal series. Chall (1967) reports the following about basal series:

Actually, a basal-reading series is one of the most expensive and extensive ventures a publisher can undertake. It attempts to give teachers and pupils a "total reading program" embodying a system for teaching reading (in the teacher's manuals), a collection of stories and selections for pupils to read (the readers), and exercise for additional practice (the workbooks) (p. 187).

Along with what attempts are made by the publishers, great importance is placed on the teacher and how the series are utilized to teach reading. Constant restructuring has to be carried out to make the program work with children. Developmental reading series are not written on an individual basis so teachers must constantly ask whether or not the grade levels are correct, whether the words are too easy or too hard, and whether the words are reinforced for learning?

Chall interviewed three authors about the grade levels of reading programs. All three were in hope that teachers would gear the instruction to the child and not hold a child back with too simple a

reader. Chall (1967) also stated, "However, two were also concerned least the bright child be advanced too fast, for even the bright children often lack the skills presented in the workbooks and teacher's manual" (p. 197). If the publishers are relying on the teacher to scrutinize the use of their series, there has to be a control of the actual level of each book in the series.

Unfortunately, the publishers do not utilize this control, so it becomes the responsibility of the person using such series to calculate the grade levels and adjust the material for the child (Spache, 1978). The key to Spache's studies is that publishers make a judgement on the grade levels and they do not actually check the materials before they are delivered to our school systems.

Spache's concentration on grade levels of books in developmental reading series has become both important and useful. His findings raise enough questions about developmental texts, that further observation of reading material is mandatory.

Readability is one of the most prominent factors of a developmental reading series in both the higher grades and the early stages of reading. Along with the concern about readability of text books is the concern for controlled vocabulary. In writing a developmental reading series, the vocabulary introduced in the primer books forms a basis for vocabulary development throughout the series. This is why control of vocabulary is so important in the early exposure to reading. The vocabulary should not only be designed for the grade level being taught, but also should be designed for the language ability of the

particular group being taught. Gillespie and Johnson (1974) state that "If the child has had no experiences with the words in the basal readers or has never heard the words spoken, he may have problems in comprehending the materials" (p. 62). These problems then extend to the child's attitude toward reading and also inhibit his/her desire for outside reading. It is of extreme importance that the vocabulary chosen for a developmental series matches the age of the children that would be learning from such series.

From studies done by Jeanne Chall (1967), publishers seem to agree that vocabulary should be controlled in developmental reading series.

Upon entering first grade the typical English speaking child has about four thousand "concepts"; he has probably been familiar with the concepts in the typical conventional basal reader for grades 1, 2, and 3 since the age of three. The hurdle for the first grader is to learn to recognize in print the concepts he already knows. This is no small feat. Therefore, all programs require some control in introducing words (p. 69).

Reed (1970) offers another idea toward vocabulary:

As long as the number of new words introduced per page is strictly limited, the basal series have no concern with the question of whether the words are already part of the child's vocabulary. The only gesture they make in this direction is to concentrate on short words in the early stages of teaching reading (p. 98).

As can be seen from the statements above, several factors enter into the type and frequency of vocabulary words in a series. Chall and Reed's concern is the actual grade level of the words introduced before each story. No more than 5 percent new words should be introduced at a time, up to the third grade; however, they can be up to one

grade level above the stories' grade level. Consideration also has to be given to the instructional level of teaching. Ekwall (1977) defines the instructional level of learning as "comprehension should average 75% and word recognition should average 95%" (p. 4). Any percentage below this will result in frustration for the student.

Control of vocabulary deals not only with grade levels of words, but also with the frequency of the vocabulary in the text. Klare (1975) found that "Words of high (versus low) frequency and/or familiarity contribute to more readable writing; found for nouns particularly, but also for other content words" (p. 15). This statement supports the idea behind readability of developmental readers and the frequency of their vocabulary. When the beginning reader is exposed to familiar words, already in his/her oral vocabulary, it can naturally be assumed that these words will be acquired with ease. The number of times the new words are reinforced influences the reader's comprehension and recognition of those words to be taught.

Comprehensive Evaluation of Readability

If educators can recognize the fact that the vocabulary in readers is not accurately controlled and that they are not always on the level which is reported, many wasted hours in the classroom can be eliminated. Several methods of evaluation are available to determine the validity of material. These techniques include: (1) computation of readability formulas; (2) use of graded word lists to make comparisons; and (3) use of a worksheet for further analysis. The use of

these methods will help the educator choose appropriate reading materials to best suit their students.

Readability Scales

Knowledge of various readability scales becomes an advantage to any teacher of reading. In existence today are several formulas that can be used to predict readability of material. Klare (1963) states that:

No formula is a perfectly accurate measuring device, but it must be constructed so that: (1) the sample measured will represent, with maximum accuracy, the entire piece of writing from which it was taken; (2) two or more analyzes of the same material whether made by the same or different analysis, will agree as closely as possible; and (3) formula scores will as nearly as they can, predict actual readability (p. 101).

The most important part of these characteristics is the validity of the formula. That is, does the measuring device actually measure what is needed to be measured?

The Spache Reading Formula and the Dale-Chall were chosen specifically for this study. These accurate readability formulas can be used simply to determine the grade level of the material a teacher is using. The Spache Formula was used for the books graded one through three and the Dale-Chall Formula was employed for books graded four and five.

The two formulas used for this study were chosen because of their validity, word lists, and structure. Klare (1963) reports "By far the highest intercorrelations have been with the Dale-Chall scores, in one case being as high as .98" (p. 118). The intercorrelation means that several examples of the Dale-Chall Formula results have been

compared with each other and that the correlation proved to be very high and strong. Most other formulas, for readability, are compared with other similar formulas to find their validity. The Dale-Chall proved valid by comparing the scores within several samples.

The Spache Reading Formula was compared with the Bloomer Readability Formula. Klare (1963) reported that "The recent formulas of Spache and Bloomer have a correlation of .82 and .72 with their criteria, in both cases, text book grade levels" (p. 113). It was suggested that these scores are high because beginning readers do have some controlled vocabulary compared to later text books, and that the formulas were administered by publishers after the books were written. The coefficients .82 and .72 are still good indicators of the validity of the Spache Reading Formula.

The administration of these formulas requires the use of a word list to find the percentage of hard words in the text material. When calculating the Spache Reading Formula, the Stone's Revision of the Dale List of 769 Easy Words taken from Good Reading For Poor Readers, is used, and for the Dale-Chall Readability Formula, the Dale List of 3000 Familiar Words is utilized.

The importance of using both word lists becomes evident in looking at the process through which they were developed. The Dale List was constructed by testing fourth-graders on their knowledge in reading of a list of approximately 10,000 words. This is why the Dale List is used to calculate material fourth grade and higher.

For the lower grade levels, the Stone's List was utilized.

Klare (1963) reports the following:

Clarence Stone felt . . . that the accuracy of Spache's formula could be increased by revising the Dale list. He proposed that 173 of the 769 words be deleted and replaced by a similar number taken from L. L. Krantz's "The Authors Word List" and his own A Graded Vocabulary for Primary Reading. This revised list, he states, yields a lower rating than the original Dale List (p. 69).

Because of the lower rating, the Spache Reading Formula is used to calculate grade levels of text first through third grades. If a calculation from the Spache comes out 3.9 or higher, the calculations should be figured again according to the Dale-Chall Readability Formula.

New Words

Readability prediction by formula is only one of the aspects included in the comprehensive evaluation of text material. Just as important are the grade levels of the new words introduced. Graded word lists can be utilized to find the actual grade level of each word introduced in the text material. The graded word lists are obtainable from professional libraries or publishing companies. For the purpose of this study, A Functional Basic Word List for Special Pupils by Tudyman and Groelle (1963) was utilized. This list was chosen because of the reasoning behind its development. Tudyman and Groella (1963) state the following:

The reading ability or reading age of many thousands of children in school today is significantly below their chronological and social age. These children have difficulty in learning to read because they are mentally retarded, slow learning, or are in need of special remedial or corrective reading techniques. Consequently, there is an increasing demand for instructional reading materials for children having the specific difficulties mentioned (p. 2).

After reviewing the ideas behind the functional word list, the criteria for its compilation becomes important. Tudyman and Groelle (1963) report:

In compiling A Functional Basic Word List for Special Pupils, the authors have used two basic criteria: (1) the words selected should be those most frequently recurring in several of the more elementary basic reading series and in several standard word lists such as the Gates, the Rinsland, the Thorndike-Lorge, and/or (2) the words selected should be of special significance to mentally retarded and slow learning children (p. 4).

The words referred to in both criteria include survival words that are found on signs and labels, simple concepts of direction, numbers, time, and the monetary system. The lists reviewed by Tudyman and Groelle to form one list, were carefully scrutinized according to frequency of words on other word lists and according to the experience of the authors. It should be noted that the words on the functional list can be used for experience charts, reading materials, and spelling experiences.

The composition of the functional list is seven separate lists, with a specific grade level attached to each; preprimer to fifth grade. This structure allows for easy checking of words according to specific grade levels.

A Worksheet for Further Analysis

Even further analysis of the material beyond readability calculations and word list comparisons can be beneficial. Using the Worksheet for Analysis of Instructional Materials Beyond Readability Scales a teacher can record the number of new words introduced in a story, and

tabulate the number of times the new words are reinforced. This evaluation sheet was developed by Schindler (1981) at The University of Tennessee. Schindler felt that an extra evaluation of developmental readers was needed beyond the calculation of readability grade levels. Schindler found that "for many years special education teachers were taught how to teach reading by the rule that children need a minimum of five reinforcements to learn a word" (1981). For a further check of this rule, Schindler developed a worksheet to tabulate reinforcements of words in published materials and the skills offered in the teacher's manual. Schindler (1981) goes further to say "to date, published materials do not reinforce words or skills like they should." Chall (1967) suggests the following that supports Schindler's rationale:

The words in the pupils' readers for grades 1, 2, and 3 should be repeated often. They should be carefully controlled on meaning-frequency principle; i.e., they should be the words that appear most frequently in general reading matter and that are within the child's listening and speaking vocabulary (p. 15).

Chall does not offer any limit on how many times the words should be reinforced but according to Harris and Sipay (1971), "Slow learners may require 30 or more repetitions before they master a new word" (p. 213). This requirement not only holds true for slow learners but could also be applied to "normal" learners as well. The new words should be reinforced for retention of the meaning without destroying the interest level of the material.

Gates made an investigation of different intelligence levels and discovered that even some bright children have difficulty learning new

words. Kirk reported Gates' findings on estimated reasonable vocabulary reinforcement for various intelligence levels.

"120-12 I.Q. - 20 repetitions

90-109 I.Q. - 35 repetitions

60-69 I.Q. - 55 repetitions" (Kirk, 1940, p. 93).

The importance of repetition can be seen from Gates' data.

The necessity of repetition for various I.Q. levels places a great responsibility on publishers as well as teachers. Harris and Sipay (1971) define repetition "as a means of helping the child to 'overlearn' the word so that he learns to recognize it rapidly" (p. 213). This overlearning happens only when the child is exposed to the new word in various forms whether visual or aural. It then becomes the task of the educator to take the new words in a developmental series and expand them for learning. Harris (1970) reminds educators that "To be effective, this repetition should not be monotonous drill, but should be presented in such a way as to maintain the child's interest at a high level and encourage accuracy of perception" (p. 319). Repetition must be handled carefully so that the child does not become frustrated with monotonous repetition or bored with drill.

In reviewing the findings of research on the repetition of new vocabulary words, it can be assumed that publishers realize its importance in reading. Betts (1957) reports disagreement with this assumption, "this repetition has been carried so far in some basal preprimers and in some reading type experience charts that the literacy quality has been reduced to the zero point" (p. 384). This

overextended repetition has been reduced in recent years. but it becomes a challenge to write very simple stories with true literary quality. Nevertheless, sentences such as Look, look, look, and Oh, oh, oh, should be eliminated from developmental reading series.

Summary

Developmental reading programs have become the backbone of reading instruction. Research indicates that published reading materials are not always accurate in grade level designations. Because of the dominance of reading programs and the evidence of research, the overwhelming need for a comprehensive evaluation technique is mandatory. The comprehensive evaluation process includes: (1) readability formulas to calculate grade levels; (2) graded word lists to check grade levels of new words; and (3) a worksheet to record new words with a frequency of reinforcements in the text material. With these tools, teachers can evaluate developmental reading programs and choose the most appropriate one for their students.

CHAPTER III

DESIGN FOR EVALUATION OF READING SERIES

Purpose

The purposes of this study were as follows:

1. To determine if there exists a difference between the publishers' estimated grade levels of five developmental reading series and their actual readability according to readability scales.
2. To determine if there exists a difference between the estimated grade level of new words and the level as evaluated by matching them with a graded word list.
3. To determine if there exists a difference between the number of times the new words are reinforced and the recommended five times.

Data Collection Procedures

The researcher chose five developmental reading series used by various school systems in the East Tennessee area. The series chosen were as follows: (1) Macmillan (1970); (2) Ginn (1961 and 1976); (3) Houghton-Mifflin (1978); (4) Distar (SRA 1974); and (5) SRA Corrective Reading Program (1974). Data were gathered on the five developmental series by Special Education students at The University of Tennessee. Many of the books were provided by the researcher and

others were acquired under requirements for major classes. Data were collected over a 15 month period.

Instrumentation

The following instruments were used to collect data for this study: (1) the Spache Reading Formula (see Appendix A); (2) the Dale-Chall Readability Formula (see Appendix B); (3) the Worksheet for Analysis of Instructional Materials Beyond Readability Scales (Schindler, 1980) (see Appendix C); (4) A Functional Basic Word List for Special Pupils (Tudyman and Groelle, 1963). All four of these instruments were chosen because of the face validity and ease of administration.

The Spache Reading Formula uses three samplings of 100 word passages from one book. The samplings are taken at random from the beginning, middle, and end of each book being studied. In the passages selected, the number of hard words not on the Stone's revised list, as recommended for use with the Spache Formula, were counted and then computed with various specified calculations. From the three passages analyzed, a grade level is found. The three grade levels are then averaged to find the actual grade level of the book.

Like the Spache Formula, the Dale-Chall uses three 100 word samplings. There is, however, a different word list, the Dale list, used to find hard words and the specified calculations are slightly different (see Appendices A and B).

The data gathering instrument, the Worksheet for Analysis of Instructional Materials Beyond Readability Scales was developed by Schindler (1980) to analyze developmental reading texts by other attributes. Again, the sampling of three sections, beginning, middle, and end, was used. The new words introduced for those stories were recorded in the far left column. The second and third columns were for recording the number of times each word was reinforced in the story and the next three stories proceeding (see Appendix C). This form also included recording the teacher's manual skills and whether or not they are taught. This section was not analyzed for the purposes of this study.

After recording the new words for each story, they were matched with A Functional Basic Word List for Special Pupils developed by Tudyman and Groelle (1963). From this comparison, grade levels were assigned to each word that appeared on the functional list and the other words that did not appear were assigned no grade equivalent.

Data Analysis

The data were analyzed by charting each developmental series according to readability scores, new word grade equivalents, and frequency of new word reinforcements.

For ease of comparison, a ratio was computed by placing publishers' estimated grade level over the actual readability level and then subtracting from 1. The number derived from this calculation was an index number for each book. A 20 percent confidence level was

utilized as the adequate tolerance for index numbers closest to 0, +.1 or -.1. A positive index value indicates that the actual readability was higher than the estimated readability, and a negative index value indicates that the actual readability is lower than the estimated readability.

Frequency of each new individual word introduced was categorized and charted for the grade levels of primer to fifth grade according to the graded divisions of the graded word list by Tudyman and Groelle. Also included was the frequency of words not on the list given the title of unknown. Percentages for each frequency were then calculated by the number of words in each grade designation over the total number of new words introduced in each book. Likewise, the frequency of the new words reinforcements was charted in three specified columns; less than five, on five, and more than five. Percentages were calculated by number of reinforcements over total number of new words introduced in each book. The results of analysis are explained in Chapter IV.

CHAPTER IV

RESULTS

Introduction

The purpose of this study was to comprehensively evaluate five developmental reading programs. The questions to be explored, from Chapter I, are divided into two separate sections. The first section contains questions one through four and is labeled Individual Analysis of Each Developmental Reading Program. The second section contains question five and is labeled Comparative Analysis of the Five Developmental Reading Programs.

The five developmental reading programs are dealt with in this order: (1) Ginn (1961 and 1976); (2) Macmillan; (3) Distar; (4) Houghton-Mifflin; (5) SRA Corrective Reading. As stated before, these series were chosen for their availability and popularity in the East Tennessee area.

It should be noted at this time that because of the structure and format of the SRA Corrective Reading Program, it could not be analyzed according to all four questions. This program is based on two separate sections: Comprehension A and Decoding B. Only the Worksheet for Analysis of Instructional Materials Beyond Readability Scales could be utilized for Comprehension A, and only readability scores could be computed for Decoding B. No additional data from the

worksheet could be gathered from this section, as indicated in Table 1.

Individual Analysis of Each Developmental Reading Program

The observations of the Ginn series according to the first four questions to be explored are as follows:

Ginn.

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

In general, seven of the eight books in the Ginn series are below publisher's estimated readability according to Spache and Dale-Chall readability calculations. The text books labeled first grade and fourth grade prove to be the furthest from publisher's estimated grade level according to the calculated index numbers. The first grade text shows .53 index meaning that the actual readability was much higher than the publisher's estimated readability. However, the fourth grade text shows a -.82 which means that actual readability is much lower than the publisher's estimated readability. From this computation, the first grade text seems to be harder while the fourth grade text is far too easy (see Table 2).

2. Does the new word analysis match the actual computed readability?

An analysis of the individual texts of the Ginn series indicates no consistent relationship between the new word analysis and the actual readability. Book 1.0, which has a positive index number of .53, shows

TABLE 1

ESTIMATED AND ACTUAL READABILITY, INDEX, FREQUENCY AND PERCENTAGE NEW WORDS,
PERCENTAGE NEW WORDS REINFORCED FOR SRA CORRECTIVE READING TEXT (1974)

Text	Estimated	Actual	Index	New Words Introduced		Percent Reinforced		
				F	Percent On/Below	Percent Above/UK	Less	5 More
Comprehension A	1	Could not calculate readability		7	14	86	56	14 30
Decoding B	2	3.1	.35					

TABLE 2

ESTIMATED AND ACTUAL READABILITY, INDEX, FREQUENCY AND PERCENTAGE NEW WORDS,
PERCENTAGE NEW WORDS REINFORCED FOR GINN TEXTS (1961 AND 1976)

Text	Estimated*	Actual	Index	New Words Introduced			Percent Reinforced		
				F	Percent On/Below	Percent Above/UK	Less	5	More
May I Come In	1.0	2.15	.53	17	64	36	48	17	35
On Cherry Street	2.0	1.70	-.18	10	90	10	80	10	10
One to Grow On	2.5	2.43	-.03	32	75	25	78	13	9
Finding New Neigh- bors	3.0	2.79	-.08	21	34	67	52	0	48
How It Is Nowadays	3.3	2.81	-.17	20	50	50	85	5	10
Inside Out (1976)	3.7	2.47	-.23	24	13	87	96	0	4
A Lizard to Start With	4.0	2.20	-.82	31	6	94	97	0	3
Tell Me How the Sun Rose	5.0	4.40	-.14	28	15	85	96	4	0
MEANS			-.14		43	57	79	6	17

*Publisher's designated readability.

64 percent of the new words on or below grade level and 52 percent of the new words reinforced five or more times. The other six books of the series all have negative index values and vary considerably in the percentages of the new word analysis. Since there is no consistency of data, other factors of readability seem to influence the discrepancy between actual readability scores and the publisher's estimated readability (see Table 2 and Figure 1).

3. What percentage of new words introduced in the text are on or below the publisher's estimated readability?

In general, there is a decrease in the percentage of new words introduced on or below grade level in comparison to the increase of publisher's estimated grade levels. As the sequence of text becomes harder according to estimated grade placement, the percentage of new words on or below grade level introduced decreases (see Figure 1).

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

For the Ginn text series, less than one half of the new words are reinforced at the recommended five times or more than five times with the exception of books 1.0 and 3.0. Text book grade 1.0 shows 52 percent of the new words are reinforced five or more times and text book 3.0 shows a marginal 48 percent reinforced five or more times (see Table 2). All other books fall substantially below the recommended five reinforcements.

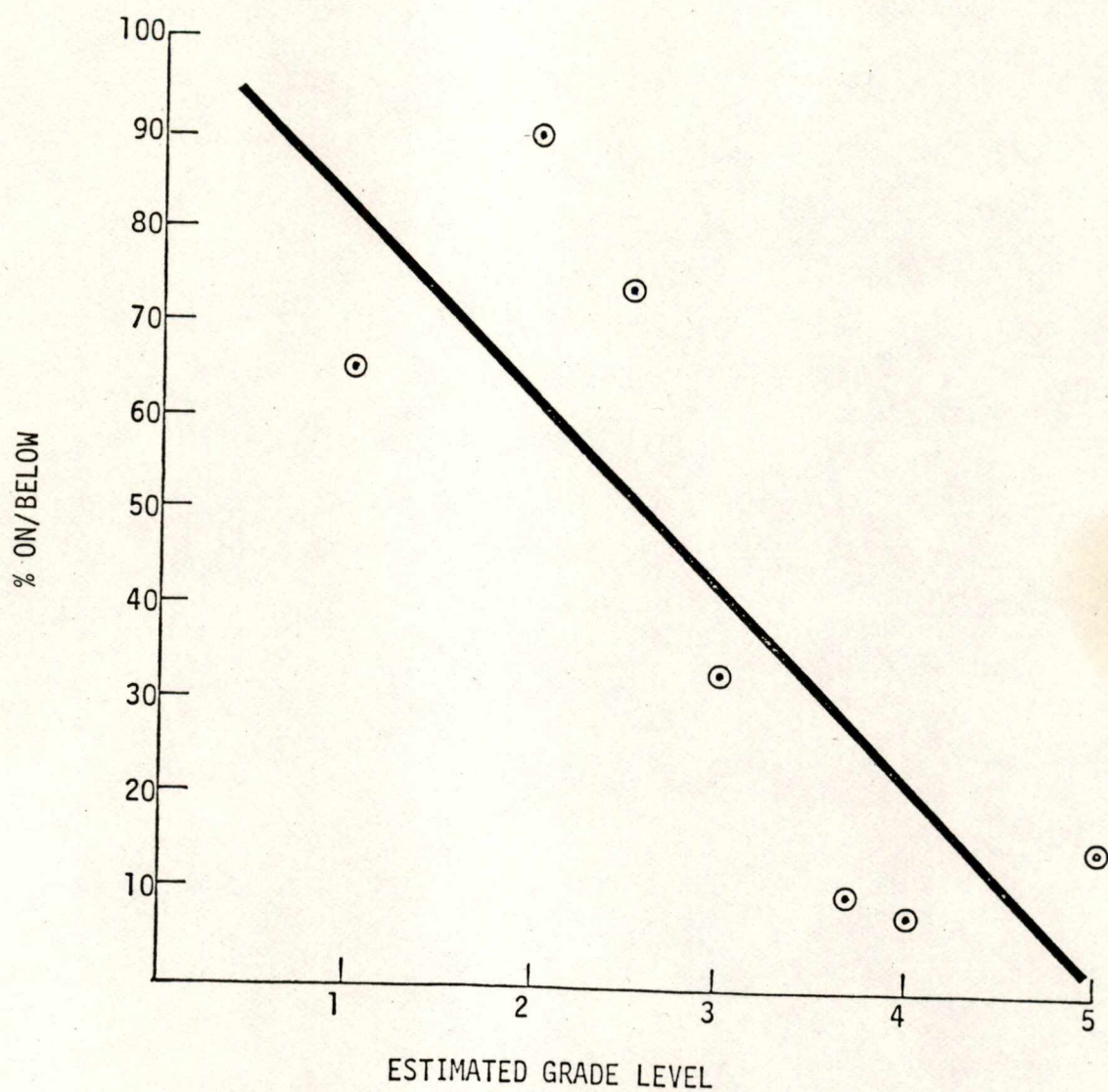


FIGURE 1. GINN TEXT: ESTIMATED GRADE LEVEL VS. NEW WORDS INTRODUCED. (% ON OR BELOW)

The observations of the Macmillan Text series, according to the first four questions, are as follows:

Macmillan.

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

According to the Spache and Dale-Chall actual readability calculations, the Macmillan series has overall higher actual readability scores in comparison to the publisher's estimated readability scores with the exception of book 3.0. Book 1.0 has the greatest discrepancy between the publisher's estimated readability and actual calculated readability with an index number of .59 (see Table 3).

2. Does the new word analysis match the actual computed readability?

A relationship can be shown between the actual readability and the summation of new words introduced with the percentages of words above grade level. This comparison (and others attempted) indicates a nonlinear but direct relationship. Purely objective summations are not possible because there is no known relative influence between new word incidence and number (or percentage) of words above grade level.

3. What percentage of new words introduced in the text are on or below the publisher's estimated readability?

The Macmillan Series had a decrease of new words introduced on or below grade level as the publisher's estimated readability increases with the exception of book 4.0 which has 62 percent new words on or below grade level (see Figure 2).

TABLE 3
ESTIMATED AND ACTUAL READABILITY, INDEX, FREQUENCY AND PERCENTAGE NEW WORDS,
PERCENTAGE NEW WORDS REINFORCED FOR MACMILLAN TEXT (1970)

Text	Estimated	Actual	Index	New Words Introduced			Per Cent Reinforced		
				F	Percent On/Below	Percent Above/UK	Less	5	More
Lands of Pleasure	1.0	2.40	.59	11	54	45	81	0	18
Enchanted Gates	2.0	2.52	.21	21	48	52	71	0	29
Shining Bridges	3.0	2.68	-.19	26	42	57	85	0	15
More Than Words	3.0	3.45	.13	13	7	83	62	23	15
Better Than Gold	4.0	4.30	.07	21	62	38	100	0	0
Bold Journeys	5.0	5.50	.09	38	25	74	100	0	0
MEANS			.85		40	59	83	4	13

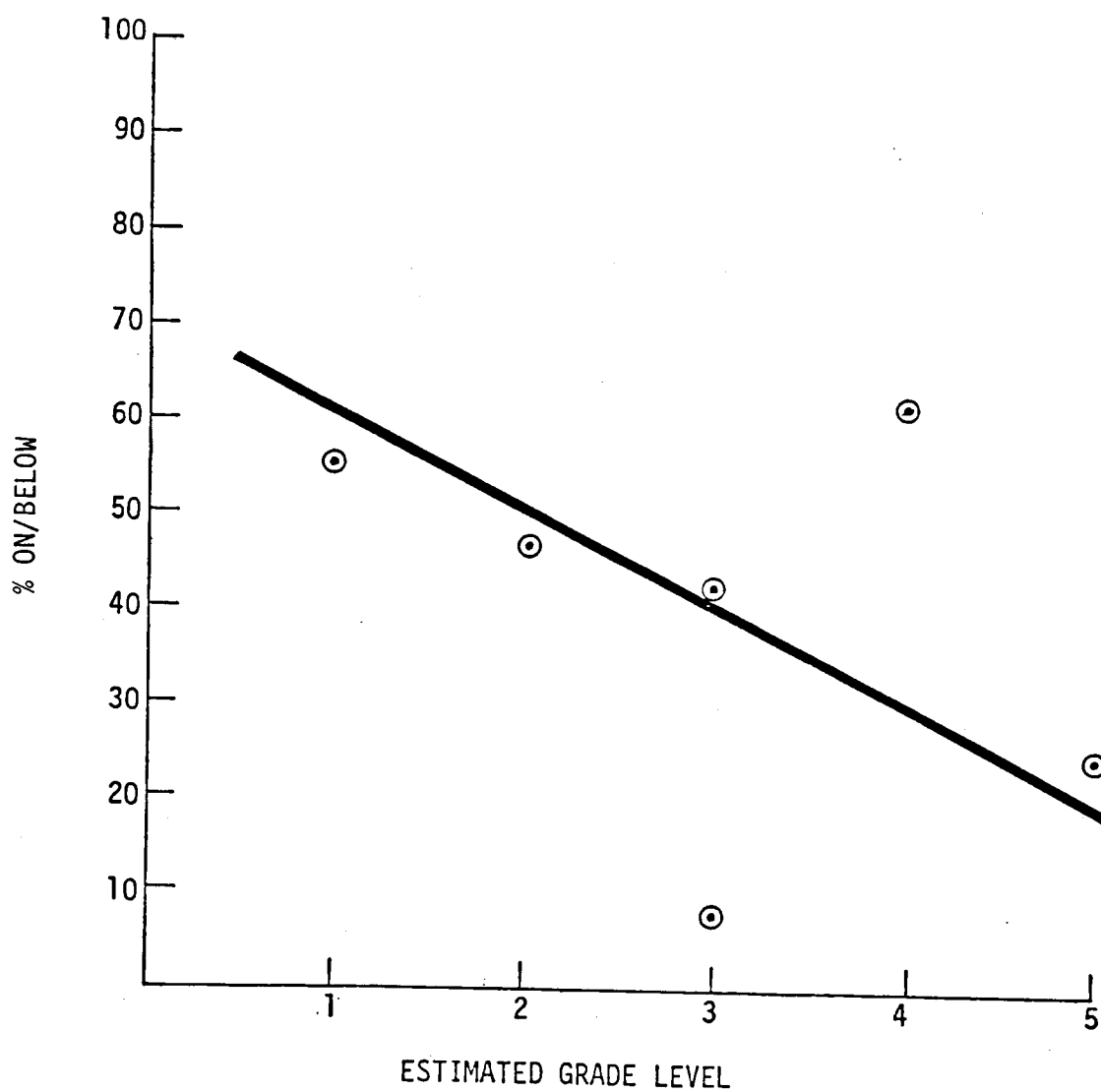


FIGURE 2. MACMILLAN: ESTIMATED GRADE LEVEL VS. NEW WORDS INTRODUCED (% ON OR BELOW)

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

For all of the Macmillan series less than one half of the new words are reinforced at the recommended five times or more. The only book that reinforces the new words five times or more is book 3.0, More Than Words, which has 23 percent on five and 15 percent above the recommended five reinforcements. The other five books have not reinforced words five times.

The observations of the Distar text series, according to the first four questions, are as follows:

Distar.

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

All of the books for the Distar series are higher than the estimated readability level according to the Spache Reading Formula with the exception of Distar III--Storybook A, which is lower than the estimated grade level. It appears that all books estimated as first grade are substantially higher in actual readability than the publisher's estimate, showing index numbers of .52, .50, and .55 (see Table 4). Distar II--Storybook IV has an index of .33 which indicates it is the hardest of the four books on grade level 2.0.

2. Does the new word analysis match the actual computed readability?

From this comparison, the books estimated at 1.0 grade level are at least one grade level higher as analyzed by the Spache formula.

TABLE 4

ESTIMATED AND ACTUAL READABILITY, INDEX, FREQUENCY AND PERCENTAGE NEW WORDS,
PERCENTAGE NEW WORDS REINFORCED FOR DISTAR TEXT (SRA 1974)

Text	Estimated	Actual	Index	New Words Introduced		Percent Reinforced		
				F	Percent On/Below	Percent Above/UK	Less	5 More
Distar I Storybook I	1.0	2.1	.52	9	11	89	100	0
Storybook II	1.0	2.0	.50	6	0	100	66	0
Storybook III	1.0	2.2	.55	7	0	100	100	0
Distar II Storybook I	2.0	2.3	.13	12	33	67	83	0
Storybook II	2.0	2.5	.20	29	21	79	86	0
Storybook III	2.0	2.4	.17	51	18	82	91	2
Storybook IV	2.0	3.0	.33	40	25	75	88	5
Distar III Storybook A	3.0	2.7	-.11	48	42	58	70	2
Storybook B	3.0	3.1	.03	102	49	51	89	1
Storybook C	3.0	2.5	.20	177	45	55	92	2
MEANS			.25		24	75	87	1
								12

The correlation becomes noticeable on the new word analysis. Of the new words introduced in Storybook I only 11 percent are on or below grade level and Storybook II and Storybook III have no words at the appropriate grade level in Distar I. These percentages tend to influence the discrepancies in the actual readability and the publisher's estimated grade level. All other texts in the series show some discrepancies with the exception of Distar III--Storybook A, which shows no direct relationship between word analysis and actual computed readability.

3. What percentage of new words introduced in the text are on or below the publisher's estimated readability?

The trend from both sets of figures shows that, as the publisher's estimate of readability increases by grade levels, the percentage of new words on or below grade level also increases. Again, the lower grade level books showed the greatest discrepancies in the two scores (see Figure 3).

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

All of the books in the Distar reading series reinforce less than 50 percent of their new words the recommended five times or more. The number of new words reinforced less than five times ranges from 66 percent to 100 percent for all books. For Distar I--Storybook I and Storybook III, none of the nine and seven new words introduced

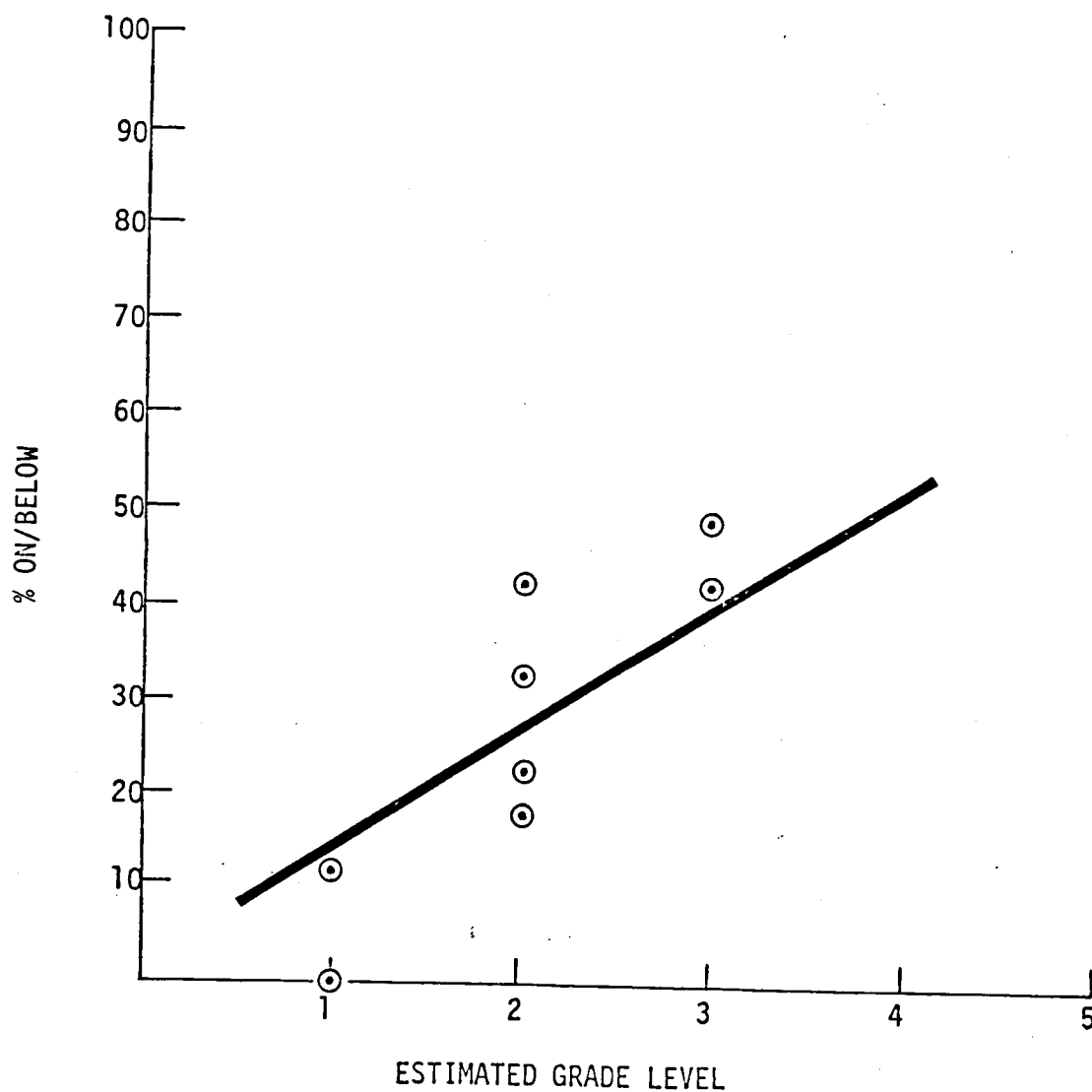


FIGURE 3. DISTAR TEXT: ESTIMATED GRADE LEVEL VS. NEW WORDS INTRODUCED (% ON OR BELOW)

respectively are reinforced five times. In fact, out of the nine new words introduced in Storybook I, five were not reinforced, two were reinforced only one time, and two were reinforced only two times. Likewise, out of the seven new words in Storybook III, four were not reinforced, two were reinforced only one time, and only one was reinforced two times. The closest books to the recommended five reinforcements are Distar I--Storybook II, with 33 percent over five, and Distar III--Storybook A, with 31 percent on or above five repetitions.

The observations of the Houghton-Mifflin series, according to the first four questions, are as follows:

Houghton-Mifflin.

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

All of the books evaluated according to the Spache and Dale-Chall formulas have a higher actual readability than the publisher's estimate of readability. The greatest discrepancy in readability scores is for the first grade reader with a .32 index number. The other books all are within the 20 percent tolerance selected for this study with the exception of book 2.3 with an index value of .23 (see Table 5).

2. Does the new word analysis match the actual computed readability?

Books 2.3-6.0 in the Houghton-Mifflin series all introduced more than 50 percent of the new words above grade level or unknown. This could influence the discrepancy between the readability scores, because

TABLE 5

ESTIMATED AND ACTUAL READABILITY, INDEX, FREQUENCY AND PERCENTAGE NEW WORDS,
PERCENTAGE NEW WORDS REINFORCED FOR HOUGHTON-MIFFLIN TEXT

Text	Estimated	Actual	Index	New Words Introduced		Percent Reinforced				
				F	On/Below	Percent	Above/UK	Less	5	More
Cloverleaf	1.5	2.2	.32	24	62	38	92	4	4	4
Sunburst	2.3	3.0	.23	39	28	72	94	3	3	3
Tapestry	2.7	2.9	.07	34	12	88	97	3	3	0
Passports	3.5	3.6	.03	28	8	92	89	7	4	4
Medley	4.0	4.9	.18	43	22	77	97	1	2	2
Impressions	6.0	6.7	.10	24	4	96	92	4	4	4
MEANS			.15	23		77	94	4	4	3

they all were calculated to be higher than the publisher's estimated grade levels. Since the number of new words introduced has a higher percentage in the above grade level or unknown frequency, the actual readability is not consistent with the publisher's estimated readability. The only book that shows no direct relationship is the first reader. It not only has the highest index number but also the highest percentage of words on or below grade level (see Table 5).

3. What percentage of new words introduced in the text are on or below the publisher's estimated readability?

As the publisher's estimate of readability of the Houghton-Mifflin books increases, the number of new words on grade level decreases with the exception of book 4.0. In book 4.0 only 22 percent of the new words introduced are on or below grade level (see Figure 4).

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

In the Houghton-Mifflin text series, only 11 percent-33 percent of the new words introduced are reinforced at the recommended five or more times for all books evaluated. Book 3.5 has the most words reinforced adequately with 7 percent reinforced five times and 4 percent reinforced more than five times. All other books in the series range from a total of 8 percent-3 percent new words reinforced five times or more. These figures may influence the readability of the material (see Table 5).

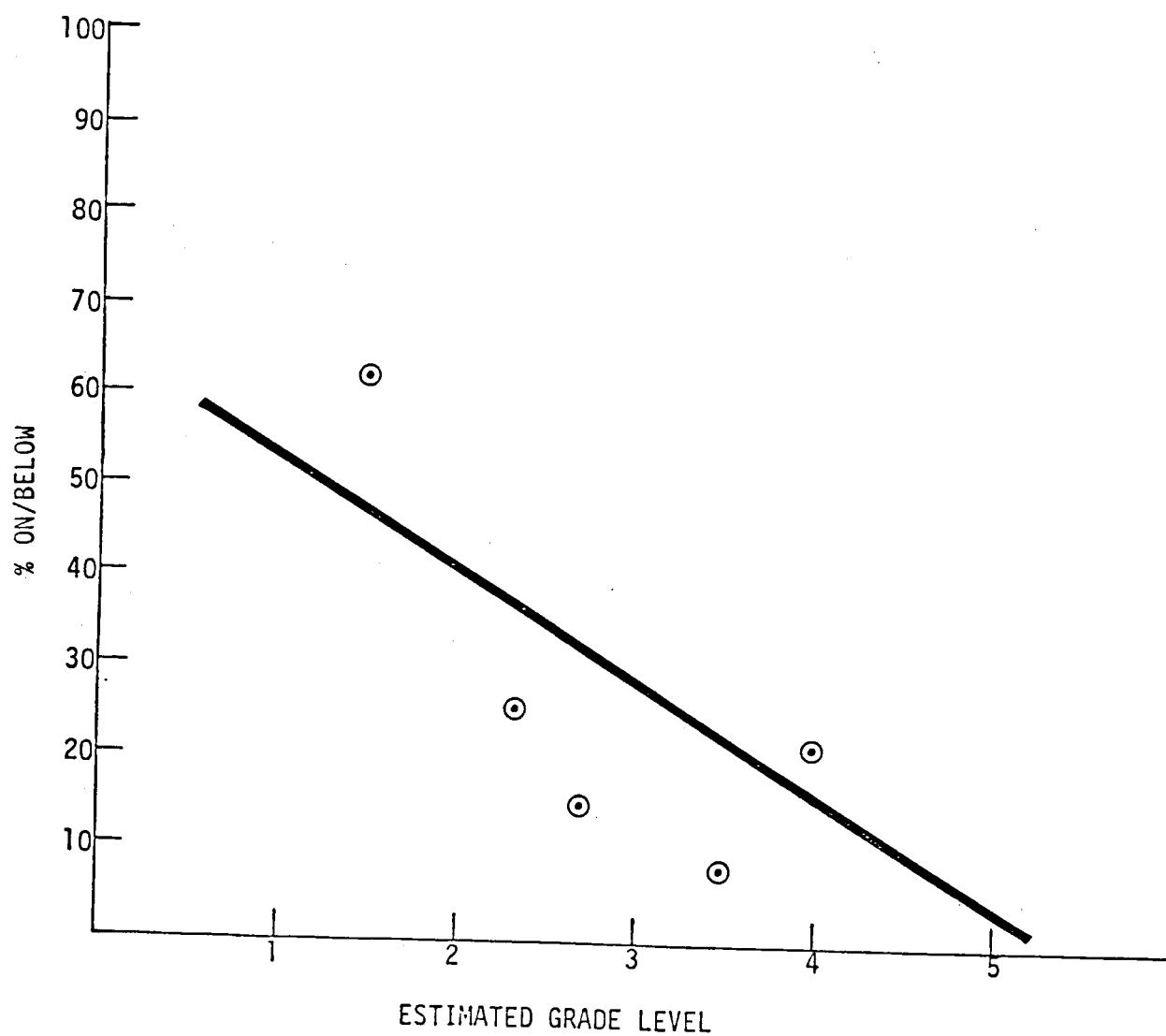


FIGURE 4. HOUGHTON-MIFFLIN: ESTIMATED GRADE LEVEL VS. NEW WORDS INTRODUCED (% ON OR BELOW)

The observations of the SRA Corrective Reading Program, according to questions one through four, are as follows:

SRA Corrective Reading.

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

As stated earlier, no readability scores could be calculated for Comprehension A because of the format of the text. However, Decoding B computed higher than the publisher's estimated readability with an index of .35. According to the Spache formula, the actual readability shows a substantially higher readability than the publisher's estimate (refer again to Table 1, page 29).

2. Does the new word analysis match the actual computed readability?

Adequate data could not be gathered to answer this question.

3. What percentage of new words introduced in the text are on or below the publisher's estimated readability?

For the text Comprehension A of the SRA series, 14 percent of the new words were on or below grade level. However, this is not a good sample since there were only seven new words introduced. The percentage of new words above or with an unknown grade level was 86 percent, which could have escalated the actual readability. Because of the structure of the manual, no actual readability could be calculated.

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

Comprehension A has 56 percent of the new words reinforced less than five times. This means that 44 percent of the new words are reinforced the recommended five or more times (refer again to Table 1).

Comparative Analysis of the Five Developmental Reading Programs

The five developmental reading programs were analyzed across text in order to answer the fifth question to be explored, from Chapter I.

Does the use of the comprehensive evaluation process suggest reasons for discrepancies of the publisher's estimated readability and the actual readability scores based on readability formulas?

Table 6 shows mean values for mean index numbers, mean frequency of words, mean percentage of new words, and mean percentages of reinforcements, for all five developmental reading programs. These means were used to compare the five series to one another.

The Ginn mean scores do not reflect the discrepancy between the publisher's estimated readability and the actual readability score, according to the comprehensive evaluation procedures. The mean index, -.14, indicates that the actual readability levels of the books were obviously lower than the publisher's estimate of readability. However, mean percent new words on or below grade level was 43 percent and those

words above or unknown were 57 percent. Of the 57 percent, 16 percent were above the publisher's estimated readability and 41 percent were not on the list. Also, only 21 percent of the new words reinforced were in the five or more times frequency. These percentages indicate that inadequate reinforcement of words could be one of the reasons for discrepancies between the publisher's estimated readability designations and the actual readability calculations of books. However, other factors thought to influence readability could be relevant.

For the Macmillan, Distar, and Houghton-Mifflin mean text scores, the mean index numbers indicate that the actual readability was higher than the estimated readability (see Table 6). The books in these series are more difficult than the publisher's estimated readability. Question number 5 cannot be accurately addressed regarding the SRA Corrective Reading Program because sufficient data are unattainable.

The comprehensive evaluation procedures indicate all three series have the highest mean percentages of new words in the above or unknown frequencies and they have the highest percentage of reinforced words in the less than five frequencies. From these comprehensive analyses, the discrepancies between publisher's estimated readability and the actual readability appear to be directly related to the new word data. The text books are harder because more words are introduced above grade level and are not reinforced the recommended five times or more than five times.

TABLE 6
 MEAN INDEX, FREQUENCY, PERCENTAGE NEW WORDS, PERCENTAGE OF NEW WORDS REINFORCED

Series	Mean Index	Mean F	Mean Percent New Words		Mean Percent Reinforced	
			On/Below	Above/UK	Less	More
Ginn (1961-1976)	-.14	23	43	57	79	15
Macmillan (1970)	.15	22	40	60	83	13
Distar (SRA 1974)	.25	48	25	75	87	12
Houghton-Mifflin (1978)	.16	32	23	77	93	3
SRA (1974)						
Comprehension A		7	14	86	56	14
Decoding B	.35					30

CHAPTER V

CONCLUSIONS

Discussion and Conclusions

The purpose of this chapter is to discuss the results of the comprehensive evaluation process in relation to the five developmental reading programs. This information gathered in Chapter IV indicates insufficient comprehensive readability for all five series.

The methods used in this study to evaluate comprehensive readability enabled the researcher to explore the text and formulate conclusions as to their appropriateness in the reading curricula. The across-text analysis showed Houghton-Mifflin, Macmillan, and Distar books higher in actual readability than the publishers' estimate of readability. Analysis of the Ginn series indicated the opposite in that the across-text readabilities were generally lower than the publisher's estimate of readability. However, in all text series, with the exception of the SRA Corrective Reading Program, the first grade books showed the greatest discrepancies between actual grade level and the publishers' estimated grade levels. All of the first grade readers were substantially harder than the publishers' reported grade levels. These discrepancies of readability in the first grade readers lead the researcher to believe that other factors have influenced the difficulty of these materials. These other factors could be sentence structure, style of writing, or the length of the sentences.

The use of comprehensive evaluation procedures provided a broader evaluation inclusive of graded new words and the frequency of the reinforcements. Ginn's first grade book had 64 percent of the new words on or below grade level and reinforced the new words 52 percent of the total recommended five or more times. These factors show no direct relationship to the high index score of the first grade book. The researcher believes that other factors influencing readability affected the discrepancy of grade level estimates. Such factors might be either the style of writing or the length of the sentences in that book.

The percentage of new words introduced in the first grade books of the Houghton-Mifflin and Macmillan series, on or below grade level, are higher than those words above or unknown. However, both series reinforce the new words only 8 percent and 18 percent, respectively, in the five or more than five frequency. This factor might influence the difficulty of the first grade readers in conjunction with style of writing or length of sentences.

The Distar reading series has the most direct relationship between the publisher's estimate of readability and the actual readability scores along with the graded new words with the reinforcements. All three of the first grade readers present the lowest percentage scores in the on or below frequency with 11, 0, and 0 percentages. The same results are shown in the reinforcement of new words. Again, 0, 33, and 0 percentages were calculated in the combined five or more reinforcement frequencies. These factors support the direct

relationship between readability scores and the new words introduced with frequency of reinforcements. Because the majority of new words introduced are above grade level and they are not reinforced properly in the text material, the text books produced have a higher readability level in comparison to the publisher's estimated readability.

No conclusions can be made about the SRA Corrective Reading Program according to both manuals. Comprehension A has only 14 percent of the seven new words introduced on or below grade level. Also, only 44 percent of the new words introduced are reinforced the recommended five or more than five times. These percentages lead the researcher to believe that the publisher's estimate of 1.0 is actually higher.

The comprehensive evaluation process of the five developmental reading programs proved to broaden the analysis beyond simple readability formulas. The most impressive result of the comprehensive evaluation process was evident in the results concerning the first grade readers. All of these readers were written substantially more difficult than a first grade reading level. The ideal text should be written on an instructional level which means 95 percent known to the child (Ekwall, 1977). These books were developed on frustration level, which would cause difficulties in reading for most children. A child's first exposure to reading needs to be the beginning of a successful reading career. If the child is frustrated by his/her first experiences in reading his/her attitude toward reading will be influenced in a negative direction. This is especially true for the slow learners or children in need of remedial reading.

From the results of this study, several factors can indeed influence the actual readability of developmental reading texts. Not all five developmental reading programs in this study showed a relationship between the comprehensive evaluation procedures and the actual readability. Therefore, other factors in the creation of these readers are evident. Chall (1967) reports the following statement of one author of basal text: "In teaching, we teach that reading is not just saying words. Even in the first grade, we try to give the pupil an appreciation of the difference between the real and the fanciful" (p. 54). The publisher's attempt to make the text material more interesting could influence the discrepancies between the actual readability and the publisher's estimate of readability.

Limitations

The limitations of this study are noticeable in two areas. First, the other factors that influence actual readability could not be evaluated by the comprehensive evaluation technique. Second, the SRA Corrective Reading Program could not be completely analyzed because of the structure of the material.

Not all of the five developmental reading programs showed a relationship between actual readability and the graded new words with frequency of the reinforcements. From these results, the researcher concludes that other factors of readability have influenced the discrepancies between actual readability scores and the publishers' estimated readability. These other factors could be sentence length,

style of writing, or sentence structure. The factors evaluated in this study and these other factors are all related to the actual readability of the text materials.

The second limitation is in the attempt to comprehensively evaluate the SRA Corrective Reading Program. A total analysis could not be performed because of the structure of this program. The comprehensive evaluation techniques are designed to evaluate basal type material. The SRA program is divided into Comprehension A and Decoding B manuals. No readability score could be calculated for Comprehension A and no comprehensive analysis was attainable from Decoding B. The lack of attainable data on this series limits any conclusions according to comprehensive evaluation.

Suggestions for Further Research

Observations of this study reveal additional possibilities for expanded research. The first option would be to use the instruments from this study to evaluate other available developmental reading series, as well as social studies, sciences, and arithmetic texts. The second option would be to evaluate developmental reading series according to the other factors thought to influence readability, for example sentence structure or style of writing.

Another option for additional research would be to evaluate the reading series by combining the factors from this study and the other factors thought to influence comprehensive readability. This type of study would broaden the analysis of all the factors in relation to the

actual readability. However, techniques would have to be developed to evaluate sentence structure and style of writing. The combination of comprehensive evaluation and the newly developed techniques would allow educators to make appropriate decisions as to the use of developmental reading programs.

Additional studies for further research could involve calculating several readability scores with one formula and comparing these scores to the publisher's estimate of readability. The most significant research study would be to explore techniques to evaluate the relationship between graded new words and the frequency of reinforcement, with the actual readability index number. This research study might show a more direct relationship between the actual readability scores and the comprehensive evaluation process. The results of this study and others would give educators valuable information about the actual readability of developmental reading programs. Educators must be familiar with expanded research before they choose a developmental reading program and use this program with their students. After all, our children's education is the most valuable investment we can make.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Betts, E. A. Foundation of Reading Instruction. New York: American Book Co., 1957.
- Chall, J. Learning to Read: The Great Debate. New York: McGraw-Hill Book Co., 1967.
- Clarence R. Stone's Revision of the Dale List of 769 Easy Words.
Spache, G. D. Good Reading for Poor Readers. Champaign, Ill.: Garrard Publishing Co., 1964.
- Clymer, T. Inside Out. Lexington, Mass.: Ginn and Co., 1976.
- 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.
- Ekwall, E. E. Locating and Correcting Reading Difficulties (2nd ed.). Columbus, Ohio: Bell and Howell Co., 1977.
- Englemann, S., and E. C. Bruner. Distar. Chicago: Science Research Associates, 1974.
- Gillespie, P. H., and L. Johnson. Teaching Reading to the Mildly Retarded Child. Columbus, Ohio: Merrill Publishing Co., 1974.
- Ginn Reading Program. Lexington, Mass.: Ginn and Co., 1961.
- 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.
- Harris, A. J. How to Increase Reading Ability. New York: David McKay Co., Inc., 1970.
- Harris, A. J., and E. R. Sipay. Effective Teaching of Reading. New York: David McKay Co., Inc., 1971.

Houghton-Mifflin Reading Program. Boston, Mass.: Houghton-Mifflin Co., 1978.

Kingsley, R. F. Adjustment of Curriculum for Handicapped Students. Presentation at Kent State University, Kent, Ohio, October 1965.

Kirk, S. A. Teaching Reading to Slow-Learning Children. Cambridge, Mass.: The Riverside Press, 1940.

Klare, G. R. The Measurement of Readability. Ames: Iowa State University Press, 1963.

Klare, G. R. A Manual for Readable Writing. Glen Burnie, Md.: REM Co., 1975.

Macmillan Reading Program. New York: The Macmillan Co., 1970.

Reed, D. W. A Glance at the Linguistic Organization of Elementary Reading Textbooks. In Doris V. Gunderson (ed.), Language and Reading. Washington: Center for Applied Linguistics, 1970.

Schindler, W. J. Worksheet for Analysis of Instructional Materials Beyond Readability Scales. Knoxville, Tenn.: The University of Tennessee, 1980.

Schindler, W. J. Personal communication, January 3, 1981.

Spache, G. D. Spache Reading Formula. 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.

SRA Corrective Reading Program. Chicago: Science Research Associates, 1974.

Tudyman, A., and M. C. Groelle. A Functional Basic Word List for Special Pupils. Pittsburgh, Penn.: Stanwix House Inc., 1963.

APPENDICES

APPENDIX A

Spache Reading Formula*

Worksheet for Application of the Spache Readability Formula for Grades 1 - 3

Article or Book _____ Date _____

Author _____ Publisher _____

Page _____ Page _____ Page _____ Page _____

From _____ From _____ From _____ From _____

To _____ To _____ To _____ To _____

1. Number words _____

2. Number sentences _____

3. Number words not on
Functional Word
List _____

4. Ave. Sentence Length
(Divide 1 by 2) _____

5. Per cent hard words
(Divide 3 by 1,
multiply by 100) _____

6. Multiply (4) by .141 _____

7. Multiply (5) by .086 _____

8. Constant .839 .839 .839 .839

9. Estimated grade
placement (Add 6,
7, and 8) _____

Average grade placement of _____ samples _____

Analyzed by _____

Date _____

*Good Reading for Poor Readers, George D. Spache. Garrard
Publishing Co., Champaign, Illinois, 1964.

APPENDIX B

Dale-Chall Readability Formula-Table II

Article:	No.	Page No.	Page No.
Author:		From	From
Publisher:		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 _____	
Average corrected grade-level	Checked by _____	Date _____	

APPENDIX C

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 stories</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

Melissa E. Watson was born in Knoxville, Tennessee on April 26, 1955. She attended a Knoxville City public elementary school and was graduated from Bearden High School in 1973. The following September she entered The University of Tennessee, and in August 1977 she received a Bachelor of Science degree in Special Education.

Melissa was employed by Anderson County public schools in August 1977 and has taught Hearing Impaired Multiple Handicapped children for four years. After one year of teaching, Melissa entered the Graduate School of The University of Tennessee, Knoxville in June 1978. She will receive a Master of Science degree in Special Education in August 1981.