

Unlocking Language

The Classic Readability Studies

William H. DuBay
Editor

Impact Information
Costa Mesa, California

Copyright

The following journal articles are in the public domain:

“Word Knowledge in the Elementary School” by Edward L. Thorndike, published November, 1921, in *Teachers College Record*. Vol. 22, No. 5, pp.334-370.

“A Method for Measuring the ‘Vocabulary Burden’ of Textbooks” by Bertha A. Lively and Sidney L. Pressey, published in 1923 in *Educational Administration and Supervision*, Vol. 9, pp. 389-398.

“An objective method of determining grade placement of children’s reading material” by Vogel, M. and C. Washburne, published in 1928 in *Elementary School Journal*, Vol. 28, pp. 373-381.

“A Technique for Measuring the Vocabulary Burden of Textbooks” by W. W. Patty, and W. I. Painter, published in 1931 in *Journal of Educational Research*, Vol. 24, pp. 127-134.

“The Reading Ability of Parents and Factors Associated with Reading Difficulty of Parent Education Materials” by Ralph J. Ojemann, published in 1934 in *University of Iowa Studies in Child Welfare*, Vol. 8, pp. 11-32 and 251-272.

“A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability” by Edgar Dale and Ralph W. Tyler, published in 1934 in *The Library Quarterly*, Vol. 4, pp. 384-412.

“Predicting Readability” by Irving Lorge, published March, 1944, published in *Teachers College Record*, Vol. 45, pp. 404-419.

“A Formula for Predicting Readability” by Edgar Dale and Jeanne S. Chall, published January 21, 1948, in *Educational Research Bulletin*, Vol. 27, No.1, pp. 11-20, 28.

“A Formula for Predicting Readability: Instructions” by Edgar Dale and Jeanne S. Chall, published February 28, 1948, in *Educational Research Bulletin*, Vol. 27, No. 2, pp. 37-54.

“A New Readability Yardstick” by Rudolf Flesh, published June, 1948, in *Journal of Applied Psychology*, Vol. 32. No. 3, pp. 221-233.

Every effort has been made to trace and contact copyright holders. The author will be pleased to correct any mistakes or omissions in future editions.

Editorial commentary © 2007 William H. DuBay. All Rights Reserved.

ISBN: 1-4196-6176-0

Readers' Comments

Please send all comments and suggestions regarding this document to:

William DuBay
Impact Information
15903 Vincent Rd. NW
Poulsbo, WA 98370 Website: <http://www.impact-information.com>

With thanks to the
dedicated friends and workers of the
Orange County Public Library

Preface

For three hundred years after the Battle of Hastings in 1066, English was mainly spoken by servants. Only in the 14th century did it start creeping into halls of school, business, law, and government.

Since then, English advanced to become the language of the greatest body of poetry ever written. A great variety of prose styles also flourished, especially during Elizabethan period. While the governing classes used the florid and polished styles, merchants, artists, farmers, and sea captains developed a straight-forward, plain style of their own.

In the 19th century, educators discovered the simpler style was better for teaching students to read. They began separating students and textbooks into different grades. Early in the 20th century, they discovered many adults had limited reading ability. They began looking for scientific methods for matching texts with readers. Some of the best minds in education dedicated themselves to this task, including Edward L. Thorndike, William S. Gray, Ralph Tyler, Edgar Dale, Irving Lorge, and Jeanne S. Chall.

That story is briefly covered in a companion book, *Smart Language: Readers, Reading, and the Grading of Text*. The purpose of this book is to bring students of reading into contact with this introductory sample of the original articles, methods, and thinking of these educators. In all of them, we see the urgency and pragmatism of the times. I hope that reading them in context will highlight their special place in the story of our remarkable language.

William H. DuBay

Costa Mesa, 10 March 2007.

Contents

1893—L.A. Sherman: The Analytics of Literature	1
1921—Harry D. Kitson: The Mind of the Buyer	4
1921—E. L. Thorndike and the Familiar Word	6
Introduction	6
Word Knowledge in the Elementary School	9
1923—The Lively and Pressey Measuring Method	44
Introduction	44
A Method for Measuring the "Vocabulary Burden" of Textbooks	46
1928—Vogel and Washburne: The Winnetka Formula	56
Introduction	56
An Objective Method of Determining Grade Placement of Children's Reading Material	58
1931—Waples and Tyler: What Adults Want to Read	67
1931—Patty and Painter: The Vocabulary Burden of Textbooks	68
Introduction	68
A Technique for Measuring the Vocabulary Burden of Textbooks	69
1934—Dale and Tyler: Difficulty of Adult Reading Materials	78
Introduction	78
A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability	80
1934—Ralph Ojemann: The Difficulty of Adult Materials	108
Introduction	108
The Reading Ability of Parents and Factors Associated with Reading Difficulty of Parent Education Materials	109
Appendix 1	131
1935—Gray and Leary: What Makes a Book Readable	162
1944—The Lorge Readability Index	166
Introduction	166
Predicting Readability	168
1948—The Dale-Chall Readability Formula	188
Introduction	188
A Formula for Predicting Readability	190
A Formula for Predicting Readability: Instructions	201
1948—Rudolf Flesch and Reading Ease	222
Introduction	222
A New Readability Yardstick	226

1893—L.A. Sherman: The Analytics of Literature

DOWN through the centuries, many had written about the differences between an “ornate” and “plain” style in English.

In 1880, a professor of English Literature at the University of Nebraska, Lucius Adelno Sherman, began to teach literature from a historical and statistical point of view. He gave the world a new method of literary criticism.

Sherman, like other teachers of his time, saw literature as a method for the moral and spiritual edification of citizens. He was one of the first to recruit science for this task and to advocate an “objective” approach to literature. In 1893, he published *Analytics of Literature: A Manual for the Objective Study of English Prose and Poetry*.

He found his approach was highly effective in giving his students an appreciation for great literature:

Students apparently without taste for reading, or capacity to discern common literary excellencies, were enabled to appreciate and enjoy poetry as well as the best. Bright scholars were also in their way benefited not less than the undiscerning. Things vague were made definite. Grounds of judgment before indeterminate or hidden were made plain. Criticism was rendered confident; and no little enthusiasm was aroused (p. xi).

In general, this method, if tried intelligently and fairly, will discover to those who suppose they have no taste for the best literature that they have such taste; and it will make those who have never found anything in poetry both feel and know something of its power (p. xii).

In defending his method against critics, he wrote:

There is a very natural antipathy to treating aesthetics by scientific methods. Yet there is in the nature of things no reason why we may not as well analyze the tissues of human speech and thought as the tissues of the human body. Within a generation science has been broadened by the use of imagination, and there is no good reason why aesthetics in turn should not have the material aid of facts and statistics (p. xiii).

The proof, he claimed, is in the results. His method immediately engages students with the text on a practical level they can understand. A

more refined appreciation of the content grows out of the familiarity with the form and structure of the text.

Sherman's work makes modern use of statistics, charts, and graphs. Most notable are his findings about the streamlining of language.

In comparing the older prose writers with the then-current writers such as Macaulay and Ralph Waldo Emerson, Sherman noticed a progressive shortening of sentences over time.

He decided to look at this statistically and began by counting average sentence length per 100 periods. In his book he showed how sentence-length averages shortened over time:

- Pre-Elizabethan times: 50 words per sentence
- Elizabethan times: 45 words per sentence
- Victorian times: 29 words per sentence
- Sherman's time: 23 words per sentence.

In our time, the average is down to 20 words per sentence.

Sherman's work set the agenda for a century of research in reading. It proposed the following:

- Literature is a subject for statistical analysis.
- Shorter sentences and concrete terms increase readability.
- Spoken language is more efficient than written language.
- Over time, written language becomes more efficient by becoming more like spoken language.

Sherman also showed how individual writers are remarkably consistent in their **average** sentence lengths. This consistency was to become the basis for the validity of using samples of a text rather than the whole thing for readability prediction.

Another of Sherman's discoveries was that over time sentences not only became shorter but also simpler and less abstract. He believed this process was due to the influence of the spoken language on written English. He wrote:

Literary English, in short, will follow the forms of the standard spoken English from which it comes. No man should talk worse than he writes, no man

writes better than he should talk.... The oral sentence is clearest because it is the product of millions of daily efforts to be clear and strong. It represents the work of the race for thousands of years in perfecting an effective instrument of communication (p. 312).

Linguistic research later confirmed Sherman's view of the relationship between spoken and written language. Sherman's most important point was the need to involve the reader. He wrote:

The universally best style is not a thing of form merely, but must regard the expectations of the reader as to the spirit and occasion of what is written. It is not addressed to the learned, but to all minds. Avoiding book-words, it will use only the standard terms and expressions of common life... It will not run in long and involved sentences that cannot readily be understood. Correct in all respects, it will not be stiff; familiar, but safely beyond all associations of vulgarity (p. 327).

—WHD

1921—Harry D. Kitson: The Mind of the Buyer

KN 1921, Harry D. Kitson, a Professor of Psychology at Indiana University, published *The Mind of the Buyer: A Psychology of Selling*. Kitson applied to selling the “New Psychology” promoted by William James. He used James’ “stream of thought” to describe the succeeding stages of the mind of the buyer in a sale: attention, interest, desire, confidence, decision and action, and satisfaction.

Kitson also focused on the part that intuition, imagination, reason, fashion, play in making a sell. He made “the psychological moment” part of the language of commerce.

In his chapter on “Selling to the Collective Buyer,” he lamented the lack of psychological “yardsticks” for measuring social characteristics mathematically. By way of example of what can be done, he turned to print advertising.

He showed how and why readers of different magazines and newspapers differed from one another. Although he may not have been aware of Sherman’s work, he found that sentence length and word length measured in syllables are important measures of readability.

Rudolph Flesch would incorporate both these variables in his Reading Ease formula 30 years later.

Although Kitson did not create a readability formula, he showed how his principles worked in analyzing two newspapers, the Chicago Evening Post and the Chicago American and two magazines, the Century and the American.

He analyzed 5000 consecutive words and 8000 consecutive sentences in the four publications. His study showed that the average word and sentence length were shorter in the *Chicago American* newspaper than in the *Post*, and the *American* magazine’s style simpler than the *Century*’s, accounting for the differences in their readership.

Kitson pointed out that the difficulty of style is not the only difference found in the publications. Readers also have different tastes and interests. He showed how the two above magazines also differed in the amount of space they gave to different kinds of articles. A magazine will reflect its

readers' interests. The advertising should reflect those interests and the editorial tone of the publication.

With Sherman's and Kitson's work both focusing on adult reading, one could have assumed assume that the first readability formulas would have been created for adult materials. One reason that they did not was the appearance in 1921 of Edward L. Thorndike's *The Teachers Word Book*. With that event, the energies of research turned toward young readers.

—WHD

1921—E. L. Thorndike and the Familiar Word

Introduction

EDWARD L. THORNDIKE is recognized as one of the most important psychologists of the 20th century. A student of psychologist and philosopher William James in Harvard, he brought the new approach of Applied Psychology to the classroom. Along with American philosopher John Dewey, he was to dominate educational theory and practice around the world for many years.

During the 1920s, two major trends stimulated a new interest in readability:

1. A changing school population, especially an increase in “first generation” secondary school students, the children of immigrants. Teachers reported that these students found textbooks too difficult.
2. The growing use of scientific tools for studying and objectively measuring educational problems.

One such tool, Thorndike’s *The Teacher’s Word Book*, which came out in 1921, was the first extensive listing of words in English by frequency.

It is hard to overestimate the effects of this book on education and communications. It provided teachers with an objective means for measuring the difficulty of words and texts. It laid the foundation for almost all the research on readability that would follow, including the first readability formulas.

Thorndike, while doing research for the Teachers College in Columbia University, noticed that teachers of languages in Germany and Russia were using word counts to match texts with students. The more frequent a word is used, they found, the more familiar it is and the easier to use.

Around 1911, Thorndike began to count the frequency of words in English texts. In 1921, he published *The Teacher’s Word Book*, which listed 10,000 words by frequency of use. Each word was given a score or “credit”, depending on how frequently it is used. A word like *the* or *to* has a credit of 208, while *advisable* has a credit of 3. In 1932, he followed up with *A Teacher’s Word Book of 20,000 Words*, and, in 1944 with Irving

Lorge, *A Teacher's Word Book of 30,000 Words*.

A vocabulary test on the meaning of words is the strongest predictor of verbal and abstract intellectual development. The knowledge of words has always been a strong measure of a reader's development, reading comprehension, and verbal intelligence. Chall and Dale wrote in 1995, "It is no accident that vocabulary is also a strong predictor of text difficulty."

It happens that the first words we learn are the simplest and shortest. These first, easy words are also the words we use most frequently. Most people do not realize the extent of this frequency. Twenty-five percent of the 67,200 words used in the 24 life stories written by university freshmen consisted of these ten words: *the, I, and, to, was, my, in, of, a, and it*. The first 100 most frequent words make up almost half of all written material. The first 300 words make up about 65 percent of it.

Educators, publishers, and teachers still use Thorndike's word-frequency lists to evaluate reading materials for schools. Extensive research on vocabulary continues to this day.

In 1968, psychologist George Klare wrote about the central role vocabulary plays in reading skill, "Not only do humans tend to use some words much more often than others, they recognize more frequent words more rapidly than less frequent, prefer them, and understand and learn them more readily. It is not surprising, therefore, that this variable has such a central role in the measurement of readability."

The following article, reprinted here in full, accompanied the publication of the first *The Teachers Word Book* in 1921. In the introduction to the book, Thorndike refers readers to this article for the background of the list and how to use it:

A full account of the methods by which this list was selected, of the reliability of the credits attached to the words, and of the uses to which the list may be put, will be found in an article of thirty-seven pages, entitled "Word Knowledge in the Elementary School," by E. L. Thorndike, published in the *Teachers College Record* for September, 1921¹.

The article also shows Thorndike's beliefs about teaching and how students learn to read. He shows how to use the list to teach high school students and those learning English as a second language.

¹ Vol. 22, No. 5, pp.334-370.

His comments on why youngsters prefer to read “trash” rather than current affairs refers to the lack of “suitable reading materials” for readers of all ages and backgrounds.

In the mid-1930’s, critics such as Dale and Tyler (see the study that follows) claimed that Thorndike’s lists did not take into account the different meanings that some words have.

Edgar Dale began working on his own word lists. In 1944, Irving Lorge published a formula using Dale’s “short list.” In 1948, along with Jeanne Chall, Dale published a readability formula using his “long list” of 3,000 easy words.

In 1981, Dale and Joseph O’Rourke published their 40,000-word *Living Word Vocabulary*, which gave not the frequency of words but their grade levels. It was a masterful tribute to Thorndike’s original work.

—WHD

Teachers College Record

Vol. XXII

November, 1921

No. 5

Word Knowledge in the Elementary School

By EDWARD L. THORNDIKE
Professor of Education, Teachers College

CONSIDER these simple questions: How many English words should the ordinary boy or girl know the meanings of at the end of Grade 8? Which words should all or nearly all pupils know at that stage? In what grades and in what connections should they be learned? If gifted and experienced teachers, supervisors, and authors of courses of study were to give answers to these questions, the answers would vary enormously. Nobody, in fact, knows the answers with even roughly approximate correctness. Nor is our condition better if we free the questions from the ambiguity of "ordinary boy or girl" and specify any particular child or type of child. The answers would still vary widely, and all of them might well be wrong. Nor is our condition better if we describe fully what else the pupil is to know and assign, say, 8.375 per cent of his time and energy to this particular feature of his education. We still cannot answer with any surety.

These questions, though concerned with details, and less inspiring than broad questions about health, morality, or citizenship, are important, as indeed all competent workers in the science of education will now admit.

It appears that one notable cause of our inability to answer them correctly is our lack of knowledge of the frequency of occurrence of words in the talk our pupil and graduate will or should hear, and the books, articles, letters, and the like, which he will or should read. Just as word counts of such material as the pupil may need to write are instructive in the pedagogy of spelling, so word counts of such material as the pupil may need to understand will be instructive in the pedagogy of reading, and indeed of all the school subjects which are presented with the aid of language.

So for about ten years I have made such counts as I could. They are as follows:

THE WORD COUNTS AND CREDITS

Children's Reading

1. Every word in Chapters 10 to .19 (and a few pages more) of *Black Beauty*, one of the most popular books if not the most popular book for children about nine years of age. About 11,500 words in all. Credits given as follows: 1 to 4 occurrences, 1, 5 to 9, 2; 10 to 19, 3; 20 to 29, 4; 30 to 49, 5; 50 or over, 6. That is, if a word occurred once, twice, three times, or four times in these chapters of *Black Beauty*, it was given a credit of 1; if it occurred from 5 to 9 times, it was given a credit of 2, etc.¹
2. Every word in Chapters 1, 2 and 3 (and 14 lines of Chapter 4) of *Little Women*. About 13,000 words in all. Credits: As in No. 1.
3. Every word in Chapters 1 to 5 and part of 6 of *Treasure Island*. About 13,000 words in all. Credits: As in No. 1.
4. Every word in *Scrooge's Christmas*, a selection from *The Christmas Carol* as reprinted in a school reader. About 8,000 words in all. Credits: As in No. 1.
5. Every word in Irving's *Sleepy Hollow*. About 13,000 words in all. Credits: As in No. 1.
6. Every word in one issue of the *Youth's Companion*, omitting advertisements and fine print. About 25,000 words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 to 29, 4; 30 or over, 5.
7. Every word in the fifty-six selections found by Hosisic to be the commonest features of school readers.² About 27,000 words in all.

¹ The reader is asked to accept arbitrarily these credits since an explanation of the method by which they were obtained is too involved to be given here.

² It would probably be more scientific to enter these along with items 22 to 30 as Standard Literature; and, in another connection, this is done. The list is as follows:

Allingham, The Fairies.	Kingsley, The Lost Doll.
Anderson, The Steadfast Tin Soldier.	Kipling, Toomai of the Elephants.
Anderson, The Ugly Duckling.	Lincoln, Address at Gettysburg.
Bjornson, The Tree.	Longfellow, The Arrow and the Song.
Browning, Home Thoughts from Abroad	Longfellow, The Village Blacksmith.

These were divided into two halves, referred to hereafter as Hosic A and Hosic B. Credits for Hosic A were; 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 to 29, 4; 30 or over, 5. Credits for Hosic B were the same.

8. Every word in ten primers or first readers (for this material in its original form, I am indebted to Mr. C. N. Smith and the teachers who aided him). As a very rough estimate, we may take 80,000 words. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 to 29, 4; 30 to 39, 5; 40 to 49, 6; 50 or over, or fewer if present in all 10 books, 7.
 9. Every word in ten second readers (for this material in its original form I am indebted to the University of Iowa and E. T. Hooch). About 150,000 words in all. Credits: Same as for No. 8.
 10. Every word in ten third readers (for the material in its original form I am indebted to the University of Iowa and W. S. Miller). 283,000
-

Browning, How They Brought the Good News from Ghent to Aix	Macaulay, Horatius
Browning, Pied Piper of Hamen	Miller, Columbus.
Bryant, Planting of the Apple Tree.	Moore, A Visit From St. Nicholas.
Bryant, Robert of Lincoln.	Saxe, The Blind Man and the Elephant.
Bryant, Song of Marion's Men	Scott, Loshinvar
Bryant, To a Waterfowl.	Scott, Love of Country (Lives there a man)
Burns, A Man's a Man.	Shakespeare, Orpheus with His Lute.
Carlyle, Today.	Shakespeare, Under the Greenwood Tree.
Drake, The American Flag	Sherman, Daisies.
Emerson, Concord Hymn	Smith, America.
Emerson, The Mountain and the Squirrel	Stevensen, My Shadow.
Franklin, Proverbs.	Stevensen, The Wind.
Franklin, The Whistle.	Tennyson, The Brook.
Gray, Elegy.	Tennyson, The Bugle Song.
Hemans, Landing of the Pilgrims.	Tennyson, The Charge of the Light Brigade
Holmes, The Chambered Nautilus.	Tennyson, Sir Galahad.
Holmes, The Deacon's Masterpiece.	Thaxter, The Sandpiper.
Holmes, Old Ironsides.	Whitman, O Captain, My Captain.
Hunt, Abu ben Adhem.	Whittier, The Barefoot Boy.
Ingelow, Seven Times One.	Wolfe, Burial of Sir John Moore.
Irving, Rip Van Winkle.	Wordsworth, Daffodils (I wandered lonely).
Key, The Star Spangled Banner.	

words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 to 29, 4; 30 to 39, 5; 40 to 74, 6; 75 or over, or fewer if present in all 10 books, 7.

11. Every word in Book One of the *Thorndike Arithmetics*, edition of 1917. About 32,000 words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 or over, 4.
 12. Every word in Book One of the original edition of the *Young and Jackson Arithmetics*. About 35,000 words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 to 29, 4; 30 or over, 5.
 13. Every word in Brigham and McFarlane, *Essentials of Geography, First Book*, first edition, pages 26 to 256. About 83,000 words in all. Credits: As for No. 12.
 14. Every word in Straubenmuller, *A Home Geography of New York City*. About 37,000 words in all. Credits: As for No. 12
 15. Every word in Thwaites and Kendal, *History of the United States*, edition of 1914, pages 26-100 inclusive. About 25,000 words in all. Credits: As for No. 12.
 16. Every word in Forman, *History of the United States*, pages 101, 102, 121, 122, etc., to the end. About 17,000 words in all. , Credits: As for No. 12.
 17. Every word in the first 25 pages of each of five standard First Books in Geography. About 40,000 words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 29, 3; 30 to 49, 4; 50 or over, 5.
 18. Every word in the first 25 pages of each of three text-books in United States History. About 20,000 words in all. Credits: As in No. 12.
 19. The vocabularies of ten books on Elementary French. Credits: A credit of 1 for a word found in 1 or 2 of the ten; 2 for a word found in 3 or 4 of the ten; 3 for a word found in 5 or 6 of the ten; 4 for a word found in 7 or 8 of the ten; 5 for a word found in 9 or 10 of the ten.
 20. The vocabularies of ten books on Elementary German. Credits: As in No. 19.
 21. The vocabularies of five books on Elementary Spanish. Credits: 0 for a word found in only one of the five; 2 for a word found in 2 of the 5; 3 for a word found in 3 of the 5; 4 for a word found in 4 of
-

the 5; and 5 for a word found in all.³

22. Strong, *Exhaustive Concordance of the Bible*. About 900,000 words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 occurrences up to a column-full,⁴ 3; a column-full but not two columns-full, 4; two columns or over, 5.
23. John Bartlett, *Concordance to the Dramatic Works of Shakespeare*. (The supplementary concordance to the poems was not used.) About 925,000 words in all. Credits: 3 to 9 occurrences, 1; 10 occurrences up to a column-full, 2; a column-full but not two columns-full, 3; two columns or more, 4.
24. Lane Cooper, *Concordance to Wordsworth*. About 400,000 words. Credits: As for No. 23.
25. Baker, *Concordance to Tennyson's Poetical and Dramatic Works*, Part II, on the Dramatic Works. About 120,000 words. Credits: As for No. 23.
26. Baker, *Concordance to Tennyson's Poetical and Dramatic Works*, Part I, on the Poetical Works. About 200,000 words. Credits: As for No. 23.
27. Neve, *Concordance to the Poetical Works of William Cowper*. About 200,000 words. Credits: As for No. 23.
28. Abbott, *Concordance to Pope* (this covers only a part of Pope's poetical works). About 90,000 words. Credits: As for No. 23.
29. Bradshaw, *Concordance to the Poetical Works of John Milton*. About 130,000 words. Credits: As for No. 23.
30. Bartlett, *Familiar Quotations*, 10th edition, pages 42-62, 201-220, 301-320, 401-420, 501-520, 601-620, 701-720, 801-820, 901-920 and 1001-1020; footnotes and all save the quotations themselves being omitted. About 32,000 words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 to 29, 4; 30 or over, 5.

Common Facts and Trades

³ I am indebted to Professor Bagster-Collins for the Spanish data and to Mr. Ben Wood for their translation.

⁴ A column means about 116 occurrences in the Bible; about 90 in Shakespeare; about 92 in Wordsworth; about 72 in Tennyson's plays, or in his poems; about 80 in Milton; 82 in Pope; and about 55 in Cowper.

31. The United States Constitution and the Declaration of Independence. About 8,000 words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 or more, 4. The words of Common List A⁵ were not counted but each was given a credit of 4. This same procedure was followed in Nos. 32 to 36 inclusive. Credits for *today*, *too*, *two*, *until* and *us* were also assigned by estimate in Nos. 31 to 36, as: 1, 2, 3, 2 and 1 respectively.
32. Farmer, *A New Book of Cookery*, pages 1, 11, 21, 31, etc. About 4,000 words in all. Credits: As in No. 31.
33. Allington, *Practical Seiving and Dressmaking*, pages 1, 11, 21, 31, etc. About 6,000 words in all. Credits: As in No. 31.
34. *Garden and Farm Almanac for 1914*, pages 7, 9, 11, 18, 19 to 32; 88 to 120; 132, 145 to 151 and 156 (but with a few omissions of notes, statistics, etc.). About 17,000 words in all. Credits: As in No. 31.
35. Five pages containing the United States postal regulations in popular form. About 1,700 words in all. Credits: As in No. 31.

⁵ The words of Common List A were as follows:

a	can	good	know	night	since	very
about	come	got		no	so	
after	could	great		not	some	
again			last	now	soon	
against			let		such	
ago	day	had	like			was
also	did	has	little	of		way
all	do	have	long	on	than	
an	does	he		only	that	
and	down	her		one	the	
any		here	made	or	their	
are		him	make	other	them	
as		his	man	our	then	
at	each	how	man	out	there	
	every		matter	over	these	
			may		they	
be		if	me		through	
been	few	in	might	place	thing	
before	first	into	more		this	
best	for	is	most		time	
better	from	it	much	same	to	
but		its	my	say		
	get			see		you
go		just	new	she	up	your
				should		

36. The first ten questions and answers in each of thirty-one trade tests chosen from those published by the Committee on Classification of Personnel in the Army. About 5,000 words in all. Credits: As in No. 31.
37. The first word of each entry in the indexes of three large mail-order catalogues. The fullest index was taken as a basis. Credits: A credit of 1 was given if the word was in this fullest index only. If it was also in one other, a credit of 2 was given. If it was in all three, a credit of 3 was given.

Newspaper Reading

38. The 44,000 words of the Eldridge count from a Buffalo paper. Eldridge did not include proper names or numerals. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 to 29, 4; 30 to 49, 5; 50 to 99, 6; 100 to 999, 7; 1000 or over, 8. For numerals and proper names, credit same as in No. 39.
39. Selections from the newspaper of an eastern city, *The Examiner-New Era* of Lancaster, Pa., taken from six issues spread over the year 1920 and taken at random from page 1, col. 1; page 1, col. 2; etc. About 40,000 words in all. Credits: 1 to 4 occurrences, 1; 5 to 9, 2; 10 to 19, 3; 20 to 39, 4; 30 or over, 5. The words of Common List A were not counted but given each a credit of 5. *Today, too, two, until, and us* were assigned credits of 1, 2, 3, 2, and 3, respectively.

Correspondence

40. The Cook-O'Shea list of frequencies derived from a count of 200,000 words of private correspondence was used. Credits: 2 to 9 occurrences, 1; 10 to 19, 2; 20 to 39, 3; 40 or over in their List III, 4; 40 to 49 in their Lists I and II, 4; 50 to 99 in Lists I and II, 5; 100 and over in List II, 100 to 299 in List I, 6; 300 to 999 in List I, 7; 1000 to 1999 in List I, 8; 2000 and over in List I, 9.
41. The Anderson list⁶ of frequencies derived from students' counts of over 360,000 words in business and private correspondence was used. He reports the results for only 3087 words, occurring with a total frequency of 5 or more, and occurring in at least three of the

⁶ W. N. Anderson, *The Determination of a Spelling Vocabulary Based upon Written Correspondence*, Aug. 1917. Manuscript in the library of Iowa State University.

six groups into which he divided his material (Professional, Business, Domestic, Miscellaneous, Personal and Farmers'). Credits: 5 to 19, 2; 20 to 49, 3; 50 to 99, 4; 100 to 499, 5; 500 to 999, 6; 1000 or over, 7.

In connection with the counts listed above, certain corrections and additions have been made where the need for them was obvious.

For example: the concordances omit altogether certain very common words like *the*, *and*, *of*, *it*. The credit for these was estimated, and usually with no risk of error, it being well above the maximum credits as stated on page 339 [14]. Sometimes the concordances give samples of a word's usage, but not complete inventories, and it is to be feared that they sometimes do this without informing the student of the fact. These cases have been treated as well as might be without elaborate study. There are doubtless errors on the part of myself and of my assistants in the counting and crediting and tabulating; but such probably act for the most part as variable errors. Where they do not, the most probable resulting error is an insufficient importance for abbreviations, and, to a less extent, for proper names. In the case of the counts from the concordances, to avoid tabulating data of no probable use in the final report, the custom was followed of not entering rare proper names. The material of the two spelling lists, and of the vocabularies of French, German, and Spanish First-Year Books, also does not include all the proper names used. One of the spelling lists includes all except those of "towns of less than 10,000 inhabitants; all strictly local publications; organizations and streets; and all family names, except those of historical characters or of men in the public eye"; the other includes only the names of the days of the week, months, and nationalities. So it may well be that on grounds of frequency alone some of the words of our last thousand or so should be displaced by names of persons and places. However, on grounds of real importance, this "error" may well have brought us nearer the truth.

The concordances, the Eldridge count, the Cook-O'Shea count, and the vocabularies rarely or never include abbreviations. Although Dr. Anderson instructed his helpers to count all abbreviations, it seems almost certain from the actual results that they did not do so. *Etc.*, *Mr.*, and *Mrs.* receive credit from him; but *doz.*, *ft.*, *kr.*, *in.*, *pt.*, *pk.*, *St.*, and the like, do not. I fear, therefore, that his instructions were often disregarded. The sums of credits in the case of abbreviations in my list are thus unreliable and in general too low.

The Cook-O'Shea list does not make separate entry for words like *am*, *are*, *ate*, *been*, *began*, *begun*, *is*, *was*, and the like. I have not corrected for this because it seems undesirable to tamper with lists selected by others, even if their procedure seems indefensible. The vocabularies of the first-year foreign language books also are taken as they stand, although they omit many derived words which really are used in the body of the text. The general effect of the omissions in them and in the Cook-O'Shea list is to reduce the credit of derived forms. This is perhaps desirable.

I shall not defend the above as an especially good selection of material, though it is by no means a bad one. Some of the items were chosen partly for other reasons than the rating of words for frequency and range of occurrence. Some were chosen, notably the concordances, because of the amount of information gained per hour or dollar spent. Nor shall I defend the system of credits used above as an especially wise system of weighting frequency and range of occurrence, on the whole or within one sort of matter, such as children's literature. Indeed I am sure that I could now improve it. The general principle of weighting range as well as mere number of occurrences is sound, and the final result from the cleverest weighting would probably not be very much better than that secured here. Let us postpone further critical study of the counts until we have inspected some of the results.

GENERAL RESULTS

First, it should be noted that a plural formed by adding *s* was not counted separately, but entered under the singular form. The same procedure was followed, except in certain cases for special reasons, with plurals where *y* is replaced by *ies*, adverbs formed from adjectives by adding *ly*, comparatives formed by adding *er* or *r*, superlatives formed by adding *est* or *st*, verb forms derived by adding *s*, *ed* or *d*,⁷ *n*, and *ing*, in cases where the derived form would probably be easily read and understood by the pupil when he experienced it, if he knew the primary word. Adjectives formed from proper nouns by adding *n* are also, as a rule, counted with the noun. For example:

days	nights	cherries	counted under	day	night	cherry
gladly	proudly	counted under	glad	proud		
stronger	weakest	counted under	strong	weak		
shows	showed	shown	showing	counted under	show	
Russian	Bolivian	counted under	Russia	Bolivia		

⁷ Including changes of *y* to *ies*, *ier*, *iest* and *ied*.

This greatly reduces the number of entries and seems desirable for our purpose. If a count is to be used as a guide to instruction in spelling, on the contrary, we need estimates of each such derived form.

It was permissible to omit from entry rare names of persons and places. Apart from this, every word or abbreviation was to be counted and was counted in most of my work. In some that was done without my direct supervision abbreviations were, I fear, somewhat neglected. In the concordances and vocabularies of text-books in French, etc., and in the counts by Eldridge, Cook and O'Shea, Miller, Housh and Anderson, abbreviations have been very largely neglected.

In all, over 20,000 words or abbreviations received a credit of 1 or more.

It should be made clear at this point that the credit assigned to a given word in each of the several counts depended upon the number of times the given word occurred in the particular book or article. It follows, therefore, that a given word might have a credit of 4 in the count of one book and a credit of 5 or more in the count of a second book. As an illustration in the count of *Black Beauty* the word *angel* has a credit of 1, in the Bible this same word has a credit of 5, while in Tennyson's plays it is credited as 2. It is evident that the credit received by the word *angel* in *Black Beauty* would not be a sufficient indication of its importance or occurrence in all the books counted.

We thus have for each of the 20,000 words which received a credit of 1 or a record like that shown in Table I for the words *and*, *angel*, *anger*, *angle*, *anguish*, and *animal*. This record is summarized in a number, the "sum of credits," found at the bottom of each column. Thus the word *and* has a "credit sum" of 210; the word *angel*, 40; *anger*, 36; *angle*, 15; *anguish*, 11; *animal*, 70.

The highest credit sum found among the 20,000 words, which belongs to the word *in*, is 211; the word which is 500th in rank has a credit sum of 75; the 1000th word has 49 as a credit sum; the 1500th word has 36 as a credit sum; the 2000th word has 28 as a credit sum; the 2500th has 23 as a credit sum. The credit sums of the next 2500 words (to the 5000th) range between 23 and 10; those of the next 5000 range between 9 and 3; and there are about 10,000 words with credit sums of only 1 or 2. That is, the distribution of words in respect to frequency and range of occurrence is as shown in Table II.

TABLE I

	<i>and</i>	<i>angel</i>	<i>anger</i>	<i>angle</i>	<i>anguish</i>	<i>animal</i>
Bible	5	5	5	1	3	
Wordsworth	4	3	2		2	2
Tennyson, plays	4	2	2			
Cowper	4	2				I
Familiar Quotations	5	4	1	1		1
Hosic, A	5	1				1
Thwaites-Kendal History, 75 pp.	5		1			2
Brigham-McFarlane Geography, 230 pp.	5			1		5
Thorndike Arithmetics	5			2		
Geography, beginning, 125 pp.	5					4
First Readers or Primers, 10	7					1
Second Readers, 10	7		3			6
Third Readers, 10	7	4	2			7
Shakespeare	4	2	2	2	1	1
Milton	4	3	2		2	1
Tennyson, poems	4	2	2	1	1	
Pope	4	2	1			
Youth's Companion	5		1			3
Hosic, B.	5	1				
Forman History, 70 pp.	5					1
Straubenmuller Geography, 218 pp.	5	1	1	1		3
Young-Jackson Arithmetic	5					
History, beginning, 75pp.	5	1				1
French 1st-year books, 10	5	2	2		1	4
German 1st-year books, 10	5		3			5
Spanish 1st-year books, 5	5		2			4
Cook-O'Shea Correspondence	7	1				2
Anderson Correspondence	7					3
Cook book	4	2				1
Sewing book	4					
Trade tests	4			1		
Farm almanac	4			2		4
Mail-order catalogues	10			2		2
Postal regulations	4					
Constitution and Declaration	4					
Eldridge newspaper	8		1			1
Examiner newspaper	5		1	1		2
Black Beauty	6	1	1			1
Little Women	6	1			1	
Scrooge	6					
Treasure Island	6		1			
Sleepy Hollow	6					1
SUM OF CREDITS	210	40	36	15	11	70

It became necessary therefore to devise some means of designating the importance of a word by adding together the credits received by that word in all the counts. The result thus obtained for each word is called the "sum of credits" or "credit sum."

TABLE II

DISTRIBUTION OF WORDS IN RESPECT TO FREQUENCY
AND RANGE OF OCCURRENCE

Sum of Credit of:

0	50,000 (approx)
1 and 2	10,000 (approx)
3 and 4	2,878
5 and 6	1,569
7 and 8	1,074
9 and 10	759
11 and 12	565
13 and 14	478
15 and 16	413
17 and 18	291
19 and 20	248
21 and 22	256
23 and 24	189
25 and 26	198
27 and 28	181
29 and 30	152
31 and 32	100
33 and 34	128
35 and 36	103
37 and 38	96
39 and 40	96
41 and 42	74
43 and 44	69
45 and 46	65
47 and 48	66
49 and 50	60
51 and 52	49
53 and 54	40
55 and 56	50
57 and 58	59
59 and 60	47

with 715 cases spread from 61 to 211 with decreasing frequency.

The meaning of these credit sums can be realized by a rapid inspection

of the list below which gives four or five words taken at random from those receiving the following credit sums: 200, 180, 160, 140, 120, 100, 90, 80, 70, 60, 50, 40, 30, 20, 10, 8, 6, and 4.

SUM OF CREDITS	SAMPLES
200	for it on one
180	an no or some were
160	go great her new well
140	did just live say
120	eye light right white woman
100	mean nothing order sea seem
90	almost ground mark rich table
80	buy dead hat known lead
70	able act animal appear cannot
60	class cup distance doubt equal
50	bit breakfast clothing (e, ed) crowd crown
40	advantage ah angel aunt bend
30	advice avenue ax (axe) beach bean
20	accordingly acquire ahead ample anyone
10	abate abolish abound accommodate acorn
8	absolve absorb absurd accommodation accordance
6	abstain adjacent adjective admirable admirer
4	abashed abridge accessory Adams adjudge

In the above list the exact credit sums of some of the words are not multiples of ten. These are: *for*, 201; *it* and *one*, 199; *an, no, or, and some*, 181; *great*, 159; *her*, 161; *live and water*, 139.

The general nature of the list may be realized from another point of view by an inspection of the following words which mark the ends of the first, second, third, fourth, and fifth thousands:

End of 1st 1000. Sum of credits, 49

chain circle condition date discover double escape fancy fence fool grave health history jump mail merry mighty noise

End of 2nd 1000. Sum of credits, 28

addition Africa Asia authority bosom brass bride broke brow button complain congress cottage create crime crush curl deceive defend devil Dick doll eagle echo eleven farewell

End of 3rd 1000. Sum of credits, 19

Alice amaze argument Arthur banish barren bat belief beloved bloody bower breadth Britain Caesar capable cedar cloudy combination com-

mend confusion conquest consequence construct crept dame

End of 4th 1000, Sum of credits, 13

accursed actor Adam affair afflict airy Alexander almighty alms ambassador amid applause arrangement attorney auto awoke bait balm beguile Benjamin Bible boundless bounty brace Carl

End of 5th 1000. Sum of credits, 10

abate abolish abound accommodate acorn admission adoption adversary affectionate . agency aisle ale allegiance aloft anniversary annoy anoint antiquity anvil apparel aspect asunder Athens attic attitude

THE RESULTS IN DETAIL

The most frequently occurring 10,000 words of our list are printed in alphabetical order with a rating of the importance of each as indicated by its sum of credits, and in a form convenient for use by teachers and others in *The Teacher's Word Book*⁸. When the term "list" is used hereafter in this article, it will mean this list of 10,000 words.

The inclusion of all the words obtaining 3 as a sum of credits would extend the list to nearly 11,000. So I have eliminated those "3" words which seemed in the combined opinion of five judges to be the least important.⁹ I have also eliminated some words¹⁰ receiving credits so exclusively from one source that it seemed absolutely certain that a more extensive count would not include them. With this exception the list represents the unmodified results of the counts.

⁸ This is a book of 132 pages so arranged that relevant facts about any of these words can be entered. In the case of the 5000 most important words, the credit sum is followed by a number and letter stating in which thousand and in which half thereof the word belongs. Thus "43 2a" means that the word has a credit sum of 43 and is in the first half of the second thousand, "21 3b" means that the word has an importance of 21 and is in the second half of the third thousand. Within the first five hundred there is a further distinction into hundreds, 1a1 meaning that the word is one of the first hundred; 1a2 meaning that it is in the second hundred; 133 meaning that it is in the third hundred, and so on. The book is obtainable from the Bureau of Publications, Teachers College, 525 West 120 Street, New York City.

⁹ As a consequence the "3" words remaining in the list are probably on the whole nearly or quite as important as the "4" words.

¹⁰ Such were, for example: Agrippa, Albion, Arabella, Castoria, chee, chiffonette, contemn, Coronado, Cyprus, Dagon, drave, Eli, Enoch, Ephriam, Galatians, Gardinar, Hepsy. Jephthah, Jip, linnet, ope, pate, pied, Titus, trow, Tubal, tweet.

THE ADEQUACY OF THE COUNTS

Before drawing conclusions from the facts of the list and planning uses for it, we must deal with certain questions about the quantitative and qualitative adequacy of the counts.

The question of quantitative adequacy may best be stated thus. What are the probable divergences of the present list from one that would be made up from counts of the same kind of selections in the same proportions, but hundreds of times as extensive? The full answer is given in Table III. The gist of it is that the present counts are adequate to determine the first one thousand words with a small probable error, and the next four thousand well enough for many educational purposes, and the last five thousand to an extent that is useful, though far from accurate.

Of the words put in the first 500 of our list, about 25 words would, by an infinitely extensive count, be put lower. Of the words in the second 500, about 31 would be put in the next higher 500, and about 57 in the next lower 500. Of the words in the top 5000, about 350 would be put lower. Of the words in the entire 10,000, about 1000 would be displaced downward, and replaced by others from the 60,000 not listed here, but almost all of them would come in the next few thousands, so that it is not a serious error to regard them as belonging within the 10,000.

There are 957 chances in 1000 that a word of credit above 100 would not by the infinite count receive a credit 13 higher or lower than it does; that a word of credit 50 would not receive a credit 11 higher or lower than it does; that a word of credit 25 would not receive a credit 8 higher or lower than it does; that a word of credit 20 would not receive a credit 7½ more or less than it does; and that a word of credit 15 would not receive a credit 6½ more or less than it does.

It is impossible to measure the qualitative adequacy of the counts and credits except by further very extensive counts. For example, only further counts of children's reading can decide whether our selections are an adequate sampling of the matter children do or should read. Only elaborate counts of correspondence will decide whether the Cook-O'Shea and Anderson counts are fairly representative of correspondence in general as to its quality. Moreover, there will always be some room for diversity of opinion as to the balance between what is read and what we should consider important to be read. As to the relative weight to attach to children's

TABLE III

THE EXTENT TO WHICH THE RESULTS OF AN INFINITELY EXTENSIVE COUNT OF THE SAME SORT MAY BE EXPECTED TO DIVERGE FROM THE RESULTS OBTAINED FROM THE PRESENT COUNT.

			Sum of Credits	P.E. of sum of Credits	Median probable displacement from true position in terms of number of words		Sum of Credits	P.E. of Sum of Credits	Median probable displacement		Sum of Credits	P.E. of Sum of Credits	Median probable displacement
4	1.50	1900	54	3.70	90		100	4.20	25				
6	1.60	1150	6	3.74	86		105	4.20	23				
8	1.75	850	8	3.78	83		110	4.20	20				
10	1.87	675	60	3.81	80		115	4.20	17				
2	2.00	550	2	3.84	76		120	4.20	15				
4	2.11	465	4	3.86	72		125	4.20	13				
6	2.22	405	6	3.88	69		130	4.20	12				
8	2.33	360	8	3.90	65		135	4.20	8				
20	2.43	325	70	3.92	62		140	4.20	10				
2	2.52	300	2	3.94	58		145	4.20	10				
4	2.60	275	4	3.96	55		150	4.20	9				
6	2.68	250	6	3.98	51		155	4.20	9				
8	2.76	225	8	4.00	48		160	4.20	8				
30	2.84	202	80	4.02	45		165	4.20	8				
2	2.92	185	2	4.04	42		170	4.20	7				
4	3.00	170	4	4.06	40		175	4.20	7				
6	3.08	157	6	4.08	38		180	4.20	6				
8	3.16	145	8	4.10	36		185	4.20	6				
40	3.24	135	90	4.12	34		190	4.20	6				
2	3.32	126	2	4.14	34		195	4.20	5				
4	3.40	118	4	4.16	30		200	4.20	5				
6	3.47	111	6	4.18	29		205	4.20	5				
8	3.54	105	8	4.20	27		210	4.20	5				
50	3.60	99											
2	3.65	94											

reading, correspondence, literary classics, the Bible, the newspaper, and so on, the diversity of opinion exists and will long remain. I hope some time to be able to publish all the original counts so that any competent person can use them with any weights that he thinks desirable.

In various details notable changes might be made. The importance of-

hath and *doth*, for example, depends almost entirely on the weight given to the reading of the Bible. If a revised version of it used *has* and *does*, they would drop down or out of our list. The importance of *acid*, *ampère*, *atom*, and the like, depends on the weight attached to reading on scientific and industrial topics as contrasted with "literary" reading.

The scientific student of the list can use the facts given about the counts and credits to amend it where he thinks wise. So it seems undesirable here to deal at any length with the principles of selection and weighting. I shall therefore simply note four principles, and illustrate in representative cases how the selection and weighting work out. Pains were taken to preserve some fair balance between importance for the boy and girl during the elementary school course, and importance for them after graduation. Rather large weight was given to appearance in a variety of sorts of reading, and relatively small weight to a large number of occurrences in any one sort alone. A balance was sought between *reliability* (attainable by a large count, as in the concordances) and *significance*. A word appearing in a small count is, other things being equal, more important, of course, than one appearing in a large count.

The practical principle of obtaining the best result per hour of time spent in the counts was considered throughout, and explains the use of counts of vocabularies in foreign-language books, and the use of some of the concordances. It is not assumed, for example, that many elementary-school graduates will read Wordsworth, or Milton, or Pope, or Cowper to any considerable extent. The concordance counts, however, are made rapidly and with a moderate weight probably benefit the total result more than would the same amount of time spent on textual counts.

To turn to some illustrative cases, there are given in Table IV the sources of the credits for five words earning 100, 90, 80, 70, 60, and 50, 40, 30, 20, and 10 respectively. The words of credit above 100 would be at or near the top by any sort of count whatever.

In the next table (Table V) are given the rank order for these fifty words by the total sum of credits, and by children's reading alone, elementary text-books alone, standard literature alone, and so on.

It should be noted that I have included neither the counts from Hosis's 57 commonest selections in school readers nor *Sleepy Hollow* under Children's Literature here, but under English Classics.

Tables IV and V show that there is a correspondence between the sums of credits for the same word from different sorts of material, but that it is

far from perfect. The coefficients of correlation (by the Spearman foot-rule) of Total with Children's Literature, English Classics, etc., are, in order, about .9, .8, .9, .5, .8, and .7.¹¹ The high values of the first three (.9, .8, .9) are in part due to the fact that they predominate in determining the Total. Common Facts and Trades has the lowest correlation partly because it has the least share in determining the total.

After this effect of the composition of the Total is allowed for, it still remains true that the vocabularies of Common Facts and Correspondence seem notably specialized in the table. That this is true in general seems certain to one who has made the counts. Indeed the Common Facts and Trades counts were selected to supply an obvious lack.

The greatest displacement for the Children's Literature is of *equal*, 18½ ranks too low ($41\frac{1}{2} - 23$); for the Bible and Classics, it is *angel*, 27 ranks too high; for the Vocabularies and Text-Books it is *class*, 19½ ranks too high; for Common Facts and Trades it is *sea*, 42½ ranks too low; for Newspapers it is *crowd*, 19½ too high; and for Correspondence it is *ah*, 28½ too low. These displacements are all such as might be expected even in an infinitely extensive count. To obtain a measure of general importance for the elementary school pupil and graduate, we have to assign weights. If we weight Common Facts more heavily, *angels* will go down and *bricks* will go up!

The sums of credits at the basis of our list down to 10, the end of the fifth thousand, are determined to over three fourths of their amount by the counts from children's reading, classics, text-books and vocabularies. The result of this weighting corresponds probably rather closely with importance as measured by the prevailing ideals of what an elementary school pupil and graduate *should* read. These ideals are, however, themselves perhaps somewhat inappreciative of science, technology, business, and politics in comparison with literature and morals. They also perhaps undervalue the present and future in comparison with the past. There is further an increasing trend toward considering what the pupils will read as well as what they should read. The present list may therefore be criticized

¹¹ The absolute magnitudes of these coefficients lack their ordinary meanings, since the words are taken at intervals of 10 in respect to the credit sum. The relative magnitudes are all that I mean to measure by them.

TABLE IV

CREDITS RECEIVED FROM DIFFERENT SOURCES BY EACH OF FIFTY
REPRESENTATIVE WORDS.

	Children's Literature	Bible and English Classics	Textbooks and Vocabularies	Common Facts and Trades	Newspapers	Correspondence	Total Sum of Credits
mean	22	25	31	4	7	11	100
nothing	28	29	24	2	7	10	100
sea	24	34	32	0	6	4	100
seem	28	35	19	1	6	ii	100
spring	22	26	30	8	5	9	100
almost	21	22	28	3	6	10	90
ground	22	25	27	6	7	3	90
mark	16	25	29	11	4	5	90
rich	18	29	30	4	4	5	90
table	23	18	29	10	4	6	90
buy	20	15	27	1	6	11	So
dead	18	34	16	1	5	6	80
hot	17	16	22	13	3	9	80
lead	13	25	24	7	7	4	80
picture	15	14	33	3	5	10	80
able	19	12	25	1	3	10	70
act	14	20	19	5	5	7	70
animal	18	8	29	7	3	5	70
appear	13	26	17	2	7	5	70
cannot	21	22	12	1	2	12	70
class	5	2	31	6	6	10	60
cup	14	18	13	7	4	4	60
distance	10	14	27	4	3	2	60
doubt	7	24	17	0	5	7	60
equal	3	19	25	7	2	4	60
bit	16	12	11	4	3	4	50
breakfast	16	6	15	4	1	8	50
demand	7	17	14	3	5	4	50
crowd	7	17	14	1	6	5	50
crown	10	25	7	4	3	1	50
advantage	1	10	17	1	4	7	40
ah	14	21	3	0	2	0	40
angel	6	27	4	2	0	1	40
aunt	13	4	11	0	2	10	40
brick	11	5	13	6	2	3	40
advice	3	8	12	1	2	4	30
avenue	4	3	13	0	5	5	30
ax (axe)	9	II	7	3	0	o	30
beach	8	9	7	1	3	2	30
bean	8	2	7	6	2	5	30
accordingly	2	6	7	1	1	3	20
acquire	0	3	12	1	1	3	20
ahead	6	0	7	0	1	6	20
ample	1	12	2	1	2	2	20
anyone	7	0	7	0	1	5	20
abate	0	8	2	0	0	0	10
abolish	0	4	3	1	2	0	10
abound	0	8	I	0	1	0	10
accommodate	1	1	4	1	1	2	10
acorn	5	2	3	0	0	0	10

Unlocking Language

TABLE V

RANK ORDER OF THE FIFTY WORDS OF TABLE IV, ACCORDING TO THE
CREDIT RECEIVED FROM EACH SOURCE

	Children's Literature	Bible and English Classics	Textbooks and Vocabularies	Common Facts and Trades	Newspapers	Correspondence	Total Sum of Credits
mean	6	11	3½	16½	3	3	3
nothing	1½	4½	16½	25	3	7½	3
sea	3	2½	2	45½	8½	31	3
seem	1½	1	19½	33½	8½	3	3
spring	6	7½	5½	4	15	11½	3
almost	8½	15½	10	21½	8½	7½	8
ground	6	11	12	10½	3	36½	8
mark	17	11	8	2	21	23½	8
rich	13	4½	5½	16½	21	23½	8
table	4	20½	8	3	21	18	8
buy	10	25	12	33½	8½	3	13
dead	13	2½	24	33½	15	18	13
hot	15	24	18	1	27	11½	13
lead	24	11	16½	6½	3	31	13
picture	19	26½	1	21½	15	7½	13
able	11	29	14½	33½	27	7½	18
act	21	18	19½	13	15	15	18
animal	13	35½	8	6½	27	23½	18
appear	24	7½	22	25	3	23½	18
cannot	8½	15½	32	33½	35	1	18
class	38½	46	3½	10½	8½	7½	23
cup	21	20½	29	6½	21	31	23
distance	27½	26½	12	16½	27	40½	23
doubt	33½	14	22	45½	15	15	23
equal	41½	19	14½	6½	35	31	23
bit	17	29	34½	16½	27	31	28
breakfast	17	38½	25	16½	43	13	28
demand	33½	22½	26½	21½	15	31	28
crowd	33½	22%	26^	33½	8½	23½	28
crown	27½	11	39	16½	27	43½	28
advantage	45	32	22	33½	21	15	33
ah	21	17	46	45½	35	47½	33
angel	36½	6	43½	25	48½	43½	33
aunt	24	41½	34½	45½	35	7½	33
brick	26	40	29	10½	35	36½	33
advice	41½	35½	32	33½	35	31	38
avenue	40	43½	29	45½	15	23½	38
ax (axe)	29	31	39	21½	48½	47½	38
beach	30½	33	39	33½	27	40½	38
bean	30½	46	39	10½	35	23½	38
accordingly	43	38½	39	33½	43	36½	43
acquire	48½	43½	32	33½	43	36½	43
ahead	36½	49½	39	45½	43	18	43
ample	45	29	48½	33½	35	40½	43
anyone	33½	49½	39	45½	43	23½	43
abate	48½	35½	48½	45½	48½	47½	48
abolish	48½	41½	46	33½	35	47½	48
abound	48½	35½	50	45½	43	47½	48
accommodate	45	48	43½	33½	43	40½	48
acorn	38½	46	46	45½	48½	47½	48

as too weak in credit to children's literature and newspapers, and much too weak in credit to words relating to common life and trades. It will, however, serve reasonably well until more extensive and specialized counts are made.

The following would be specially worth making:

A count of 250,000 to 1,000,000 words taken at random from a standard encyclopedia.

A count of 250,000 to 1,000,000 words taken at random from newspapers and weekly magazines.

A count of 250,000 to 500,000 words taken at random from boys' magazines.

A count of 250,000 to 500,000 words taken at random from girls' - magazines.

A count of 250,000 to 500,000 words taken at random from reading of a definitely utilitarian nature, such as printed directions for the use of machines, tools, medicines, etc.; laws and ordinances; proclamations; hand books; railroad guides; civil service examinations for policemen, firemen and the like.

USES OF THE LIST IN THE TEACHER'S WORD BOOK

THE TEACHING OF WORDS

Conscientious teachers now spend much time and thought in deciding what pedagogical treatment to use in the case of words which offer difficulty to pupils. In the third readers which they use they find, according to Miller, over nine thousand different words.¹² Some of these probably should not be taught at all in that grade; some should be explained at the time to serve the purpose of the story or poem, but then left to their fate; some should be thoroughly taught and reviewed. The Teacher's Word Book helps the teacher to decide quickly which treatment is appropriate.

The same service is performed, of course, in each of the school grades. Consider, for example, these words taken from Shelley's "Ode to a Skylark," a poem often found in school readers for grade seven or eight.

¹² Many of these, however, are derived forms.

Unlocking Language

blithe	harmonious	rainbow	unbidden
bright'ning	hidden	rapture	unbodied
chaunt	hymeneal	shrill	unpremeditated
chorus	hymns	soar	vaunt
divine	ignorance	spirit	wert
dost	madness	sprite	wherein
forth	matched	strains	wrought
fountain	melody	sunken	
gladness	panted	thine	
heeded	profuse	triumphal	

If the reader will decide for himself in which thousand each of these belongs, keep account of the time spent to reach a decision, and then compare his ratings with those on page 357 [31] derived from the list, he will have a sample of the gain in time and correctness of judgment due to using the list.

Even expert teachers have very inadequate and inaccurate notions of the relative frequency and importance of words. For example, thirteen expert teachers were asked to rank certain words as 10, if in the first thousand for importance; 9, if in the second thousand; 8, if in the third thousand; and so on, using 0 for words below the tenth thousand. They differed enormously from one another in their ratings for the same words; and any one of them gave widely different ratings to words which are of closely equal importance by our count or by the average voice of the thirteen. For example, their ratings for ten words all having 10 as a sum of credits, and thus being all in or near the lower half of the fifth thousand, were as shown below:

RATINGS OF 10 WORDS BY 13 TEACHERS (I, II, III, ETC.)

(The + sign means "more" or "over.")

	I	II	III	IV	V	VI	VI I	VI II	IX	X	XI	XI I	XI II	Range for the same word
abate	1	3	6	4	7	7	0	3	6	8	9	3	6	9000 or +
abolish	1	5	8	3	8	9	9	6	8	10	10	7	7	9000
abound	6	6	8	6	6	8	4	6	8	10	6	3	7	7000
accom- modate	8	10	7	4	8	6	10	8	8	10	10	8	7	6000
acorn	1	1	7	10	7	7	9	7	9	5	10	8	8	9000
admission	7	6	9	4	9	8	10	8	8	10	8	10	10	6000
adversary	2	4	7	3	6	5	5	0	7	7	5	2	5	7000 or +

1921—E. L. Thorndike and the Familiar Word

agency	4	6	5	8	8	9	9	7	8	10	10	6	6	6000
aisle	5	5	4	6	10	6	10	7	7	10	8	3	8	7000
allegiance	3	2	4	..	9	7	9	10	8	10	8	6	10	5000
														or +
Range for the same teacher in 1000's	6	9	5	7	4	4	10 or +	7	3	3	5	8	5	

RATINGS OF THE WORDS FROM THE "ODE TO A SKYLARK" BY THE LIST, THE
NUMBERS REFER TO SUCCESSIVE THOUSANDS. "NOT" MEANS THAT
THE WORD IS NOT IN THE 10,000.

blithe	7	harmonious	7	rainbow	3	unbidden	not
bright'ning	not	hidden	4	rapture	4	unbodied	not
chaunt	10	hymeneal	not	shrill	4	unpremedi- tated	not
chorus	7	hymns	4	soar	5	vaunt	7
divine	2	ignorance	4	spirit	1	wert	4
dost	3	madness	4	sprite	6	wherein	4
forth	1	matched	2	strains	3	wrought	3
fountain	2	melody	4	sunken	10		
gladness	5	panted	3	thine	3		
heeded	3	profuse	7	triumphal	7		

In teaching arithmetic, history, geography, civics, or elementary science, there will be found in the book lessons many words which some of the pupils will not understand. Which are these and in which cases should the occasion be used to master a word for future use? Decision obviously depends in part upon how important the word is. For example, the first twenty-five pages of a standard geography for elementary schools contain these words:

accumulate	cable	consequently
alluvial	canal	continual
ant	capital	continue
area	capitol	definition
barrier	camel	delta
beaver	cascade	department
boulder	churn	deserted
bounce	clay	ditch
Buffalo	cliff	drain
burrow	climber	dweller

What the teacher should do with each of these words depends partly on its special importance for geography, and partly on its general importance then and later for pupils of the grade in question. It will be found, with respect to the latter, that some of these words rank as high as the first thousand, while some of them do not appear in our list at all, and probably would not appear even in a list fifty per cent larger.

The *Teacher's Word Book* does not, of course, rate correctly for any one community, words which are very important locally (as, for New York City, *subway*, *elevated*, *Brooklyn*). By its very existence, however, it directs attention to this issue, and stimulates the educational authorities to extend and amend it in respect to words of special local importance. In the case of spelling, the publication of general lists has been notably effective in producing the reaction of attention to special local lists; and we may expect the same effect from this reading- or meaning-list.

THE ESTABLISHMENT AND CLARIFICATION OF STANDARDS

The list makes it much easier than it has been in the past to put standards for word knowledge, by grades, by ages, or by mental ages, into clear, definite comprehensible form. For example, we may say that at a certain mental age or grade the minimum standard should be knowledge of the meanings of 95 per cent of the first 2500 worlds, 80 per cent of the next 1000, 60 per cent of the next 1500, and 20 per cent of the next 5000. If it seems desirable, we can specify still more narrowly, for example, 100 per cent of the first 1000, 95 per cent of the next 1000, 90 per cent of the next 1000, 80 per cent of the next 2000, with or without stipulation of a knowledge of the second 5000.

The actual learning of meanings is probably best accomplished by a large amount of relatively easy reading, plus a much smaller amount of harder reading with recourse to the dictionary, plus a still smaller amount of specific teaching of meanings as such. This actual learning of meanings also may be accomplished by means of varying and unsystematic stimulation of individual pupils. But the testing and keeping account of the knowledge gained does need to be in terms of specific word knowledge, and the list is a great aid in defining and testing such word knowledge.

THE EVALUATION OF TEXT-BOOKS

Within very recent years there has been quietly developing a demand for objective, scientific evaluation of text-books and other instruments of instruction. A first book in reading, for example, is being judged by a system of credit points for type, spacing, number of words used, quality of the English, interest of the selection to little children, and the like. One element in such an evaluation of almost all text-books is the suitability of their vocabulary to the grade for which they are intended. This can be measured with absolute impartiality with the aid of the list. One has only to make a word count of a sufficient sampling of pages from the book in question, and look up the ratings of the words on the list. For example, when it is found that of two contemporaneous beginners' books in arithmetic to be read by pupils in the first half of grade 3, one has in the first fifty pages eleven words that are not in our 10,000 at all, and twenty-five more that are not in the first 5000, whereas the other has five and fifteen as the corresponding numbers, it is obvious where credit belongs for wisdom and care in the choice of words.

SELECTION AND GRADATION IN READERS

The importance of the words in readers and other reading material and the gradation of this material should, of course, be one element in its evaluation. Without a word list such as this, however, the judgments have necessarily been subjective and rather vague. With this list, they can be absolutely impartial and precise to any desired degree. The results of the studies, which I hope this list will stimulate, are likely to be far-reaching in their exposure of imperfect selection and gradation of material in even our best instruments of instruction. Until such a list was available, indeed, the labor of inspecting material for details of vocabulary was too great perhaps to be expected from authors. As an illustration, I have taken at random the last pages of Book II, the first pages of Book III, and the last pages of Book III, counting approximately 4500 words in the case of each, from one of the best of present series of readers.

The vocabulary of the beginning of Book III is actually wider than that of the end of the book: 842 words to 736! From the end of Book II to the beginning of Book III there is a jump of over 20 per cent, from 681 to 842.

TABLE VI

WORDS IN THE FIRST 10,000 BECAUSE OF THEIR PRESENCE IN
FIRST, SECOND AND THIRD READERS¹³

bagpipe	Blitzen	bushtail	cluck
banter	bluebell	buzzard	cob
begone	Bobby		codfish
belfry	bonny		comical
bin	bo-peep		coon
blink	bowlful	cackle	coverlet
blur	Boxer	calyx	Cox
bonfire	boxwood	cipher	cozily
bun	brown	crier	cozy
baa	brawny	cruise	crape
badger	Bremen	caller	creak
ball-room	brig	caramel	cress
bantam	brimful	catnip	crock
barnyard	Brom	can	crocus
bask	broth	changeful	cub
beanstock	Brownie	chanticleer	cupola
beck	browse	charley	curd
bedside	bruce	chee	curt
beefsteak	brunette	cheerless	czar
belated	Bruin	chick	czardom
bespoke	bumble	Chinaman	
betook	bumble-bee	chirk	
Betsy	bummel	Cinderella	
birdie	Bunny	clapper	
blare	burrow	clement	

Attention to the importance and difficulty of words in selection and gradation is only one of many possible desiderata to be considered in a series of readers. It may well be sacrificed from time to time for the sake of literary excellence, or interest, or informational value, or other worthy qualities. But it should not be sacrificed to no purpose; and it has just claims for much consideration not only in grades 1 to 3, but to the very end of the elementary course.

The list, when used in connection with a word count of any instrument

¹³ Some of these are not in the 10,000 as printed since there were over 1700 words of credit 3, which carried the list to about 10,800. So the less important of the "3" words were left out of the final list.

of instruction, will probably often lead to constructive recommendations of some importance. Two such may be noted here. The first is that primers and first readers should try to secure interest and adaptation to childish ideas and activities, without recourse to rare and even fictitious words. The second is that they should try to provide for phonic experience and practice without recourse to such rare or fictitious words. There are words in primers and first readers which do not even rank in our 10,000, or would not rank there except for the credits they get by virtue of their use in primers and first readers. Children are taught to read words in the first year of school which they may actually not see again for years.

TABLE VII

WORDS IN THE FIRST, SECOND AND THIRD READERS WHICH ARE
NOT IN THE 10,000

backwoods	bluejay	calmy	cleak
benighted	bluebird	Campbell	climax
baff	blumb	cantaloupe	climber
bagful	Blynnkin	Canute	clop
baggy	bobtail	carman	clump
bab	boo	Caroline	cockleshell
bannock	Borneo	Carrara	columbine
barbecue	Bose	Caspar	commoner
barrelful	bossy	Charlotte	Conrad
basketful	Bowden	chatterer	Cora
beanshooter	bricklayer	cheep	corncob
bearskin	Brinker	chipmunk	Corrola
beehive	brooklet	chirrup	counterpane
Bert	broomstick	chore	craggy
Biddy	burdock	chubby	creepy
biff		chug	crispy
bight		chump	croon
birdling		churnful	crotch
blithesome		citron	crumple

A list of words which would not be in the 10,000 except for their presence in the first, second or third readers is instructive from many points of view. This list for words beginning in *b* and *c* is given on page 33 [Table VI]. Above [Table VII] is a list of words which are in the first, second, and third readers, but are not even in the 10,000.

MATERIAL FOR PHONIC DRILLS

It is interesting to note those words which are suitable to develop phonic insights and habits and are among the thousand most important words according to our count. I have, therefore, taken about seventy of the most useful phonograms, and entered after each phonogram the words from the first thousand of our list which present it clearly. Some of these words are not equal in interest to the words now used in beginning reading for the purpose, but on the whole, they will form a very serviceable basis for phonic drills; and every one of them is well worth learning for its own sake.

ace—face place race space
ack—back black
ade—made shade trade
ail—sail
ake—cake lake shake take
all—all ball call fall hall small tall wall
ame—came game name same
an—an can cannot man manner plan ran than
and—and band command demand hand land sand stand
at—at fat hat matter sat that
ate—gate late state
ay—away bay day gray lay may pay play say stay today way
bl—black bless blind blood blow blue
br—branch brave bread break breakfast bridge bright bring
 broad broken brook brother brought brown
ch—chair chance change charge chief child children choose church
cl—class clean clear clock close cloth clothing clothes
cr—cried cross crowd crown cry
dr—draw dream dress drink drive drop dry
eam—dream
eep—deep keep sheep sleep
eet—meet street sweet
ell—bell fell fellow tell well
en—men pen ten then when
ent—cent center different enter entire sent went
est—best nest rest yesterday
et—get let letter met set settle yet
fl—floor flow flower fly
fr—free French fresh friend from front fruit
gl—glad glass
gr—Grace grain grant grass gray great green grew ground grow
ice—nice price
ick—prick quick sick stick thick
ide—beside decide divide guide hide ride side wide
ight—bright delight fight light might night right sight
ill—bill fill ill kill mill still till will
in—begin in inch increase indeed Indian instead interest into

skin thin win
ine—line mine nine shine
ing—being bring coming during evening going king morning
ring sing and many others
ip—lip ship trip
it—fit it its sit
ite—quite white write
oat—boat coat
ock—clock lock rock stock
old—cold gold hold old sold told
ong—along belong long song strong wrong
ook—book brook cook look took
oon—noon soon
op—shop stop top
ot—hot not
ound—around found ground pound round sound
out—about out outside shout
own—brown crown down town
ox—box
pi—place plain plan plant play pleasant pleasure
pr—practice press pretty price prince promise proper prove proud
qu—quarter queen question quick quiet quite require
sh— shade shake shall shape she sheep shine ship shoe
shop shore short should shoulder shout show shut
sk—skin sky
sl—sleep slow
sm—small smoker
sp—space speak spend spirit spoke spot spread spring
st—stand star start state station stay step stick still stock
stone stood stop store storm story study
str—straight strange stream street strength strike strong
tr—trade train travel tree trip
wh—what wheat wheel when where whether which while
white who whole whom whose why

A STANDARD VOCABULARY FOR THE TEACHING OF FOREIGNERS

By the elimination of certain specially childish or "literary" words from the first 500 of our list and the addition of certain words of special importance to the newcomer to America, such as *danger*, *poison*, *cent*, *dollar*, *entrance*, *exit*, we shall have a basic list of great value in teaching foreign adults to read English. A tentative first 500 for foreigners may be formed simply by omitting *ball*, *being*, *pretty*, and *soldier* from the list on page 365, and adding *danger*, *poison*, *cent* and *dollar*. A second 500 can be formed in much the same way. The use of two such lists, with any additions necessary to arouse interest and meet local needs, may be expected to improve and facilitate the teaching of English to foreigners above the age of twelve.

THE VOCABULARIES IN ELEMENTARY BOOKS FOR TEACHING A FOREIGN LANGUAGE

Investigations by Bagster-Collins have shown that first-year books in French, German, and Spanish differ enormously in the words used, and that many of the words in any one of them are such as should not be learned by the beginner. It is also probably the case that in any one of them some very important words will be given little or even no attention. Until counts for the foreign language itself, comparable to our count for English, are available, it will be worth while to check the vocabulary of a foreign language text-book against our list. Words of little importance according to our list should receive little emphasis in teaching, unless they are of clear service to the student. Words of much importance according to our list which are given little or no practice by the text-book should be provided for by additional exercises unless they are words like *acre*, *baseball*, or *inch* where use is restricted chiefly to English speakers and writers.

THE FIRST 500 WORDS

A about above across add after again against air all almost alone along also always am among an and another answer any apple are arm around as ask at away back bad ball bank be bear beautiful because become bed

been before begin behind being believe best better between big bird black blow blue body book both box boy bread bring brother brought build burn but buy by call came can care carry case cause certain change child children church

city clear close cold color come company corn could country course cover cross cut dark day dead dear death deep did die do does done door down draw dress drink drive drop during each ear early earth east eat egg

end enough even every eye face fair fall family far fast father fear feel feet few field fill find fine fire first five floor flower fly follow food foot for form found four free fresh friend from front full

garden gave general get girl give given glad go God gold good got great green ground grow had hair half hand happy hard has have he head hear heart heavy help her here high hill him himself his hold home

hope horse hot hour house how hundred I if in into is it its

just keep kill kind king know known land large last late laugh
law lay lead learn leave left length less let letter lie life light
like

line little live long look lost love low made make man many
mark matter may me mean measure meet men might mile milk
mind mine miss money month more morning most mother mountain
move much must my name near need

never new next night no north not nothing now number of off
often old on once one only open or order other our out over
own paper part pass pay people person picture piece place plain
plant play please point

poor power present pretty put quick rain raise reach read ready
reason receive red remain remember rest rich ride right river
road rock roll room round run said sail same save saw say
school sea second see seem seen send

sent serve set several shall she ship short should show side sight
silver since sing sister sit six sleep small so soft soldier some
something sometime son soon sound south speak spring stand
start state stay step still stone stop

story street strong such summer sun sure sweet table take talk
tell ten than thank that the their them then there these they
thing think third this those though thought thousand three through
till time to today together too took top

town train tree true try turn two under until up upon us
use very visit voice wait walk wall want war warm was
watch water way we week well went were what when where
which while white who whole why

wide will wind window winter wish with without woman wood
word work world would write year yet you young your

TEXT-BOOKS FOR HIGH SCHOOL PUPILS

The instruments of instruction in mathematics, science, history, civics, and even literature, used for pupils in high schools, especially in Grade 9, may well be scrutinized from the point of view of the Teacher's Word Book. Other things being equal, it is better not to burden a subject like algebra or chemistry with unnecessary linguistic difficulties. The slight gain from a widened vocabulary is more than balanced by the loss in ease of comprehension of the principle to be taught.

As an illustration, note some of the words used in the first fifty pages of two well-known text-books in algebra which are not in the list at all. Some of these should probably not have been used. In the case of others the preparation for the lessons in question should include special attention to the meanings of the words.

WORDS IN THE FIRST FIFTY PAGES OF A STANDARD ALGEBRA WHICH ARE NOT IN THE 10,000

abbreviate	Diophantus	literal	prefix
Ahmes	dissimilar	mathematical	Pythagoras
algebraic	distinctive	mathematician	quadrilateral
algebraist	elementary	minuend	rearrange
arithmetical	et	monomial	redwood
assets	exponent	multiplicand	reintroduce
binomial	facilitate	multiplier	rewrite
bricklayer	formula	ninth	scalepan
casting	formulae	nitrogen	simplify
coefficient	Harriot	numerical	Stifel
complementary	haw	Oughtred	subtrahend
computation	Herigone	parallelogram	supplement
debit	Hindu	parenthesis	supplementary
Demosthenes	ingenuity	Pell	trinomial
denominate	integer	perimeter	Vieta
Descartes	Leibnitz	polynomial	Widmann
digit	likelihood	potentia	workmanlike

THE AMENDMENT AND EXTENSION OF SPELLING LISTS

Ours is not a spelling list, and the order of importance of words for spelling will often diverge widely from that of this list. Words like *ache*, *cough*, *hoarse*, *doctor*, *medicine*, *coal*, *shoes*, *waist*, *dear madam*, *yours truly*, *Mass.*, *Conn.*, *Ill.*, *Neb.*, will figure relatively much more frequently in writing than in reading. In a spelling list also the derivatives may best be counted separately.

This list will, however, be useful to correct certain notable omissions from the spelling lists, such as parts of irregular verbs from the Cook-O'Shea list, or names of the days, months, states, large cities, and the like from the Ayres list. It will be still more useful to extend spelling lists beyond 2500 or 3000, as seems necessary with pupils in high schools and pupils who expect to become stenographers; and this may also be desirable for others. The work being done at the University of Iowa by Horn

and Ashbaugh will soon, I trust, provide us with a list even more adequate for spelling than this list is for reading.

THE IMPROVEMENT OF READING SCALES AND TESTS

This list makes it possible to devise tests and scales for word knowledge which will be very much superior to any that we now have. Existing tests and scales are made better at once because we can give to each element of the test a rough measure of importance.

It also becomes much easier than it would otherwise be to extend tests of word knowledge by alternative forms; and to assign a provisional gradation for difficulty within each form. Finally, certain very promising new methods of testing word knowledge become practicable, when we have ten thousand words graded fairly well as to importance.¹⁴

OTHER USES

It is hoped that the Teacher's Word Book will be of service to students of education in many other ways. It seems, for example, to offer an excellent chance to measure the relative importance of words of Latin derivation, and of the extent to which a given knowledge of Latin may be expected to help a pupil of a given degree of ability to understand the present meanings of these words.

The frequency of the use of words in the reading matter of any given time for any given group is of some value as an index of the knowledge and interests of that group at that time. The words which are not in concordances of Shakspeare and Milton but are very common today would, I think, make an instructive list. The words which are not in Shakspeare, Milton, Pope, Cowper, Wordsworth, or Tennyson, but are common today, are perhaps more instructive. The words which figure largely in the two newspaper counts but only slightly elsewhere in our material form an interesting group. Among them are, for example, *hosiery*, *millionaire*, and *nomination!*

There is, of course, a notable lag of a list like this behind the actual reading material of 1920. It may well be that automobile, auto, and Ford, are now read oftener than horse, though the credit sum on our list is 108

¹⁴ It may be noted here that the List will be of value in the arrangement of psychological tests in general, by enabling their authors to be sure that the words in the instructions for a test are sufficiently easy to understand.

for horse to 33, 13, and less than 10^{15} for the other three, respectively. The lag in school readers is perhaps even greater than that in our list.

Children read about giants and fairies, knights and castles, kings and queens, forms of work and fighting, and ways of thought and belief, which are really a sort of paleontology of civilization. This does not, of course, do much harm, since the misleadings about the facts of human nature and the world are probably easily curable, and the literature of folk-lore, feudalism, and militaristic societies is reputed to have great merits of simplicity, interest, and literary quality. In general, material finds entrance to our readers a generation or more after it is written. Possibly this is wise. Literary critics as a rule agree that you must wait for the long test of time to decide what is really great and fine in literature. They would presumably not make such a confession of their own lack of acuity if it were not true. The contrast with science, where the elementary student may learn about, say, the electrons, a year or two after their discovery, should, however, receive our attention. Since school histories also usually spend 95 per cent of their energies on the world minus the decades since the pupil and his older brothers and sisters were born, the pupil's academic reading acquaintance with human affairs is almost entirely out of date. It may be well for it to be so. It surely is well that teachers should understand that it is so. The *Teacher's Word Book* may also help us to understand why so many children and so many adults read what the cultivated man condemns as trash. Dr. Jordan¹⁶ has shown that boys will wait for an hour at public libraries to get books by certain present-day writers whose names the cultivated man has never even heard. It is commonly assumed that children and adults prefer trashy stories in large measure because they are more exciting and more stimulating in respect to sex. There is, however, reason to believe that greater ease of reading in respect to vocabulary, construction, and facts, is a very important cause of preference. A count of the vocabulary of "best sellers" and a summary of it in terms of our list would thus be very instructive.

PREFIXES AND SUFFIXES

It was formerly customary to teach the meanings of prefixes and suffixes. Such teaching became discredited, partly because it extended to rari-

¹⁵ 18 for ford and Ford together. At least 8 are for ford.

¹⁶ Jordan, A. W.. *Children's Interests in Reading*. Teachers College, Contributions to Education, No. 107.

ties like *agogue* and *ambulist*, partly because the prefixes and suffixes were divorced from their connections in real words, and partly because of the ambiguities of many of them. The word counts reveal the very great importance of certain prefixes and suffixes, and suggest that they are worth teaching, in proper connections, even though the pupil will be occasionally misled. For example, there are in our list 170 words beginning with *un*. Of these, 15 begin with *under*, leaving 155. The remainder of the word is, in 140, or 90 per cent, of these 155 cases in the list as a word by itself. In only about four per cent will the simple negative or oppositional meaning of *un* mislead pupils (save perhaps a very few very dull pupils). If pupils were taught¹⁷ *un*—after learning, say *unhappy*, *unkind*, and *unwilling*—it would probably represent a considerable economy over leaving them to discover its meaning haphazard in the course of further reading. The advisability of teaching any prefix or suffix should be considered in the light of similar data about it now readily available in the list.

This report has already overrun its allotment of space, but I must at least mention the fact that one chief service of the *Teacher's Word Book* will be to aid in the production of some much better list, from wider counts, to replace it. The entire procedure in counting, entering and crediting words is made very much easier, once we have an approximately correct list for use in recording the counts economically.

¹⁷ The cases where such misleading might be considerable are: *unanimous*, *uniform*, *unit*, *unless*, *unto*, and *untoward*.

1923—The Lively and Pressey Measuring Method

Introduction

BERTHA A. LIVELY and Sidney L. Pressey of Ohio State University were concerned with the practical problem of selecting science textbooks for junior high school. Books at that time were so overlaid with technical words that teachers spent all class time teaching vocabulary.

In 1923, they published their study, “A Method for Measuring the ‘Vocabulary Burden’ of Textbooks” in the journal *Educational Administration and Supervision*.¹

They argued that it would be helpful to have a way to measure and reduce the “vocabulary burden” of textbooks.

Their study tested three different methods for measuring the vocabulary load of a thousand words of text:

1. The first method used the number of different words (the vocabulary range).
2. The second method used the number of “zero-index words,” words **not** in *The Teacher’s Word Book*, the Thorndike list of 10,000 words.
3. The third method used the median of the index numbers of the words taken from the same Thorndike list of 10,000 words.

They tested the three methods on 15 textbooks of different difficulties, along with one newspaper. The low end included a second and a fourth-grade reader and Stevenson’s *Kidnapped*. The high end included a college physics textbook and an elementary chemistry textbook.

They found that the median index number was the best indicator of the vocabulary burden of these reading materials: the higher the index number, the easier the vocabulary; the lower the index, the harder the vocabulary.

¹ Vol. 9, No. 7 (October 1923), pp. 389-398.

1923—The Lively and Pressey Measuring Method

The Lively-Pressey study demonstrated the effectiveness of a statistical approach for predicting text difficulty. It had a great influence on the readability formulas that would follow and also use the Thorndike word list. As the authors announced, “The fundamental value of Thorndike's contribution is obvious; the ‘Word Book’ has opened up a whole new field for investigation.”

—WHD

Educational Administration and Supervision

INCLUDING TEACHER TRAINING

Vol. IX

October, 1923

No. 7

A Method for Measuring the "Vocabulary Burden" of Textbooks

BERTHA A. LIVELY AND S. L. PRESSEY

Ohio State University

I. NEED FOR A SYSTEMATIC METHOD FOR INVESTIGATING VOCABULARY BURDEN

THE PRESENT study was begun as a result of a minor investigation regarding the number of technical words in a certain junior high school science book. The study revealed an astounding number of technical terms—a number so large (as testified by teachers using this book) that the course often became quite as much a study of scientific vocabulary as of scientific facts. This investigation brought out in striking fashion the importance of the question as to comparative vocabulary burden, in public school textbooks. The problem is perhaps most acute in connection with junior high school science books. But it is also an important problem in reading; some method for measuring vocabulary difficulty in supplementary reading material should be decidedly worth while. The present paper presents an effort to develop a method capable of dealing with these problems, together with results of application of this method to certain representative textbooks from second grade readers to a medical school physiology.¹

II. METHOD AND MATERIAL

The method finally adopted, after an extended series of preliminary investigations which need not be presented here, may be briefly described. Two questions were involved in the elaboration of procedure; (*a*) How

¹ The writers wish to acknowledge their obligations to the Graduate Council of the University for funds to assist in the clerical labor.

many words must be included in any sampling from a textbook in order to obtain a reliable indication regarding vocabulary—and how should these words be selected? (b) How can the difficulty of the words in this sampling be best measured?

The method of vocabulary sampling finally adopted dealt with thousand word units obtained from a systematic sampling throughout the text. First of all, the investigator noted the number of pages in the book and the approximate number of words per line; he then estimated the number of pages which should be sampled, taking one line per page, in order to cover 1000 words, and chose pages so that the sampling would be evenly distributed throughout the book. Thus if the book contained approximately 500 pages, and there were about 10 words to the line, a line on each fifth page throughout the book would make up about a thousand words. The investigator then went through the book, counting up the number of words found on the third line of each fifth page until exactly 1000 words were obtained. The third line was used as conveniently found on a page. If the book were shorter, every other page might be taken, if much longer, every tenth; pages were chosen simply to give a systematic sampling throughout the book. The lines chosen through to the thousandth word were now gone over, and all the different words found in this thousand listed and alphabetized.

Once the thousand-word count was made, the total number of different words per thousand was first noted. This gave what has been called vocabulary range. Next, these words were looked up in the Thorndike "Word Book"² and the index number for each word was found. The number of words was now counted, in the thousand-word sampling, not appearing among the most common 10,000 words; these words were listed as zero value words and may be taken to indicate the size of the technical vocabulary. The weighted median index number was finally calculated.

² Thorndike, E. L.: "The Teacher's Word Book," Teacher's College, Bureau of Publications, Columbia University, New York City. This book lists the 10,000 most common words of the English language, as determined on the basis of an elaborate investigation by Professor Thorndike. In this "Word Book" each word is followed by an index number indicative of its commonness. Thus such a common word as "and" has an index number of 210; a relatively uncommon word like "atom" has an index number of 4; still more rare words such as "neolithic" do not appear in the word book at all—those are listed as zero value words. Words with credit numbers of 49 or over occur in the first 1000 words, in frequency. Words with index numbers of 10 or over occur in the first 5000 words—and so on; for a more detailed statement, the reader is referred to the "Word Book" itself.

This is simply the median index number with zero value words counted twice. Evidently the higher the median index number the easier the vocabulary.

Sixteen different types of reading matter were thus studied; Three second-grade readers (Jones, Aldine, Horace Mann), three fourth-grade readers (Jones, Aldine, Horace Mann), Stevenson's "Kidnapped" Thackeray's "Vanity Fair" the Columbus *Dispatch* (as a sampling of newspaper vocabulary—only the first page was taken), Muzzey's "American History" Clark's "General Science" and "Introduction to Science" (as representative of Junior High School books in science), Hunter's "Elements of Biology" McPherson and Henderson's "Elements of Chemistry," Kimball's "College Physics," and Howell's "Physiology."

The results below summarize the findings regarding these materials as to (a) range of vocabulary, (b) size of highly technical vocabulary (zero value words), and (c) weighted median index number. For each book two counts were made, in order to determine the reliability of the method. The second sampling was made exactly as the first except that a different page was used. Thus, if the first count used the pages 5, 10, 15, 20, then the second count used pages 1, 6, 11, 16, the third line on each page being studied in each case. As will be noted, the method seems fairly reliable. If it is desired to increase the reliability it is suggested that additional thousand-word counts be made and the results averaged. This would seem more satisfactory than increase in the size of a single sampling since the thousands-word count is a very convenient unit, and after a number of thousand-word counts are made, comparisons from one count to another are of some interest.

SUMMARY REGARDING VOCABULARY BURDEN—16 TYPES OF MATERIAL

Counts	Range		Zero value words		Weighted median	
	1	2	1	2	1	2
Second readers: Jones.....	371	350	4	4	86	88
Horace Mann.....	412	421	9	9	78	83
Aldine.....	367	353	7	6	77	79
Fourth readers: Jones.....	471	454	12	20	71	62
Aldine.....	450	455	24	11	63	69
Horace Mann.....	466	472	15	17	65	66
Stevenson: Kidnapped.....	402	415	21	30	67	65
Thackeray: Vanity Fair.....	490	459	43	34	43	54
Columbus Dispatch.....	528	581	49	45	33	37
without local names.....	514	560	35	24	38	45
Muzzy: American History.....	533	506	24	30	38	40
Clark: Introduction to Science.....	483	491	22	25	52	50
Clark: General Science.....	480	463	30	30	43	45
Hunter: Elements of Biology.....	464	467	57	57	28	34
Elementary Chemistry.....	399	358	67	69	22	14
Kimball: College Physics.....	393	405	60	59	24	22
Howell: Physiology.....	422	473	108	94	4	10

III. RESULTS

(a) Range of Vocabulary.—The following table summarizes all results for these materials. As will be noted, the table shows (a) range, (b) zero value words, and (c) weighted median index numbers—and for each one of these items the results of the two counts are shown.

The results are, however, best shown in graphic form. Chart I shows the difference in range.

As might be expected, the range of vocabulary in second-grade readers is small. But two or three unexpected findings do appear. Thus the range of vocabulary in "Kidnapped" is also small. Stevenson evidently gets his effects not by using a large number of words, but by sentence structure, and other devices. Range in the science books is low; these books appar-

ently use their technical terms over and over again, and use besides these technical terms relatively simple words. It is suggested that these figures may be used as tentative bases for comparison in further counts, the average range of the two samplings being employed. If the method seems of value, it is intended that norms for readers in the various grades, and for various types of books, should be developed.

(b) *Number of Zero Value Words.*—Again, graphical presentation is the most satisfactory.

As would be expected, the second readers show the smallest number of words outside the 10,000 most common words. It is somewhat startling, however, to find that second readers do include a few such terms. The number of zero value words in the newspaper is somewhat high. This is due in part to the number of local names included; the table presents also figures for the newspapers with these local names left out. It was found very difficult, however, to decide just which names might best be eliminated; so the chart presents the results without such elimination. The large number of technical terms in the Junior High School biology is of decided interest, and is indicative of the vocabulary burden of this book. The Medical School physiology also has a huge number of such terms, as would be expected.

(c) *Weighted Median Index Numbers.*—These results are presented in graphical form in Chart III. It is felt that the weighted median index number is probably the best measure of vocabulary burden. It will be noted that "Kidnapped" is about at fourth grade reading difficulty; it is suggested that "Kidnapped" might well be used as supplementary reading at about the fourth or fifth grade. Other details regarding the comparative standing of the various books are obvious from the chart and need no comment.

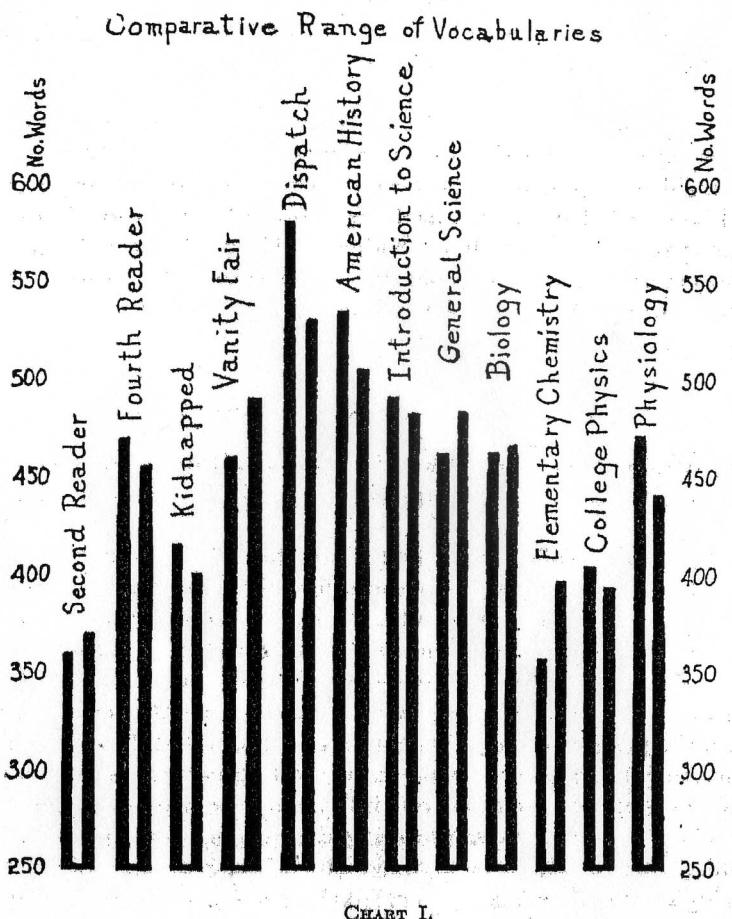


CHART I.

IV. POSSIBLE DEVELOPMENTS OF THE METHOD

The question now is as to the values and limitations of the method as thus illustrated. But perhaps the limitations should be pointed out first.

It should be re-called in the first place that the reliability of these thousand-word samplings—as determined by comparisons of two samplings for each book—was called “fair.” From trying experience in the field of tests, it has come to be realized that reliabilities first considered fair might be by no means as good as they should be. It can at least be said of the present study that data are presented which make possible some judgment as to what reliability, in a given instance, may be expected. More im-

portant, however, is the possibility (as was suggested), that the reliability may be increased as desired by taking further thousand-word samplings. Presumably, the reliability desired will depend upon the nicety of the distinctions which it is desired to make; very likely also the reliability of the weighted median index number and number of zero value words will be conditioned somewhat by the range of vocabulary. It is one important merit of the general procedure suggested that it is elastic, and thus adapted to such various demands or conditions.

It should also be listed in the catalogue of limitations that the description of the sampling in terms of range, number of zero value words, and weighted median index number is undoubtedly a description which leaves out certain important elements. This is, of course true of any method of statistical summary; features appear, when the complete distribution is studied, which are lost in the scheme of averaging. Thus in the present study, the history seems to involve a greater proportional number of words in the last 5000 of the 10,000 most common words than any of the other books studied. But no special features of any of the distributions have appeared of sufficient prominence to demand special treatment.

It should also be mentioned as a third handicap that the method is so involved in use of the Thorndike "Word Book" as to partake of any faults that that book may have. The writers' work has emphasized the extent which Thorndike has weighted his investigation in the direction of literary and even poetical vocabularies.³

³ Comparisons of thousand-word counts. (not reported here) on the "Golden Treasury" and other literary materials with the newspaper vocabularies, as to overlapping, have brought out this feature.

Number of Zero-Value Words

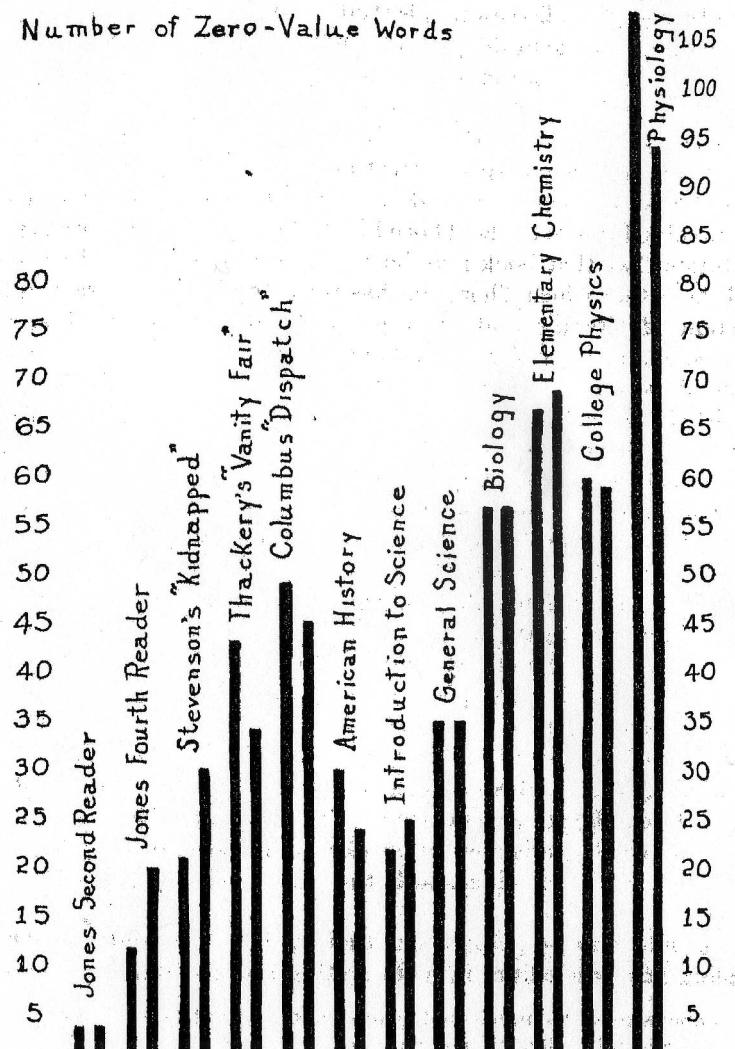


CHART II.

The study has also made clear the advantages which would have accrued to Thorndike's work, if a more systematic sampling on his part had permitted the inclusion, with the index numbers, of the frequencies with which each word occurred, per 1,000,000 words.

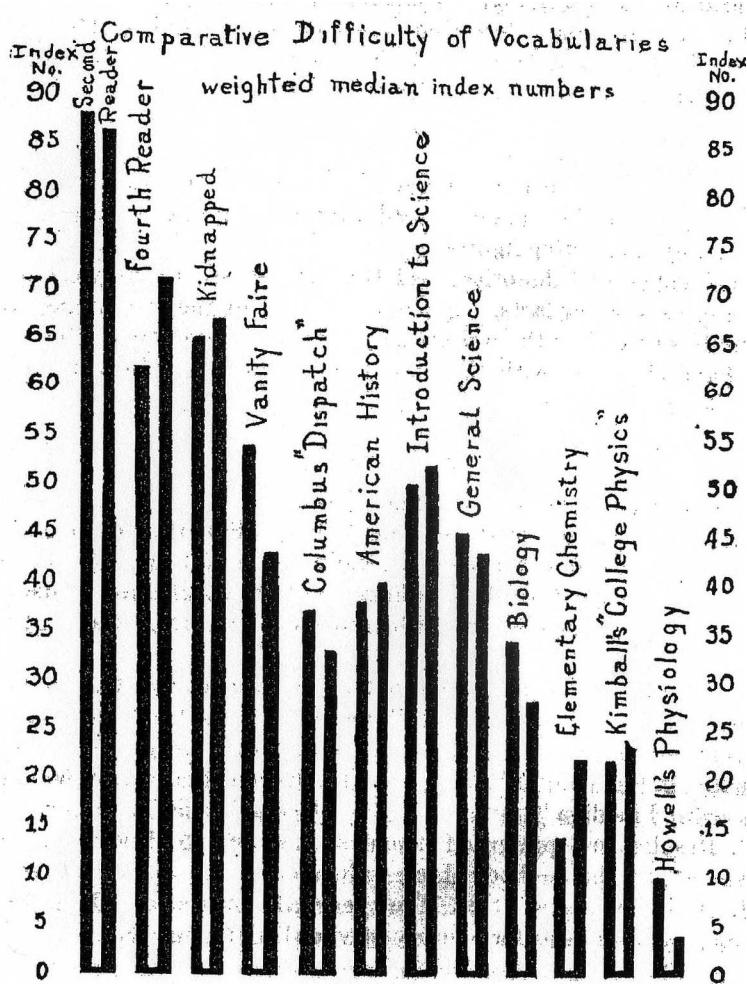


CHART III.

But now for some of the potentialities of the method. It would seem obvious in the first place that, with the development of such further data as might be considered to establish norms, such procedure should be of considerable use in evaluating texts or other reading material. A thousand-word count can be put through in about three hours. Even though several counts might be necessary, in order to obtain sufficient consistency from one count to another to give one confidence in the findings, still the time cost would not appear excessive.

It should also be pointed out that such a systematic method of sampling has possibilities in investigating the distribution of vocabulary burden through a book. Many texts appear to have the vocabulary load at the beginning. A thousand-word count in each chapter should make possible interesting comparisons regarding this matter. Finally, it should be mentioned that certain further developments of the method are possible. For instance, under certain circumstances over-lapping from one book to another may be of importance. Thus in studying the additions to technical vocabularies involved in chemistry, after a course in general science has been taken, comparison of a thousand-word counts for over-lapping, from one book to another yields many interesting figures.

The reader will doubtless feel that these are great expectations on the basis of a few facts. Quite so it is. But the study has seemed distinctly suggestive; the writers are therefore presenting what results they have obtained, with the hope that others may be interested to work along these lines.

SUMMARY

The paper may be briefly summarized.

1. It is suggested that .the vocabulary difficulty or vocabulary burden of a book or other piece of reading material, may be evaluated by taking thousand-word samplings of the vocabularies used and examining these samplings with reference to the type of word employed.
2. Three methods are suggested for summarizing the facts with regard to such a sampling: (a) Range of vocabulary, or number of different words per 1000 words sampled, (b) number of words not occurring in the Thorndike list of the 10,000 most common words, and (c) weighted median Thorndike "Word Book" index number.
3. Results are presented from study of 15 books, and one newspaper, by the methods above indicated.
4. It is suggested that the general procedure has decided possibilities, as a basis for a study of vocabulary burden.

1928—Vogel and Washburne: The Winnetka Formula

Introduction

IN 1928, Mabel Vogel and Carleton Washburne of Winnetka, Illinois, published one of the most important studies of readability.¹ They were the first to study the structural characteristics of the text and the first to use a criterion based on an empirical evaluation of text. They studied ten different factors including kinds of sentences and prepositional phrases, as well as word difficulty and sentence length. Since, however, many factors correlated highly with one another, they chose four for their new formula.

Following Lively and Pressey, they validated their formula, called the Winnetka formula, against 700 books that had been named by at least 25 out of almost 37,000 children as ones they had read and *liked*. They also had the mean reading scores of the children, which they used as a difficulty measure in developing their formula. Their new formula correlated highly ($r = .845$) with the reading test scores.

With this formula, investigators knew that they could objectively match the grade level of a text with the reading ability of the reader. The match was not perfect, but it was better than subjective judgments. The Winnetka formula, the first one to predict difficulty by grade levels, became the prototype of modern readability formulas.

A Word about Correlations

In reading research, investigators look for **correlations** instead of **causes**. A correlation coefficient ($r =$) is a descriptive statistic that can go from +1.00 to 0.0 or from 0.0 to -1.00. Both +1.00 and -1.00 represent a perfect correlation, depending on whether the elements are positively or negatively correlated.

A coefficient of 1.00 shows that, as one element changes, the other element changes in the same (+) or opposite (-) direction by a corresponding amount. A coefficient of .00 means no correlation, that is, no corre-

¹ Vogel, M. and Washburne, C. 1928. "An Objective Method of Determining Grade Placement of Children's Reading Material." *The Elementary School Journal*, Vol. 28, pp. 373-381.

sponding relationship through a series of changes.

For example, if a formula should predict a 9th-grade level of difficulty on a 7th-grade text, and, if at all grade levels, the error is in the same direction and by a corresponding amount, the correlation could be +1.00 or at least quite high. If, on the other hand, a formula predicts a 9th-grade level for a 6th-grade text, an 8th grade level for a 10th-grade text, and has similar variability in both directions, the correlation would be very low, or even 0.00.

Squaring the correlation coefficient ($r^2 =$) gives the percentage of accountability for the variance. For example, the Vogel and Washburne formula above accounts for 71% ($= .845^2$) of the variance of the text difficulty.

—WHD

Elementary School Journal

Vol. XXVIII

January, 1928

An Objective Method of Determining Grade Placement of Children's Reading Material

MABEL VOGEL

Director of Research Office, Public Schools, Winnetka, Illinois

CARLETON WASHBURNE

Superintendent of Schools, Winnetka, Illinois

EVERY TEACHER has to face the problem of fitting reading material to children's reading ability. Any attempt in the past to give children suitable material has been largely a matter of guesswork. An objective method of determining what material is appropriate for children of given reading ability is needed by classroom teachers. Similarly, in selecting textbooks and supplementary-reading material, the superintendent or supervisor should have a means of knowing whether the books are within the reading grasp of the children for whom they are intended. The writers of textbooks and other books for children need to have an objective method of determining whether their vocabulary and sentence structure are such as will offer no serious obstacles to the children who are to read what they write.

Two years ago the foundation was laid for a study of the objectively measurable differences that exist among books read and enjoyed by children of various levels of reading ability. Thirty-six thousand seven hundred and fifty widely scattered children reported on all the books which they had read during the preceding year. The ballots which they filled out were brought together, and the results of their judgments make up the *Winnetka Graded Book List*.¹

The *Winnetka Graded Book List* is a list of seven hundred books on

¹ Carleton Washburne and Mabel Vogel, *Winnetka Graded Book List*, Chicago: American Library Association, 1926.

each of which twenty-five or more children's judgments were received. This list is graded not according to the actual school grade of the children but according to the grade to which their reading ability corresponds. The paragraph-meaning section of the Stanford Achievement Test was used as a measure of silent-reading ability. The reliability and the validity of the grading of the books in the *Winnetka Graded Book List* have been amply checked. Since the grade placement of these seven hundred books is known, it is possible to use this information in determining the grade placement of other books.

One hundred and fifty-two books were chosen from the *Winnetka Graded Book List* as a basis for the present study. About half of them are the most popular books in the various grades. The other half are books well liked by both sexes and read by an equal number of boys and girls closely concentrated around the median in reading ability. This number of books was chosen arbitrarily as a fairly good representation of the books in the list.

The Winnetka teachers' seminar, composed of twenty volunteer teachers, examined these books for every conceivable element of difficulty which might influence the grade placement. Examination was made of the following elements:

1. Vocabulary difficulty (according to Pressey's technique²)
 - a) Number of different words occurring in a sampling of 1,000 words
 - b) Median index number (based on Thorndike's indexed word list³) of 1,000-word sampling
 - c) Number of words in 1,000-word sampling not occurring in Thorndike's list
2. Sentence structure of seventy-five sample sentences
 - a) Sentence use—declarative, exclamatory, imperative, and interrogative

² Bertha A. Lively and S. L. Pressey, "A Method for Measuring the "Vocabulary Burden" of Textbooks," *Educational Administration and Supervision*, IX (October, 1923), 389-98.

³ Edward L. Thorndike, *The Teacher's Word Book*, New York: Teachers College, Columbia University, 1921.

- b) Sentence form—simple, complex, compound, and complex-compound
 - c) Dependent clauses—noun, adjective, and adverbial
 - d) Phrases—adjective, adverbial, infinitive, and participial
3. Parts of speech occurring in 1,000-word sampling— nouns (common and proper, abstract and concrete), pronouns, verbs (action and non-action, transitive and intransitive), infinitives, participles, gerunds, adjectives (sensory and non-sensory), articles, adverbs, prepositions, conjunctions, interjections, and expletives
 4. Paragraph construction
 - a) Number of sentences and words per paragraph of conversation and non-conversation
 - b) Percentage of seventy-five sentences containing conversation
 5. General structure
 - a) Number of words to a line, number of lines to a book, and number of words to a book.
 - b) Length of chapters
 6. Physical makeup⁴
 - a) Weight
 - b) Size of type
 - c) Length of line
 - d) Distance between lines

After all the elements were tabulated and counted for each book, each element was graphed to determine whether there was a definite rise or fall from grade to grade. Those elements showing the most definite rise or fall from grade to grade were chosen for further study, and the others were cast aside.

Table I⁵ shows the correlation between each of the elements selected

⁴ The items under this heading were not fully explored as it was felt that a different technique of study was needed to determine optimum size of type, length of line; and leading appropriate to each grade. A separate study of these elements is being made.

⁵ This is one of a number of similar tables constructed during the study. One table showed correlations with the reading grade instead of the reading score; another showed correlations with chronological age; etc. More than one hundred coefficients of correlation were found. The most satisfactory correlations were with the reading score, as shown in Table I.

for further study and the median reading score of the children who read the 152 books. Table II shows the intercorrelations of the ten most promising elements in Table I. The aim in choosing these ten elements was to find elements which would correlate as little as possible with one another and as highly as possible with the median reading score of the children who read and enjoyed the books measured.

TABLE I
CORRELATION OF VARIOUS ELEMENTS WITH MEDIAN READING SCORE*

Element	Correlation
Number of different words occurring in a sampling of 1,000 words	.770
Median index number (based on Thorndike's list) of 1,000-word sampling	—
Number of words in 1,000-word sampling not occurring in Thorndike's list	.674
Number of words in book	.592
Number of phrases in 1,000-word sampling	.576
Number of verbs in 1,000-word sampling.	—
Number of words per paragraph	.527
Number of prepositions in 1,000-word sampling	.518
Number of phrases of all kinds in 75 sample sentences	.518
Number of phrases and clauses of all kinds in 75 sample sentences	.474
Number of adverbial phrases and clauses in 75 sample sentences	.467
Number of adverbial phrases and clauses in 1,000-word sampling	.463
Number of adjective phrases and clauses in 75 sample sentences	.461
Number of adverbial phrases in 75 sample sentences	.458
Number of words in 75 sample sentences	.453
Number of simple sentences in 75 sample sentences	—
Number of conjunctions in 1,000-word sampling	.371
Number of adverbial clauses in 75 sample sentences	.296
Number of nouns in 1,000-word sampling	—
	.262

*Because of the difference in sentence length, the number of words in seventy-five sample sentences varied greatly. To reduce phrase and clause counts to a common basis in certain cases, the number of phrases or clauses was divided by the number of words in seventy-five sentences and the quotient multiplied by 1,000. The result showed the number of phrases or clauses there would be in seventy-five sentences if these sentences contained exactly 1,000 words. This procedure was used only when it yielded a better correlation than did a simple phrase or clause count of seventy-five sentences. It was not used in the case of any of the four elements that make up the final regression equation.

Various combinations of the ten elements shown in Table II were tried and a series of multiple correlations found. The best multiple correlation

(.845), combining four elements, was made the basis of a regression equation which predicts with a high degree of reliability the reading score necessary for the reading and understanding of any given book. The standard error of estimate in using this equation is 8 points on the paragraph-meaning section of the Standard Achievement Test.

TABLE II
INTERCORRELATION OF VARIOUS ELEMENTS

	Differ-ent Words in 1,000	Prepo-sitions in 1,000 Words	Verbs in 1,000 Words	Words per Paragraph	Words in 75 Sen-tences	Simple Sen-tences in 75	Uncom-mon Words	Ad-verbial Clauses in 75 Sen-tences	Nouns in 1,000 Words
Median read-ing score	.770	.518	.527	.518	.453	-.371	.674	.291	-.262
Different words in 1,000546	-.572	.516	.442	-.306	.692	.308	-.177
Prepositions in 1,000 words	-.777	.462	.398	-.134	.412	.131	.002
Verbs in 1,000 words	-.517	-.543	.285	-.431	-.192	.017
Words per para-graph706	-.503	.322	.565	-.356
Words in 75 sen-tences	-.741	.244	.818	-.399
Simple sen-tences in 75106	.674	.552
Uncommon words in 1,000074	.069
Adverbial clauses in 75 sentences	-.424

This means a difference of less than a grade in the lower grades and a difference of slightly more than a grade in the upper grades. This is a very reasonable standard error since it was found that any book that was read and enjoyed by children in a given grade could be read and enjoyed by children one grade above or below.

The elements which have a multiple correlation of .845 are as follows: number of different words occurring in a sampling of 1,000 (X_2), number of prepositions (including duplicates) occurring in 1,000-word sampling (X_3), number of words (including duplicates) in 1,000-word sampling not

occurring in Thorndike's list (X_4), and of simple sentences in 75 sample sentences (X_5).

By making a count of these elements, any teacher can determine the grade placement of any book. The technique used is as follows:

1. Make a sampling of 1,000 words from the book as follows:
 - a) Determine the number of pages in the book.
 - b) Determine the number of words per line by counting the number of words in ten lines scattered through the book and dividing by 10.
 - c) Divide 1,000 (the number of words needed) by the number of words per line. For example, if there are eight words per line, Item c will be 1,000 divided by 8, or 125, the number of pages from which sample lines are to be chosen.
 - d) Divide the number of pages in the book (for example, 432) by the number of pages from which samples are to be chosen (for example, 125). In the example given, the quotient is 3.5. Therefore, the sample lines will be taken from every third page.
 - e) Copy on a separate card (cards cut 2 inches by 3 inches are a convenient size) every word from the top line (or any other given line) of every page to be sampled. Put a *p* in the corner of each card containing a word used as a preposition.
 - f) After copying the words from a given line on the number of pages estimated in *c*, count the cards. If there is not an even thousand, discard any excess, or add cards by copying words from additional lines until an exact thousand is reached.
 - g) Arrange the cards in strictly alphabetical order so that all duplicates of any given word come together. Eliminate all duplicate cards, writing the total number of such cards on the one card that remains. For example, if there are thirty cards containing the word "the," write the number 30 on one "the" card and discard the other cards containing this word.
2. Count the cards after the duplicates have been eliminated, thus obtaining the number of different words in 1,000. Call this number X_2 .

3. Count the total number of prepositions in the 1,000 words. If the preposition "in," for example, occurs fifteen times, it should count as fifteen prepositions. Record the total number of prepositions as X_3 .
4. Check each word card with Thorndike's word list. Count the total number of words, including duplicates, which do not count in Thorndike's list. In this connection it must be remembered that derived forms of words included in the Thorndike list are considered as being themselves included in the list. For example, the word "sing" occurs in the Thorndike list. The word "singing" would be counted as being included in the Thorndike list although it will not be found there in this form. Thorndike's introduction to his word list should be carefully read to determine which derived forms he has not included. Record the total number of words not included in Thorndike's list as X_4 .
5. Make a sampling of seventy-five sentences from the book as follows:
 - a) Count the total number of pages in the book, excluding picture pages.
 - b) Divide the number of pages in the book by 75 to determine which pages must be chosen. For example, if there are 150 pages in the book, a sentence should be taken from every other page to make up the 75 needed sentences. If there are 250 pages in the book, a sentence should be taken from every third page.
 - c) Tabulate as simple or not simple the first complete sentence on every page to be sampled. A simple sentence is defined as one in which there are no dependent or co-ordinate clauses; it contains only one subject and one predicate.
6. Count the number of simple sentences in the 75 sentences sampled. Record this number as X_5 .
7. Apply the following regression equation to the data, X_1 being the reading score, X_2 , the number of different words in 1,000; X_3 , the number of prepositions in 1,000 words; X_4 , the number of uncommon words in 1,000, and X_5 , the number of simple sentences in 75:

$$X_1 = .085X_2 + .101X_3 + .604X_4 - .411X_5 + 17.43$$

The answer to the equation score will be the score on the paragraph-meaning section of the Stanford Achievement Test necessary for reading the book measured. The reading score may be translated into reading grade according to Table III.

TABLE III

GRADE STANDARDS—PARAGRAPH-MEANING SECTION
OF THE STANFORD ACHIEVEMENT TEST

Score.....	Grade
4–16.....	II
18–34.....	III
36–52.....	IV
54–62.....	V
64–70.....	VI
72–78.....	VII
80–86.....	VIII
88–94.....	IX
96–102.....	X
104–112.....	XI

Let us take *The Japanese Empire*⁶ by Harry A. Franck as an example of the application of the equation.

$$\text{Number of different words in 1,000} = 445$$

$$\text{Number of prepositions in 1,000 words} = 117$$

$$\text{Number of uncommon words in 1,000} = 22$$

$$\text{Number of simple sentences in 75} = 20$$

$$X_1 = (.085)(445) + (.101)(117) + (.604)(22) - (.411)(20) + 17.43$$

The reading score necessary for the ready comprehension of the book is 72.14. As can be seen from Table III, this book is suitable for children whose reading ability is that of the average child at the beginning of Grade VII.

Any book for use in the elementary grades may be similarly analyzed. It is therefore possible to determine the correct grade placement for any book so far as structural difficulty is concerned. When books are so graded and children's reading ability is measured, it is possible to give chil-

⁶ Harry A. Franck, *The Japanese Empire: A Geographical Reader*. Dansville, New York: F. A. Owen Publishing Co., 1927.

dren books which fit their ability. Furthermore, in writing a book for children in a given grade, an author can check his writing by the regression equation and simplify it if necessary. For the latter purpose Table IV will be found helpful. This table gives the medians and upper and lower quartiles, grade by grade, for each of the four elements measured. If, through the use of the regression equation, the author finds that his material is too difficult for the grade in which it is to be used, he can compare the word and sentence counts with Table IV and see which elements need simplification.

TABLE IV
STANDARDS FOR USE IN WRITING BOOKS FOR CHILDREN

	Grades					
	III	IVs	V	VI	VII	VIII
Number of different words in 1,000:						
Upper quartile	338	401	417	435.5	457.5	460.8
Median	316	377	407.5	416.5	440.5	458
Lower quartile	258	329	386	397.5	411	447
Number of prepositions in 1,000 words:						
Upper quartile	79	106	110	114	116	123
Median	71.3	96	100	107	99	115.5
Lower quartile	63	79	83	100	93	101
Number of uncommon words in 1,000:						
Upper quartile	8	14	20	24.5	34.5	40
Median	6	11	14.5	19.5	26	32.5
Lower quartile	2	6	12	17	18.5	28.5
Number of simple sentences in 75:						
Upper quartile	49	42.5	26	25	29	34
Median	39	30	21.5	19	22.5	26
Lower quartile	36	22	18	11	18	21.5

Since reading is the most basic of all school subjects and giving children material which is too difficult in structure tends toward wrong methods of visual perception, lack of interest, and faulty understanding and is responsible for many school failures, an objective method of measuring the structural difficulty of reading matter for children in the elementary grades is of primary importance.

The present study deals only with structure. A similar study dealing with content is well under way.

1931—Waples and Tyler: What Adults Want to Read

DURING the Depression in the '30s, adult education and the increased use of libraries stimulated studies in reading. Sociologists studied "who reads what and why over consecutive periods," looking at reading as an aspect of mass communication.

In 1931, Douglas Waples and Ralph W. Tyler published *What People Want to Read About*, a comprehensive, two-year study of adult reading interests. Instead of using the traditional library circulation records to determine reading patterns, they interviewed people divided by sex and occupation into 107 different groups. It showed the types and styles of materials that people not only read but also want to read. It also studied what they did not read and why.

They found that the reading of many people is limited because of the lack of suitable material. Readers often like to expand their knowledge, but the reading materials in which they are interested are too difficult.

This study confirmed what many studies before and since have found, that people gravitate towards that type of literature that matches their level of reading ability.

—WHD

1931—Patty and Painter: The Vocabulary Burden of Textbooks

Introduction

IN 1931, W. W. Patty and W. I. Painter¹ discovered the year of highest vocabulary burden in high school is the sophomore year. Believing that the length of a text affects the vocabulary burden, they questioned the Lively and Pressey method of sampling 1,000-word passages from a text,

They believed that taking a percentage of words from each text would give a better sample. The new method they devised took the words from the third line of each fifth page.

Their formula determined the relative difficulty of textbooks using a combination of frequency as determined by the Thorndike list and vocabulary diversity (the number of different words in a text).

—WHD

¹ Patty, W. W. and W. I. Painter. 1931. "A Technique for Measuring the Vocabulary Burden of Textbooks." *Journal of Educational Research*, Vol. 24, pp. 127-134.

Journal of Educational Research

Vol. XXIV

September, 1931

No. 2

A Technique for Measuring the Vocabulary Burden of Textbooks

W. W. PATTY AND W. I. PAINTER

Indiana University

AMONG the important factors that should be considered when selecting textbook for high-school use is that of vocabulary burden. It seems evident that ease of reading and understanding the words of a textbook is an important index to its learning difficulty. A few studies have been made previously in this field.

LIMITATIONS OF SOME VOCABULARY STUDIES

Lively and Pressey¹ selected one thousand word units from each of fourteen different types of reading material. These words were then assigned the values given in Thorndike's *Teachers' Word Book*² and comparison was made on the basis of total values for each type. Their method seems to be very effective for the comparison of the vocabulary difficulty

¹ Lively, Bertha A. and Pressey, S. L., "A Method for Measuring the Vocabulary Burden of Textbooks," *Educational Administration and Supervision*, IX (October, 1923), 389-98.

² Thorndike's *The Teachers' Word Book* is an alphabetical list of ten thousand words which were found to occur most widely in a count of about 625,000 words from literature for children; about 8,000,000 words from the Bible and English classics; about 300,000 words from elementary school textbooks; about 50,000 words from books about cooking, sowing, farming, the trades, and the like; about 90,000 words from daily newspapers and about 500,000 words from correspondence. Forty-one different sources were used. A measure of the frequency of each word's occurrence is given by the credit-number following it. If this credit number is 49 or over, it means that the word is in the first 1,000 for importance. A credit-number of from 29 to 48 places it in the second 1,000. A credit-number from 19 to 28 places it in the third 1,000, and a credit-number of 14 to 18 places it in the fourth 1,000. A second column is given in the *Word Book* which indicates by number the thousand in which the first five thousand fall. It also indicates by the letters a and b whether the word is in the first or second half of that thousand.

of texts or of the vocabulary burden of units equal in length. It fails, however, to consider the extra burden imposed on the reader by additional length.

Dolch³ makes a very good point in working out a ratio between the number of different words to the number of total words in a book so as to take into consideration word repetition. He suggests that the book having the wider range of vocabulary is certain to possess a vocabulary farther from the region of everyday language. This is undoubtedly true in the majority of cases, but it would not seem to be necessarily an absolute fact.

Ward⁴ took the total count of words in a section of a text and compared it with the Lively and Pressey method. He found a wider range of words in the total count and assumed that a thousand word count would not be a sufficient measure of any book. Since he has attempted to compare the results of two different techniques, it does not seem that he has proved his point. The average difficulty of words within samples would seem more desirable as a basis for comparing the vocabulary difficulty of one book with that of another than would the extreme ranges of words within the whole texts.

It is not to be expected that, in any sampling method, we will arrive at results which we can set up as a fixed standard. Instead, we should only expect to arrive at results on various texts, which, when compared, would bear the same ratio as would a comparison of burdens of the entire vocabularies of these texts.

A SUGGESTED TECHNIQUE

The following is suggested as a desirable technique for measuring the factor of vocabulary burden of high-school textbooks written in the English language. This technique was developed and used in a research project at Indiana University, in which all state-adopted texts for Indiana, with the exception of foreign language texts, were measured.

³ Dolch, Edward William. "Vocabulary Burden," *Journal of Educational Research*, XVII (March, 1928), 170.

⁴ Ward, J. L. "Measuring Vocabulary Burden," *American School Board Journal*, LXXI, page 98.

LENGTH OF TEXT MUST BE CONSIDERED

Word samples were taken from each of the texts studied. Since some of the books are considerably longer than others, it did not seem quite fair to compare their reading burden by selecting a definite unit of words from each book as a sample. The length of a book would undoubtedly affect the vocabulary burden of that book as compared with other books of different lengths. Where the difficulty of the average word is approximately equal, it, at least, is a greater task to read a long book than it is to read a short one. Also, by taking a definite unit, we would not take into consideration the proportional repetition of these words in any other unit similarly selected from the same text.

Thus, the different words in a one-thousand-word unit from a short text bear a lower ratio to the total one thousand than the different words in the entire text might bear to the total words in it. In other words, the longer the text the greater is the probability of its having a high percentage of word repetition. By taking a definite percent sample from each text, the ratio of different words to total words in the sample would be more nearly representative of the actual ratio of different words to total words in the entire book. It, therefore, seems that a proportionate word sample is the only valid sampling basis for comparing texts of unequal length and that the results so derived are more reliable than where a definite unit is used.

SELECTION OF THE SAMPLE

In order to get what was considered a fair proportionate word sample each book, the words from the third line of each fifth page were listed. If this were not a full line, the next full line was taken except in cases where the previous five pages were partly given over to pictures, graphs, diagrams, etc. If this piece of line seemed to be proportionately comparable to the amount of printed matter on those pages, it was used. When the fifth page was given over entirely to non-printed material, then the next printed page was sampled; the regular order, however, was resumed in taking the succeeding samples. These words were then tabulated alphabetically and their frequency numbered as they appeared in each sample. The number of words then in each list was the number of different words in each sample. This was called the range of the words.

USING THE TEACHERS' WORD BOOK

Each of these different words was then looked up in Thorndike's *Teachers' Word Book*. The values that Thorndike had estimated were set down opposite each word in column arrangement, as is shown in the sample following.

The figures in the column at the left indicate the frequency of the word as found in this word sample. The figures in the first column to the right indicate the word value or the Thorndike-index number. Where there are figures in the next column the first figure to the left represents the ranking of a thousand of the most common ten thousand words to which this particular word belongs; the letters *a* and *b* indicate whether it belongs to the first or the second half of this thousand; and the number at the right, when such number is found in this column, indicates the quarter of that thousand into which the word falls.

SAMPLE PAGE OF WORD SAMPLE TABULATION

F.	Word	T.I.N.	Position of in	
			thousand	W.V.
14	are	181	1a1	2534
1	accounts	63	1b	63
41	and	210	1a1	8610
61	a	208	1a1	12688
1	against	114	1a1	114
1	Archimedes	0		0
1	acetylene	0		0
19	as	204	1a1	2652
8	at	203	1a1	1624
2	atmosphere	11	5b	22
5	air	91	1a4	455
1	also	119	1a2	119
8	another	116	1a2	348
1	attached	20	3b	20
1	apparatus	7		7
2	arrange	35	2b	70
1	away	125	1a2	125
1	absorbed	8		8
1	aids	47	2a	47
1	along	99	1a3	99
1	animals	70	1b	70
1	act	70	1b	70

A WORD WEIGHTED VALUE

In the last column we find what we have termed a weighted value of each word. This weighted value is a product of the Thorndike-index-number and the frequency of the word in that particular sample. We might represent this by the following formula: $W.V. = T.I.N. \times F.$ In this formula, $W.V.$ represents the weighted value; $T.I.N.$, the Thorndike-index-number and $F.$, the frequency. The purpose of calculating such a weighted value is to take care of word repetition within, the sample, permitting each word to be considered in proportion to the frequency of its use.

DIFFICULTY VARIES INVERSELY WITH RECORDED VALUES

It must be borne in mind that Thorndike's index numbers were based largely on the frequency of the use of the words included in his *Word Book*; the higher the value which he places on a word, the more commonly that word is used in everyday language. The lower values found throughout this study, then, indicate a greater difficulty, or a greater vocabulary burden; that is, the difficulty or burden varies inversely with the values recorded.

COMPARISON OF AVERAGE WORD DIFFICULTY

The next step undertaken was to calculate an average-word-weighted-value so that it would be possible to compare the average word difficulty of one book with that of another as accurately as this is possible by a sampling method. This was done by totaling the weighted values for all words of a sample and dividing the result by the total number of words in the sample. This might be represented by the formula: $A.W.W.V. = T.W.V. \div T.W.S.$ In this formula, $A.W.W.V.$ represents the average-word-related-value; $T.W.V.$, the total weighted values; and $T.W.S.$, the total words in the sample. It can be readily seen that this average-word-weighted-value is merely the arithmetic mean of the Thorndike-index-numbers for all -words of the sample.

INCLUSION OF THE RANGE OF THE WORDS

This average-word-weighted-value, however, does not seem to be quite a fair measure for the total burden of one book as compared with that of another, since it does not take into consideration the relation of the num-

ber of different words, or range, to the total number of words in the sample. We would riot know whether the A.W.W.Y. was the result of a few words used a number of times, or whether it was the result of a number of different words of about the same degree of difficulty used only a few times each. The reader will undoubtedly concede that the latter would be the greater burden, and that it would be especially noticeable in reading a long book of such proportionate range.

In order to take this range of words into consideration, an index number was found by dividing the average-word-weighted-value by the range of the words within the sample. This gives the ratio of the different words to the difficulty of the average word, as is shown in the following example:

Text	T.W.S.	R.	T.W.V.	A.W.W.V.	I.N.
Book I	646	379	77.045	119.26	.315
Book II	1051	433	125.289	119.20	.275

In this example, T.W.S. equals total words in the sample; R. equals the range; T.W.V. equals the total weighted value; A.W.W.V. equals the average word weighted value; and I.N. equals the index number. It can be readily seen that, in the sample given, the A.W.W.V. is approximately equal in both books, but that Book II is considerably longer and has a larger range of words; consequently its reading burden would be heavier. By dividing our A.W.W.V.'s by their respective ranges we reduce these books, which have approximately an equal A.W.W.V., to index numbers which include not only the weighted of the words, but also the total words in the sample and the range of these words, varying inversely with the latter; that is, the larger the range the smaller is the index number, indicating a greater reading burden. We might develop this into a formula, as follows:

$$A.W.W.V. = T.W.V. \div T.W.S.$$

$$I.N. = (T.W.V. \div T.W.S.) \div R \text{ or}$$

$$I.N. = (T.W.V. \div T.W.S.) \times (1 \div R) \text{ or}$$

$$I.N. = T.W.V. \div (T.W.S. \times R)$$

This same process should give satisfactory results in all cases.

A 8UB-CLASSIFICATION OF THE SAMPLE

Each sample list was next gone over, and the number of words found in each of the first three thousand was totaled, as well as was the number of those found between the third and tenth thousand, and of those which were not found at all in Thorndike's most common ten thousand words. The last group were called the zero value words, since there were no values listed for them. Each of these numbers was then divided by the total number of words in the sample in which it was found in order to determine the percent which each was of the total. This can be used to compare the difficulty of the words in each book according to their distribution. It will also help to clarify the meaning of the index number which was first determined as well as to serve as a check on it.

DETERMINING THE MOST DIFFICULT YEAR OF SCHOOL

Another phase of the procedure was the sending out of a questionnaire to one hundred commissioned high schools in the state of Indiana. These schools were selected by taking every eighth commissioned high school, regardless of size, from the Indiana state school directory. The apparent range was from the smallest to the largest commissioned high schools of the state.

The final form of the questionnaire resolved itself into a check list containing the names of the forty-three different texts adopted by the state, and space for the of those used within the school and for checking the year or years of school in which each was used. Replies were received from seventy percent of the schools. A tabulation of the results gave us the frequencies of the use of each text in each year of high, school. By comparing this with our index number, as worked out in the first division of our procedure, we were able to find out in which year the texts having the greatest vocabulary burden were used.

In order that we might more objectively compare the difficulty of one year of high school with that of another, we worked out an average index number, or an arithmetic mean of the index numbers, of the total state adopted texts used in any one year. First, we multiplied the frequency that a text was used in each year by the index number for that text, deriving what we called a frequency-index-product for that text for each year used. The frequency-index-products for each year, as well as the frequencies of the use of state adopted texts in each year, were then totaled.

Unlocking Language

SAMPLE OP TEXTBOOK CHECK LIST

Year of H.S.
1 2 3 4

Our English	
Written and Spoken	
English Literature.....	Clippinger
Outlines of English Lit.....	Long
Outlines of English Lit. readings	Long
American Literature	Long
Outlines of American Lit	Long
Outline of American Lit. Readings.....	Long
Early American History, Rev.....	Webster
Modern European History, Rev.	Webster
Modern Times and Living Past.....	Elson
History of the U.S.	Beard, etc.
Government in the U.S.	Smith, etc.
Elementary Principles of Chem.	Brownlee, etc.
First Book in Chem., 1928 ed	Bradbury
Chemistry and Its Uses	McPherson, etc.
Beginners Chem. and Its Uses	Irwin, etc.
Elem. Prin. of Physics.....	Fuller, etc.
Elements of Physics	Millikan, etc.
Essentials of Mod. Physics	Dull
Physics in Everyday Life	Henderson
Commerce and Industry.....	Smith
High-School Geography	Whitbeck
First Course in Algebra.....	Nyberg
Second Course in Algebra	Nyberg
Modern Plane Geometry	Clark, etc.
Modern Solid Geometry	Clark, etc.
Arithmetic of Business.....	Smith
Applied Arithmetic	Smith
Farm Projects and Problems	Davis
Animal Husbandry	Harper
Soils and Crops	Mosier
Studies in Horticulture	Lloyd
Engineering on the Farm.....	Stewart
Practical Botany (Agri.)	Bergen
Plant Life and Plant Uses.....	Coulter
Elem. Studies in Botany.....	Coulter
First Course in Botany	Pool, etc.
Textbook in Botany.....	Allen, etc.
Animal Studies.....	Jordan, etc.
General Zoology	Linville, etc.
Healthful Living.....	Williams
Practical Zoology	Hegner

Please be especially careful to check whether you are using the text or would use it, if you were not using an old adoption, and whether you are or would use it in years 1, 2, 3, or 4.

Finally these four total-frequency-index-products were divided by the total of the frequencies for their respective years. The result was an average-index-number which indicated the average vocabulary burden for each year of high school. We might also work this out as a formula, as follows:

B.F. \times I.N. equals F.I.P., in which B.F. represents the book frequency; I.N., the index number; and F.I.P., the frequency index product.

A.I.N. equals T.F.I.P. \div N. in which A.I.N. represents the average index number or the arithmetic mean of the index numbers; T.F.I.P., the total frequency index products; and N., the total number of book frequencies in each year.

It is recognized that the measurement of vocabulary burden of high-school textbooks is only one of several desirable devices for ascertaining their relative suitability for class use. The foregoing method is presented only as an apparent improvement in technique in one phase of measurement of the quality of the texts. Methods of equal or better quality should be developed for evaluating other features that affect the worth of books as aids to learning.

1934—Dale and Tyler: Difficulty of Adult Reading Materials

Introduction

AFTER WORKING with Waples, Ralph Tyler became interested in adults of limited reading ability. He joined with Edgar Dale to publish in 1934¹ their own readability formula and the first study on adult readability formulas.

Most of the research up until that time focused on crucial task of providing students with materials that match their reading levels. Teachers had found that without materials that match both the interests and reading levels of students, they will fail to catch fire and take reading seriously. The Dale and Tyler study reflects a similar urgency in matching materials with adult levels of reading ability. Researchers had discovered that the average reader in the U.S. was an adult of limited reading ability.

Dale had found problems with the Thorndike *Word Book* and started looking for better alternatives. The specific contribution of this study was the use of materials specifically designed for adults of limited reading ability.

Their criterion for developing the formula was 74 selections on personal health taken from magazines, newspapers, textbooks, and adaptations from children's health textbooks. They determined the difficulty of the passages with multiple-choice questions based on the texts given to adults of limited reading ability.

From the 29 factors that had been found significant for children's comprehension, they found ten that were significant for adults. They found that three of these factors correlated so highly with the other factors that they alone gave almost the same prediction as the combined ten. They were:

- Number of different technical words.
- Number of different hard non-technical words.
- Number of indeterminate clauses.

¹ Dale, E. and R. Tyler. 1934. "A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability." *The Library Quarterly*, 4:384-412.

They combined these three factors into a formula to predict the proportion of adult readers of limited reading ability who would be able to understand the material. The formula correlated .511 with difficulty as measured by multiple-choice reading tests based on the 74 criterion selections.

This study and the Ojemann study that follows are the classic studies on how to develop materials for disadvantaged adults of low reading skills.

—WHD

The Library Quarterly

Vol. IV

1934

A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability

EDGAR DALE AND RALPH W. TYLER

Bureau of Educational Research, Ohio State University

THE PROGRESS of adult education depends to a large degree upon the development of reading materials which are adapted to the abilities of adults who are limited in their skill in reading.

Present materials are inadequate in this respect. Librarians have often maintained that much of the material available for adults of limited reading ability is too difficult. Furthermore, there are no scientific techniques by means of which to make an accurate estimate of the reading difficulty of books and pamphlets on the library shelves. The inevitable result is the fact that adults of limited reading abilities do only a slight amount of reading of non-fiction materials. The inadequacy of appropriate materials constitutes one major cause of this deficiency.

This situation can be improved through the development of methods by which the easier reading materials can be identified and which would serve to guide writers in preparing materials which can be understood by adults of limited reading ability. In other words, investigations are needed to discover the characteristics within the reading materials themselves which affect their ease of comprehension. A critical analysis of the widely varying results of previous studies indicates the impossibility of determining the factors in the reading materials which make them understandable unless the investigations separate the influence of factors within the reading material from those outside. The reader's interest in the topic treated in the reading matter, his ability to read, the kind of comprehension appropriate to the purposes of the reading matter, and the difficulty of the ideas developed in the reading matter are all factors which greatly affect

his comprehension of the material read but are distinct from the characteristics involved in the materials themselves which may be changed so as to make these ideas understandable to adults of limited reading ability.

Some of the effects of these external factors are easily noted. For example, stenographers, reading on a topic like "marriage" in which they have evinced high interest, are more strongly motivated and put forth greater effort to comprehend than when they are reading about the international aspects of the gold standard. Long sentences usually provide greater difficulty for young children than for adults with more mature reading abilities. The characteristics of materials which tend to cause the reader to feel the rhythm of the writing, as in poetry, may be different from those which make it easier for the reader to note specific details, as in a scientific treatise. Writing which in itself is easy to follow may result in more meager comprehension when dealing with the difficult concepts of relativity than when dealing with the simple concepts of barter. The various factors not in the reading materials themselves must be controlled in order to determine the effects of factors within the materials.

The study which follows is intended to illustrate the possibility of determining the factors which influence the difficulty of reading materials when the topic treated, the group of readers concerned, and the purpose of the reading is held constant. This is only a single exploratory study. Obviously, a large number of similar investigations are necessary to provide more intelligent guidance in writing and selecting reading materials which will be comprehended by given groups of people. Furthermore, to check the results of this study, experimentation is needed in selecting more materials and in writing materials according to the specifications indicated to discover whether they are consistently useful in selecting and preparing reading matter which is more easily comprehended by adults of limited reading ability. Five steps were involved in this investigation:

1. Samples of reading materials dealing with personal health were selected for use in the experiment.
2. Tests were developed which would measure the ability of adults to comprehend these materials.
3. The materials and the tests were given to groups of colored adults of limited reading abilities.
4. Correlations were run between a series of factors contained in these selections and the index of difficulty as discovered for each selec-

tion by the use of the reading tests.

- 5 A multiple regression equation was developed by means of which the lay worker or librarian can estimate the difficulty of personal health materials not previously tested.

This study was limited to reading materials dealing with problems of personal health. In comparing the difficulty of various materials it was necessary to choose materials which dealt with the same topic since the interests of adult groups vary with the subject treated in the materials, and variations in interest are likely to affect the effort put forth by the reader, thus affecting his comprehension. Since differences in comprehension due to differences in interest are not inherent in the reading materials but are due to the reader's attitude toward the topic, it was important that materials be chosen which dealt with the same topic and that the subject be one of great interest to these adult groups so that their reading would be adequately motivated.¹ I have shown that the topic of personal health is of high interest to all adult groups. Furthermore, one of the writers had previously collected data on the technical vocabulary of literature on health which facilitated the analysis of materials dealing with this subject.

METHODS USED TO DETERMINE THE DIFFICULTY OF THE READING MATERIALS

Difficulty of reading materials is to be interpreted in this study as meaning that the average comprehension on the part of readers of such materials is low. An examination of a wide variety of health materials indicates that comprehension may conceivably be measured in terms of one or all of these possible major outcomes:

1. An emotional reaction, the nature and potency of which depends upon the difficulty or the interest of the article.
2. The ability to recall specific elements of the reading.
3. The ability to form general conclusions on the material read.

The emotional outcome has not been tested at all in this study, except in so far as we can assume that it is positively correlated with ability to comprehend the reading materials. There might be value in having the reader indicate his interest in the article after he has read it, and then de-

¹ D. Waples and R. W. Tyler, *What people want to read about* (University of Chicago Press, 1931), p. 286

termine the relationship existing between such interest and ability to comprehend the article satisfactorily. It might also be desirable to discover other emotional reactions of the reader and to note as before the relationship of these reactions to comprehension. No attempt, however, was made in this study to measure such outcomes.

It has been assumed that the most fruitful outcome for study in this investigation is the ability of adults to read various types of non-fiction materials and to abstract there from the general conclusion which may reasonably be drawn from the content. This type of reaction is more often demanded of the reader in the field of health than is the reaction to the specific elements. Further, since the ability to draw general conclusions depends in some measure upon ability to react satisfactorily to the specific elements, there is little doubt that some portion of the latter outcome has been tested by the technique used. It should further be noted that the terms "specific" and "general" are relative, not absolute. A reaction that is specific for one paragraph may be a general conclusion in some other paragraph.

SELECTION OF READING MATERIALS

In selecting materials to be tested certain criteria constituted the major considerations.

First, authoritative materials were chosen. It is manifestly undesirable to utilize reading materials which, because of their inaccuracy, might have harmful attendant outcomes. Second, the materials chosen were of a type easily accessible to adults. If a study were made of the reading difficulty of materials which are rarely found in current textbooks, books of non-fiction, newspapers, and magazines, the results would have slight value unless techniques were thereby developed which could be used to evaluate available non-fiction materials. The use of these two criteria led to the selection of a series of current syndicated health articles by Drs. Brady, Evans, and Copeland to be utilized in one phase of the study (see Appen. I).

Long articles were avoided, since there is evidence that adults of limited education prefer to do their non-fiction reading in short bits instead of long ones. As a matter of fact, most of the newspaper health articles examined were approximately four hundred words in length.

For other phases of the investigation selections dealing with health

were taken from textbooks and magazines. They represented a wide variety of authors and topics. Every variant which might possibly affect difficulty was sought in collecting the materials. The material selected was never less than one paragraph in length, and usually it was more. An attempt was made in all cases to get a unit of thought. Most of the material from the health columns of daily newspapers included the entire article.

The first series of tests showed that the material from the health columns was too difficult for our readers. Hence, an attempt was made to get simpler materials for a second test. Short excerpts were then chosen from the first of a series of elementary-school textbooks in health, a few from junior high school textbooks in health, and two items from books that were being used for Americanization work. We wished to test more of the material which is actually being used in adult education work, but the materials available on the topic of personal health were meager in amount and poor in quality. However, the textbook materials used in the second test were probably somewhat more appropriate for adults of limited schooling since they were prepared for children in the upper grades of the elementary school (see Appen. II).

In spite of these precautions, the reading materials for the second test also proved too difficult. In the next attempt, therefore, one of the writers prepared as easy materials in the field of health as it was possible for him to write with the present limited knowledge of the factors which made the materials difficult (see Appen. III). In general, the principles followed in writing the paragraphs for the third test were these:

1. Elimination as far as possible, of vocabulary difficulties both of technical and non-technical words.
2. Writing the materials in an informal style, making extensive use of anecdotes about persons, conversation, and answers to questions.
3. Elimination of material not contributing directly to the main ideas in the selection.

MAKING THE TESTS

The material in the first of the series of three tests was selected from the health columns of Drs. Brady, Copeland, and Evans. These articles are syndicated in a number of newspapers throughout the country. They represent one of the best and the cheapest opportunities afforded the adult of limited education to extend his range of information in the field of

health. By testing the reader's comprehension of such articles, it is possible to determine within the limits of the errors of sampling and measurement the extent to which such health columns can function as a device for adult education.

It was assumed that there was implicit or explicit in these units of material a generalization which would adequately summarize the article. It was also assumed that a good test of the difficulty of such materials for adults of limited education would be the success with which they were able to formulate such a conclusion. It was further assumed that skilful reading of such a unit of material involved the ability to tell what was not in the unit. In other words, a person who has a clear-cut notion of what is found in a unit of reading material ought to be equally well aware of what is not found there. It should be further pointed out that to have an individual make two separate reactions to the same paragraph greatly increased the reliability of the test results.

In order to have a finer measuring instrument it was assumed that some credit should be given to the adult who could draw a conclusion that was close to either of the two correct answers for each paragraph. Therefore, five choices were included for every test item: the best conclusion, the worst conclusion, and three others—all of them distributed as far as possible in equal intervals of difficulty.

A variety of methods might have been used for testing the ability of readers to draw satisfactory conclusions from the reading materials. Adults of limited reading ability could be asked to read the selections and then state or write out the best conclusions which they could draw from their reading. This technique was ruled out because it would involve expenditures beyond the scope of the budget, since it would necessitate the evaluation of a large number of responses by a group of judges. Furthermore, the ability to express a conclusion orally or in writing may not be a true measure of the ability to draw such a conclusion mentally.

The technique used was to provide five responses for each selection, from which the reader was to choose the one he thought was the best conclusion to the selection and the one he thought was the poorest. These responses were obtained by having two expert readers independently formulate what they believed to be the best and worst conclusions to each selection and then in conference to iron out minor differences in the statements. Their conclusions were submitted to a small group of adults

whose suggestions for changes were followed when they seemed pertinent. The three conclusions lying between these best and worst conclusions were formulated by one of the writers and his research assistant. In formulating these conclusions an attempt was made to formulate them in language of less difficulty than the language used in the reading selections.² This precaution was necessary so that the limiting factor in the readers' comprehension would be the difficulty of the reading material in the selection and not the difficulty of comprehending the meaning of the conclusions. In this study the writers assumed that they formulated conclusions in language easier to comprehend than the language used in the reading selections. Hence, only the reading selections themselves were analyzed for elements of difficulty and not the series of five conclusions which follow each selection.

ADMINISTERING THE TESTS

Investigators who wish to discover certain facts regarding the abilities of children can secure in the public school an unfailing source of subjects. They have but to convince the principal or teacher that their project promises to secure useful data, and entree is secured. *Rapport* with the pupils is easily established, and children usually do such tests with a great deal of zest. Moreover, there are writing desks upon which pupils can place their test blanks.

The investigator of the abilities of adults, however, is not so fortunate and must contend with a series of difficulties, not only in securing the subjects, but also in securing favorable mechanical conditions under which to give the test. With the exception of the comparatively small number of adults in night schools, he has no adults undergoing educational training to whom he can turn. Under normal conditions of employment, adults must be tested in the evening, since they are employed during the day. Female subjects can sometimes be secured through the co-operation of clubs and organizations, but certain difficulties arise here, as will be pointed out shortly. Further, adults of limited reading ability are diffident about exposing their ignorance. For example, the negroes who found themselves unable to read our tests usually explained the deficiency as follows: "Ah'm sorry, boss, but ah left mah glasses at home." Finally, when adult groups volunteer to take the test without pay, they sometimes take the test with such a lack of seriousness that doubt is cast on the valid-

² Samples of the test are given as Exhibits A, B, and C.

ity of the findings.

Most of the foregoing difficulties were faced by the investigators when they attempted to secure subjects for the tests. The night schools conducted under public-school auspices were not open during the summer when this investigation was being conducted. Co-operation was promised, however, should the investigation be extended into the fall and winter.

The following experience with a vocabulary test given to a group of women belonging to one of the clubs in the community house in a poor section of the city of Columbus illustrates one of the difficulties enumerated in the preceding paragraph. It was a volunteer group composed of native-born American women, almost none of whom had ever taken work in high school. A list of words was given them and they were to mark the words whose meaning they knew. They were then to take a multiple-choice test on these words. After the directions had been given and the group had begun the test, they conversed, glanced at each other's papers, looked upon the whole thing as a lark, and the only value of the final results was to suggest very strongly that such groups were not satisfactory research subjects.

Another method of securing subjects offered high promise, but proved inadequate in the one trial which we gave it. We believed that church groups composed largely of adults of limited education might be very willing to co-operate with us in consideration of a donation to the church funds. One colored pastor in Columbus promised us a group of thirty people, members of his Sunday-school class of young married people. When the day for testing arrived, the number had been reduced by the pastor to twelve. And, to cap the climax, when the appointed hour for testing arrived, the pastor forgot his appointment, took a nap instead, failed to wake up in time, and we were unable to give the tests.

Volunteer groups having failed as a source of subjects, we decided to try our luck at securing subjects by paying them for their time. One of the workers at the Godman Guild in Columbus, a community house in a district composed almost entirely of negroes, suggested that when indigent adults came to the Godman Guild once a week to apply for aid from the community fund, she would inform them to return the following Friday at 2:00 P.M. to take some reading tests. She also told them that they must be able to read and write, that they would take some reading tests, and that they would be paid twenty-five cents for so doing. The social worker at

this community house informed the investigators that the twenty-five cents was psychologically a good amount to offer and that a smaller amount, even for less time, would not be so satisfactory.

This method of securing subjects proved satisfactory. On only a limited notice we were able to assemble from thirty-five to sixty adults each time the tests were given. Voluntary, unpaid subjects would hardly have been willing to put in continuous work for one hour or more on such tests on July afternoons when the temperature was around 95° F. Further, our paid subjects came when it was convenient for the investigators to use them. Since they were being paid for what they were doing, they were under an obligation to follow instructions. This method is to be recommended, therefore, when sufficient funds are available to carry it out.

When the subjects first assembled, the Monroe Silent Reading Test was administered to determine their general level of reading ability. They were then given the first set of reading selections on health, including the test of their comprehension of these materials.

Tests 2 and 3 were administered with greater facility than Test 1. Each individual was met at the door and asked whether he could read and write and was also questioned concerning the extent of his school work. If the replies suggested strongly that he would fail in the test, he was not allowed to take it. Further, the name of each individual was secured at the door and his previous record on the Monroe Reading Test and his score on the test covering the first set of health materials were noted in order to determine whether he would be able to do the subsequent tests.

ANALYZING THE FACTORS CAUSING DIFFICULTY

The purpose of the study was: first, to secure a valid criterion of the reading difficulty of a series of paragraphs; then to make quantitative studies of the factors in the paragraphs which might be influencing the difficulty; and, finally, to run, a series of correlations between the factors and the index of difficulty. A number of factors were chosen, therefore, for quantitative evaluation. Each of the factors is here described.

1. *The technical vocabulary.*—Technical vocabulary is defined, for the purposes of this study, as those words having their only meaning in the field of health or having a specialized meaning when applied to this field. Technical words in health can be divided perhaps somewhat loosely into what one might call their structural and functional aspects. The structural

vocabulary consists primarily of nouns, such as "acid," "alcohol," "alkali," "mercury," and "iodine." The functional words connote activity, namely, "cohere," "decay," "sterilize," etc. Some words have both a structural and a functional aspect, since they can be noun in one form and verb in another.

It should be pointed out that the term "health" is very much like the term "education." Strictly speaking, there is no technical vocabulary in education or in health. There are only technical vocabularies in chemistry, zoology, bacteriology, and allied scientific fields.

A number of words that might be considered technical are eliminated from consideration as technical health terms, because we have chosen to label as "easy words" those 769 words which are common to the word list of the International Kindergarten Union, and the most frequent thousand words in Thorndike's *Teacher's word book*. The nature of these "easy" words will be commented on later in this section.

It seems plausible that there will be an increase in the difficulty of reading materials in health when the technical vocabulary is increased in amount or difficulty. The increase in amount of different words, where there is no increase in difficulty requiring the reader, as it does, to keep a wider variety of ideas in mind at one time, is likely to decrease his success in reading such materials. That an increase in the unfamiliarity of the technical terms is likely to increase the difficulty of understanding the reading materials in which such technical terms appear is patent. In certain cases this will be due to the fact that the experiences themselves are utterly foreign to the readers, and in others it will be due to the fact that the symbols used are not sufficiently well related to the experiences which the reader has already had. An example is the word "absorb." The equivalent, "soak up," would be well known to almost everyone. The term "alkali," however, could not be so easily explained. The technical vocabulary used in this study was obtained from a wide variety of sources. The major source was a study of the health vocabulary and health problems found in twenty-seven different books and textbooks on health. To it was added the technical vocabulary found in an analysis of thirty different articles selected from health columns in newspapers. This final composite list included more than two thousand terms. It was now reduced to approximately one thousand terms, excluding those terms not likely to occur in the health materials written for the layman.

At the same time that Test 2 was given, this list of one thousand terms was presented to the group. Each reader was asked to put an X in front of every word in the list which he knew. The subjects were told that it did not make any difference how many or how few words they checked. They were also told that the three examiners would go about the room as they were doing the test, and question them on the words which they checked as known. The examiners discovered only two persons who were "fudging." It was thought that adults of limited education might be inclined to exaggerate what they knew, but this random check-up did not disclose this tendency. The thirty-eight papers secured in this fashion were thrown into two random halves, and the number of persons in each half checking each word was tabulated separately. The two sets of numbers were then correlated, which gave a coefficient of 0.90. When the Spearman prophecy formula was applied, it was found that the estimated correlation of this group of thirty-eight papers with another thirty-eight papers from persons of similar abilities would yield a coefficient of 0.94. This means that these persons agree extremely well among themselves as to the words which they believe they know. It does not settle the question of validity, although this close agreement does suggest that the method is fairly valid. The studies of a number of investigators have shown that the judgment of the difficulty of a word will usually correlate 0.80 or better with actual tests of the difficulty of such words. For fifty of these health words the coefficient of correlation between the group judgment and the difficulty determined by a multiple-choice test was 0.79.

The percentage of persons "knowing" each word in the total list was now tabulated and became the index of the difficulty of the technical health words.

2. *Easy words*.—It would appear that one of the most important factors in the relative difficulty of reading materials would be the percentage of the total words which are "easy" words or the number of different "easy" words contained therein. Two reasons can be given for such an assumption: the more easy words, the more familiar the content is likely to be, and, conversely, the fewer hard words, the less unfamiliar the content will be.

A word was considered "easy" if it was common to the most frequent thousand words in Thorndike's *Teacher's word book* and the word list of the International Kindergarten Union. Seven hundred and sixty-nine words were found to be common to both lists. Since these words are

probably known by children before entering the first grade, and since they are among the thousand words most frequently used in general reading materials, it is possible to consider them as of zero difficulty to adults of limited reading ability. It should be pointed out that there are a few practical difficulties in the use of such a list. Many of the words in the Thorndike word list are homographs, e.g., words spelled alike but of different meaning. It was necessary, therefore, to note after certain of these 769 words the meaning which was assumed to be the easy one.

3. *Hard non-technical words.*—There are a number of words in reading materials on any technical subject which are hard although non-technical in their nature. It seems plausible that the number of running words of this kind and the number of different words of this nature might be one of the factors influencing the reading difficulty of a paragraph.

It was believed that Dale's study of the unfamiliarity of eight thousand common words for pupils in Grades IV, VI, and VII would be of assistance at this point. This study made use of the judgment of pupils in these grades as to the familiarity of words. To check the applicability of this list for use with adults there was included in the vocabulary test used with the colored adults slightly more than one hundred non-technical words, which were also in the Dale list. When the scores on these words for the adults were correlated with the corresponding scores for the children, the coefficients of correlation were as follows:

Adult scores with scores of fourth-grade pupils	0.42
Adult scores with scores of sixth-grade pupils	0.65
Adult scores with scores of eighth-grade pupils	0.64
Adult scores with average score of fourth- and sixth-grade pupils	0.59

Since the words known by adults correspond most closely with the words known by pupils in the sixth and eighth grades, a record was made of the non-technical words that were rated in the Dale list as difficult for sixth- and for eighth-grade pupils.

4. *Types and length of sentences used.*—It seems likely, when sentences are used which involve suspension of one's judgment as to the outcome until the entire sentence has been covered, that the difficulty would be increased. It would thus appear that length of sentence might correlate positively with difficulty and that complexity of the sentence as measured

by whether the sentence was complex or compound would also increase its difficulty over that of the simple sentence. In final analysis of the health selections so few complex-compound sentences were found that these were classed with the compound, thus leaving three categories: simple, complex, and compound or compound-complex sentences.

5. Number of clauses and prepositional phrases.—Inclusions of clauses and prepositional phrases in a paragraph of reading materials may increase the difficulty either by its effect on making the sentence longer or through increasing the number of ideas which must be kept in mind while one is trying to get the thought of a sentence. A count was made of the number of clauses and prepositional phrases appearing in each reading selection on health. A separate count was made of clauses beginning with what and whatever, where and wherever, when and whenever, in contexts which did not imply simple interrogation and did imply indefiniteness or indeterminateness of time or place or subject. This definition was adequate to enable two clerks, working independently, to obtain the same values when counting these clauses in the various selections.

6. Number of personal pronouns.—The hypothesis that the informality with which an article was written would increase the ease with which it could be read seemed a tenable one. Hence, a count was made of the number of pronouns of the first, second, or third person appearing in each health selection so that the relationship between such appearances and the difficulty of the article might be determined.

7. Number of monosyllabic words.—Several writers have maintained that there is a relation between the percentage of monosyllabic words in a passage and the reading difficulty of that passage. In this study, therefore, a count was made of the percentage so that the degree of relationship existing between this factor and the difficulty of the reading selections might be discovered.

8. Other factors affecting difficulty.—There are a number of other factors which were compared with the difficulty of these selections on health. One of these was the percentage of words beginning with *e*, which has been suggested by Lewerenz as a very good predictor of reading difficulty. However, these factors either had but slight relation to difficulty or else were closely related to other factors which had already been used. None of these factors added any appreciable amount to the size of the multiple correlation coefficient obtained from the factors previously de-

scribed.

ESTIMATING THE DIFFICULTY OF READING MATERIALS

The method used in determining the degree of comprehension with which adults of limited education read selections of differing difficulties has already been described. A group of persons of limited education were sampled by means of this method and a record made of the average degree of comprehension with which the group read each of seventy-four selections dealing with health problems. Although the group selected for study was not large enough to be a completely adequate sample of similar populations at large, the scores by which the difficulty of the selections was determined have a reliability of 0.75 as computed by splitting the entire group of persons into two random halves, calculating the scores on the selections for the two half-groups separately, correlating the two sets of scores, and predicting by the Spearman prophecy formula the probable correlation between two groups, each of which is the size of the entire group sampled. This reliability is high enough for the scores to furnish an index of relative difficulty of the selection, but the study will need to be supplemented in the future by a larger and more inclusive sampling of adults of limited education.

After the average score for each selection had been computed, thus furnishing an index of its reading difficulty, it was possible to compare the difficulty of the selections with each of the several factors characterizing these reading materials which were described in the previous section. The several factors analyzed and the coefficient of correlation between each factor and the group score for the selection is shown in Table I. It is necessary to remember that the higher the score on a selection, the greater the group comprehension of the selection. Hence high scores indicate selections easily comprehended, and low scores selections difficult to comprehend. Thus a negative correlation coefficient means that the factor is positively correlated with difficulty while a positive correlation coefficient means that the factor is negatively correlated with difficulty of comprehension.

An examination of this table reveals the fact that there were many factors correlated with reading difficulty in these selections dealing with health. The factor most closely related to difficulty of comprehension was the *number of different technical words in the selection*. The factor next most closely related to difficulty was the *number of hard, non-technical*

words in the selection not known by go per cent of sixth-grade pupils. However, the *number of different hard, non-technical words in the selection* was almost as closely correlated with difficulty. It will be noted that some of the factors analyzed show practically no relation to difficulty of comprehension. If these selections are typical of the non-fiction reading material for adults, the latter factors do not greatly affect the difficulty of reading materials. Of the twenty-five factors shown in the table, ten are correlated with difficulty of comprehension to the degree represented by coefficients of 0.30 or above. All of these ten are easily computed for a given selection by any layman with the aid of the Dale vocabulary lists.

If all these ten factors are tabulated and combined to make the best prediction of difficulty possible by their use, the coefficient of correlation between the resulting prediction and the group scores for comprehension is 0.561. The meaning of this correlation may better be shown by illustration. There were twenty-two (30 per cent) of the seventy-four selections which were so well comprehended by the group of persons tested that they were understood by 60 per cent or more of the group.

By combining these ten factors and choosing the twenty-two selections which would rank least difficult with reference to these factors, one would find that he had selected sixteen of the twenty-two which were actually easiest to comprehend as shown by later test. Now since all of these selections were chosen by the investigators because they were thought to be easy enough to be read by adults of limited education, it is probable that if one were dependent upon his offhand judgment in attempting to select the 30 per cent which are easiest to comprehend, he would be likely to get only seven right out of twenty-two. Using these ten factors, therefore, materially improves the layman's ability to choose selections which will be easiest for adults of limited education to comprehend.

For most practical purposes ten factors are too many to use in rating the difficulty of a selection. Since many of these factors are closely related to each other, as shown by their inter-correlations, it is apparent that the tabulation of these factors simply gives the same measure over again. The most simple method for predicting the difficulty of these selections is to count the *number of different technical words in the selection*, the *number of different hard non-technical words*, and the *number of indeterminate clauses*.

TABLE I**COEFFICIENTS OF CORRELATION BETWEEN VARIOUS FACTORS
FOUND IN READING MATERIALS AND THE AVERAGE
COMPREHENSION-SCORES FOR THESE MATERIALS**

Factor found in Reading Material	Coefficient of Correlation Factor Found in Reading Material between the Factor and Comprehension Scores
Number of different technical words in the selection.	-.462
Number of hard non-technical words in the selection not known by 90 per cent of pupils in Grade VI as shown in the Dale list	-.380
Number of technical words in the selection known to less than 75 per cent of pupils in Grade VI as shown in the Dale list	-.377
Number of different hard non-technical words in the selection.	-.374
Percentage of monosyllabic words in the selection	.367
Percentage of easy words in the selection	.352
Number of prepositional phrases in the selection.	-.345
Number of words in the selection beginning with <i>i</i>	-.319
Percentage of hard non-technical words in the selection not known by 75 per cent of pupils in Grade VI as shown in the Dale list	-.309
Percentage of technical words in the selection	-.302
Average number of words in each sentence of the selection	-.293
Number of second person pronouns in the selection.	.251
Percentage of bisyllabic words in the selection	-.227
Number of words in the selection beginning with <i>w</i>	.226
Number of indeterminate clauses in the selection	.226
Average difficulty score for the technical words in the selection as computed from the Dale list	.225
Number of words in the selection.	-.195
Percentage of hard non-technical words in the selection	-.193
Number of words in the selection beginning with <i>e</i> .	-.164
Percentage of compound-complex sentences in the selection	-.149
Percentage of simple sentences in the selection.	.146
Number of easy words in the selection	-.114
Number of sentences in the selection.	.112
Number of first person pronouns in the selection	-.091
Number of words in the selection beginning with <i>b</i>	-.076
Number of words in the selection beginning with <i>h</i>	-.075
Number of third person pronouns in the selection	-.069
Percentage of complex sentences in the selection.	-.028
Percentage of compound sentences in the selection	-.024

The three factors when combined give a coefficient of correlation of 0.511 with the actual difficulty of comprehension as shown by the tests. These three factors are about as good for purposes of prediction as are the ten factors. For most purposes librarians and others interested in selecting reading materials of given difficulty will find these three counts most valuable. Their use makes possible the prediction of the difficulty of selections comparable to those used in this study, estimates which will indicate the proportion of the group who can comprehend the selections. In more than 85 per cent of the cases such an estimate will be in error by less than 20 per cent.

If one is merely trying to judge the relative difficulty of selections, it is sufficient to make a count of these three factors and to select as least difficult those with the fewest different technical words, with the fewest different hard non-technical words, and with the most indeterminate clauses. On the other hand, if one wishes to predict the difficulty of reading materials in terms of the proportion of a group made up of persons of third-to fifth-grade reading ability who are likely to comprehend the materials, the following regression equation may be used:

$$x_1 = -9.4x_2 - .4x_3 + 2.2x_4 + 114.4 \pm 9.0 .$$

In this equation x_1 stands for the percentage of a group of adults with third- to fifth-grade reading ability who, when tested on the reading material, will comprehend it; x_2 equals the number of different technical words found in a sample of the size of the selections used in this study, viz., about four hundred words in length; x_3 equals the number of different hard, non-technical words found in a sample of the size of the selections used in this study; x_4 equals the number of indeterminate clauses found in a sample of the size of the selections used in this study. The percentage of the group who will probably comprehend the material when predicted in this fashion has a probable error of 9.0, which means that probably one-half of all predictions thus made would differ from the actual results obtained by testing by not more than 9 per cent.

The use of this regression equation may easily be shown by illustration. Suppose one wished to select some reading materials which are easy enough so that they would be comprehended by at least 80 per cent of adults who have from third- to fifth-grade reading ability. Samples of these selections of similar size to those used in this study could be examined and the number of different technical words, the number of different hard,

non-technical words, and the number of indeterminate clauses counted. These counts could then be multiplied by -9.4, -0.4, +2.2, respectively. The resulting products could then be added to 114.4. If all the selections were chosen in which the resulting sums were 80 or higher, we should have those in which the predicted difficulty of comprehension would be such that the selections would be understood by 80 per cent or more of the adults who have third- to fifth-grade reading ability. For these selections we should probably find half of them within 9 per cent of the difficulty predicted. Four-fifths of our selections would probably be found by tests to be within 18 per cent of the difficulty predicted. Hence this regression equation does give a reasonably accurate prediction of the difficulty of reading materials similar to those used in this study. It should be pointed out that the samples tabulated should be of similar size to the selections used in this study. For long articles several such samples should be examined and the average used in the regression equation.

PROBLEMS REQUIRING FURTHER INVESTIGATION

As a source from which satisfactory methods of providing reading materials for adults of limited education may be obtained, the study which has been described requires additional supplementary investigations. A technique for attacking the problem has been developed, but this study has been limited to an analysis of reading materials dealing with health problems, and the difficulty of these materials has been measured only in terms of the comprehension of these materials by a group of colored adults in the city of Columbus. The investigation should be extended by analyzing in similar fashion reading materials dealing with several other topics of high interest to adult groups. By determining whether the factors found to be significant in this study are also useful in estimating the difficulty of reading materials dealing with other topics the value of the present investigation would be greatly enhanced. Furthermore, the application of this technique to other typical groups of adults who have limited skill in reading would give evidence as to the usefulness of this method of determining reading difficulty for adults of limited education generally.

Another logical extension of this study would involve experimentation in the writing of materials appropriate for adults who have limited skill in reading. Several typical groups should be chosen, and the non-fiction topics in which they were greatly interested could then be determined. It would be possible, by means of the technique developed in this study, to select reading materials on these topics which would be expected to be

easily comprehended by these adult groups. These materials could then be distributed, and the degree to which they were understood by the individuals in the group could be determined by tests and by interviews. Additional materials dealing with these topics could then be prepared, written in such a way that the factors discovered in this study to have close relationship to difficulty would be present in the amount which would make for ease in comprehension. These new materials could then be distributed to the groups and the degree to which they were comprehended determined again by tests and interviews. Such an investigation would not only throw light upon the adequacy of the technique as a means of selecting appropriate materials for adults of limited reading ability, but it would also develop a method by which satisfactory materials could be prepared for use by these groups.

If these additional studies demonstrated that certain definite factors were in general closely related to the difficulty of the reading materials, it would be possible to experiment with a central agency, such as the American Library Association, which would be responsible for preparing lists of reading materials appropriate for adults of limited education. It would be much more economical for a central agency to estimate the difficulty of reading materials dealing with topics of interest to adult groups than to have each branch library in the country make its own estimate of difficulty. The method of having a central agency prepare a list of appropriate materials could be tried out for a brief period, and evidence could be collected from individual libraries regarding the usefulness of these lists. Such a study might require several field-workers who would spend some time in the branch libraries interviewing patrons to discover the degree to which the materials suggested in the reading lists were appropriate for the adults of limited reading ability representing the clientele of these libraries. The usefulness of these lists could also be checked by other adult-education agencies. Such extensions of this type of investigation would greatly enhance the value of the study.

EXHIBIT A

SAMPLES OF MATERIAL TAKEN FROM TEST 1

BUREAU OF EDUCATIONAL RESEARCH
OHIO STATE UNIVERSITY

EDGAR DALE AND RALPH TYLER, *Investigators*

ADULT READING STUDY

PROJECT NO. 301

1. Write your name here
First name _____ Last name _____

2. Put a circle around the age nearest yours:

15 20 25 30 35 40 45 Over 50

3. Put a circle around the last grade in which you went to day school:

1 2 3 4 5 6 7 8 9 10 11 12

4. Put a circle around the length of time you went to night school:

Months						Years			
0	3	6	9	12	18	1	2	3	4

5. Put a circle around the number of minutes or hours you read every day;

Minutes					Hours		
0	15	30	45	60	1½	2	3 or more

On the following pages you will see paragraphs from health articles which have been in newspapers. Following these paragraphs are five sentences about the reading materials in the paragraphs.

Read these paragraphs carefully. Then read the sentences which follow. Put a CHECKMARK (✓) after the sentence which best tells what the article has said. One of the sentences is all wrong, or tells something which was not in the paragraph. MARK IT WITH A ZERO (0). Do not mark the other three sentences.

Now, remember to put a CHECK (✓) after the sentence which best tells what was in the paragraph, a ZERO (0) after the sentence which is all wrong, and do not mark the other three sentences. Here is an example:

Children should not eat with adults. They should eat either before or after their elders. This plan permits the parents to eat undisturbed. It does away with the tempta-

tion of giving "tastes" of this and that to the youngsters.

Summary of Paragraph

1. Young children should not eat with their parents.
2. Parents should eat undisturbed.
3. Children should always eat earlier than their parents.
4. When children eat with their parents they want to eat everything that their parents eat.
5. Parents should allow their children to taste all foods.

Time

Article C-6

Sinus disease is one of the most prevalent and common of body disorders. Though no actual figures have been compiled, so far as I know, its prevalence at times is appalling.

At least one person in every family seems to be the victim of this complaint. What can the reason be and how can we prevent this annoyance?

Sinus disease is usually the result of neglected head colds. It often follows sore throat, ear disease, grippe, bronchitis, pneumonia, or any infection of the upper air passages.³

Summary of Paragraphs

1. No actual figures have been compiled on sinus ills.
2. Sinus trouble is very common and is usually caused by neglected head colds,
3. Sinus trouble is the same as sore throat, ear disease, and grippe.
4. What is the reason for the prevalence of sinus trouble?
5. Sinus follows infection of the air passages.

Time

Article B-7

A medical colleague offers this interesting contribution: In a country practice of thirty years I have had my share of diabetic patients, and for years I have noticed that the use of tobacco is rare among them. In fact, I can recall only one case, that of a doctor who died in coma; he was a voracious eater and a heavy user of tobacco.

For about ten years I have been advising my diabetic patients to smoke—that is, the men, but I suppose I might include the women as well now. I believe I have observed consistent benefit from it,

I do not know whether tobacco has any definite therapeutic effect on the pancreas

³ Royal S. Copeland, "Keeping well" (King Feature Syndicate).

or the secretion of insulin. The only theory I have is that tobacco diminishes appetite to some extent and in that way tends to prevent or to moderate the disease.

Few heavy smokers or chewers of tobacco are heavy eaters.

In my experience, overeating is the chief cause of diabetes, and I believe that is why but few users of the weed have it.

Non-users are constantly eating something—candy, fruit, etc. Not because they're really hungry or it is regular mealtime, but just for something to do or just to satisfy a vague craving.⁴

Summary of Paragraphs

1. Diabetes is probably caused by the use of tobacco.
2. The heavy smoker is not a heavy eater.
3. The diabetic person smokes and therefore does not eat so much candy.
4. The doctor has had his share of diabetic patients.
5. This doctor believes that the use of tobacco helps diabetic patients because it causes them to eat less.

Time

Article E-2

The field laborer in hot countries, in the very countries from which these people come, has taught the world just the proper garment to wear for hot-weather comfort. It is of cotton, is light in weight, open in weave, and white in color. It is made into two garments—a shirt (with the tail worn out, making a blouse) and the trousers loose at the bottom and not very tight at the middle. However, Mexican and other Spanish-American women are not permitted by custom to follow this example set by their fellow-countrymen.⁵

Summary of Paragraph

1. Custom does not permit women to wear coats and trousers.
2. The men field laborers in hot countries cannot wear white cotton because it is against their customs.
3. The shirt should be worn outside the blouse.
4. The best hot-weather clothing is light-weight, open-weave cotton, made into a trousers and shirt that fit loosely.
5. We should follow an example of field laborers.

⁴ William Brady, "Personal health service" (National Newspaper Syndicate).

⁵ Dr. W. A. Evans, "How to keep well," *Chicago Tribune*.

EXHIBIT B

SAMPLES OF MATERIAL TAKEN FROM TEST 2

TIME

Milk is so nearly a perfect food that people who are ill can live on it for many weeks. But it is not a complete food. One reason is that it lacks iron, which we need for the blood. If we do not have iron we begin to grow pale and ill. Fortunately, fruits and leafy vegetables have a fairly good supply of iron. Newborn babies thrive on milk alone for a long time because they have enough iron in their bodies at birth to last for many weeks. After that time they must get iron in some way in order to be healthy.⁶

Paragraph Summary

1. Fruits and leafy vegetables give you iron.
2. Milk is a perfect food.
3. If milk had iron, it would be a complete food.
4. Newborn babies thrive on milk alone.
5. The blood needs iron.

TIME

When you burn yourself you need not bother to wash the burn with water, for the heat has killed the germs. The first thing to do is to keep the burn covered from the air. To do this, cover it with a paste made of baking-soda mixed with water. Lard or vaseline is good. Keep on the grease or baking-soda until the pain is gone.⁷

Paragraph Summary

1. Smear grease on a burn.
2. Wash a burn carefully so that there are no disease germs left.
3. Grease or baking-soda paste is good for a burn because it keeps out the air.
4. Mix the baking-soda with water before you put it on the burn.
5. Lard or vaseline is very good for a burn.

⁶ Andress and Evans, *Health and success* (Ginn & Co.), p. 185.

⁷ Ibid.

TIME

It has long been known that certain diseased conditions in man are due to the lack of particular foods. Thus, sailors whose vessels were kept at sea by storms and contrary winds until their supply of fruits and fresh vegetables gave out, often came down with the scurvy. They became weak, sores broke out on their bodies, and many died of the disease. If, however, they could put into some island where they could find fruits, like lemons or oranges, or fresh green stuff to eat, their scurvy would disappear. This was particularly true in the days of the sailing vessels when ships were often at sea for months between ports.⁸

Paragraph Summary

1. Scurvy may sometimes be gotten by eating fruits and vegetables that are a bit too fresh.
2. Sailing vessels were often at sea for months between ports.
3. Lemons and oranges grew on some of the Islands visited by sailors.
4. Sailors who ate fruit and fresh vegetables did not get scurvy.
5. Sometimes the supply of fresh fruits and vegetables gave out.

TIME

Anna had a raincoat and rubbers. She wore them to the picture show and did not take them off. She was so warm after the show that she took off her raincoat and sat on the porch for a while. She caught cold and had to stay in bed for two days.⁹

Paragraph Summary

1. Never wear a raincoat and rubbers when you are on the way to a show,
2. Anna caught a bad cold.
3. Anna had to stay in bed for two days.
4. Anna sat on the porch after the show.
5. Anna caught a cold because she didn't take her raincoat and rubbers off in the show.

TIME

In the army where men do so much marching, the feet of the men must be kept free of sores. The shoes are not tight, and the men are told to cut the toenails straight across, instead of rounding them. If the nails are cut straight across, or almost straight, and if the shoe is not too small, you will not be bothered by an "in-growing" toenail.¹⁰

⁸ Elliot R. Downing, *Science in the service of health* (Longmans), p. 282.

⁹ Williams and Dansdill, *Health and happiness* (Sanborn), p. 101.

¹⁰ Burkhard, Chambers, and Maroney, *Health habits* (Lyons & Carnahan), I, 35.

Paragraph Summary

1. Cut the nails straight across.
2. Men must do much marching in the army.
3. The feet of soldiers must be kept from being sore.
4. In-growing toenails are caused by tight shoes and rounding the toe-nails when they are cut.
5. Buy shoes that fit.

TIME

Cotton is probably the best material for underclothing, as it is comfortable to the skin. It is cool and it easily lets the heat and air out from the body. It does not need to be washed as often as wool. It washes easily, that is, it does not shrink much. Cotton underclothing can be worn not only in summertime, but unless one has a weakness of some kind, it can be worn without danger in the winter as well. It has an extra advantage in that it is less expensive than other kinds of underclothing.¹¹

Paragraph Summary

1. Some people wear cotton underclothing in winter.
2. One bad thing about cotton is that you have to wash it more often than wool.
3. Cotton is less expensive than other kinds of underclothing.
4. Cotton is cool.
5. Cotton is probably the best material for underclothing.

¹¹ *Ibid.*, p. 99

EXHIBIT C

SAMPLES OF MATERIAL TAKEN FROM TEST 3

TIME

Tony went to the drug store for some iodine. The doctor had told him to put some on his finger, which he had cut with an old knife. Iodine is very good for a new cut. On the outside of the bottle the druggist had put a label to show that the iodine was poison. Tony forgot to put the bottle away, and left it on the window sill. Soon his little baby brother, who was just beginning to walk, came along and started to drink the iodine. Just then the mother saw the baby, and screamed in time. The baby had swallowed only a little of the iodine. Tony heard his mother scream. He jumped out of bed, and, half-dressed and barefooted, he ran to the drug store and telephoned for an ambulance.

The druggist, who had heard Tony, told him to run home and give the baby some flour paste and then some mustard water. Tony ran back as fast as he could go and told his mother what the druggist had said. She snatched a spoon from the table and stirred some flour in a half-cup of water. Then as she held the baby, Tony poured this into the baby's mouth. In another cup she stirred a half-spoonful of mustard into some water, and poured several spoonfuls of this mustard water into the baby's mouth. This made the baby vomit. Then the ambulance arrived, and the baby and his mother were soon in the hospital. When the good doctor there took charge of the baby, he said that Tony and the mother had really saved the child's life.¹²

Paragraph Summary

1. If a person has swallowed poison give him some flour paste and mustard water. This will make him throw up.
2. The baby and his mother went to the hospital in an ambulance.
3. Tony went to the drug store for some iodine.
4. You put the flour in a half-cup of water.
5. Never put iodine on a new cut. It is poison.

¹² Myers, *The Language of America*, II, 27.

TIME

Is freshly baked bread more harmful than bread a day or two old?

When you eat fresh bread it very quickly forms into a ball which is then swallowed. Bread a day or two old doesn't form a ball so easily and you can chew it longer. The digestive juices can work on a lot of little pieces better than they can on one big ball. That's why fresh bread isn't so good for you as bread that is a day or two old.¹³

Paragraph Summary

1. Fresh bread forms a ball in your mouth.
2. You chew old bread longer.
3. Bread which is a day or two old is easier to digest than very fresh bread because you chew it longer.
4. Digestive juices can work better on bread a day or two old.
5. Bread should be baked hard on top.

TIME

Is graham bread more healthful than white bread?

If you were to eat only white or only graham bread, then the graham bread is better. But if you eat the right kinds and amounts of other foods with your white bread there is no difference. Graham bread, however, has more bran, salts, and vitamin content than white bread.¹⁴

Paragraph Summary

1. Graham bread has bran in it.
2. If you eat the right kinds and amounts of other foods, white; bread is as healthful as graham bread.
3. Graham bread has salts in it.
4. Graham bread is better because it is baked longer.
5. There are vitamins in graham bread.

¹³ Adapted from *Hygeia—the health magazine*, March, 1931, p. 288.

¹⁴ *Ibid.*

TIME

The doctor and I watched the long row of men waiting in the bread line for their cup of coffee and doughnuts. The doctor shook his head and said, "Why are the people in charge of this bread line so foolish as to feed these starving men coffee and doughnuts. Those men won't get enough vitamines and maybe they'll start getting scurvy and pellagra. What those poor fellows need is thick vegetable soup and milk. That would be much better than coffee or doughnuts."¹⁵

Paragraph Summary

1. The city ought to feed coffee and doughnuts to all people out of work.
2. Coffee and doughnuts are not the right kind of foods to keep a person healthy.
3. The men need thick soup.
4. These men might start getting scurvy and pellagra.
5. There were many men waiting in the bread line for their food.

TIME

A man went into a drug store and asked for something to cure a headache. The druggist held a bottle of ammonia under the man's nose. The ammonia was so strong that it nearly knocked the man over. It made the tears come in his own eyes. When he began to get over the shock, he got very mad at the druggist. He said, "I'll punch your nose." The druggist said, "What's the matter with you? Didn't I help your headache?" Then the man said, "Help my headache? Why, I didn't have any headache. It's my wife at home that's got the headache."

Paragraph Summary

1. This is a joke. The druggist thought the man had a headache and it was the man's wife that had it.
2. Ammonia was held under the man's nose.
3. Ammonia makes tears come in your eyes.
4. The man was going to punch the druggist's nose.
5. Always bring your wife with you when you get medicine. You will not get into trouble if you do.

¹⁵ Adapted from *ibid.*

1934—Ralph Ojemann: The Difficulty of Adult Materials

Introduction

THE YEAR 1934 marked the beginning of more rigorous standards for the formulas. Ralph Ojemann did not invent a formula, but he did invent a method of assessing the difficulty of materials for adult parent-education materials¹. His criterion was 16 passages of about 500 words taken from magazines. He was the first to use adults to establish the difficulty of his criterion. He assigned each passage the grade level of adult readers who were able to answer at least one-half of the multiple-choice questions about the passage.

Ojemann was then able to correlate six factors of vocabulary difficulty and eight factors of composition and sentence structure with the difficulty of the criterion passages. He found that the best vocabulary factor was the difficulty of words as stated in the Thorndike word list.

Even more important was the emphasis that Ojemann put on the qualitative factors such as abstractness. He recommended using his 16 passages for comparing and judging the difficulty of other texts, a method that is now known as **text scaling** or **leveling**. Although he was not able to express the qualitative variables in numeric terms, he succeeded in proving they could not be ignored.

The study again addresses urgent problem of providing suitable reading materials for adults of limited reading ability. As Ojemann states, at that time only half of the adult population had more than elementary school education.

Although there have been great advances, the problem today is no less urgent. Over a quarter of adults have not finished high school and their reading level is below the fourth grade. Providing parents in that category with suitable materials remains a leading issue for teachers, educators, and policy makers.

—WHD

¹ Ojemann, R. J. 1934. "The Reading Ability of Parents and Factors Associated with Reading Difficulty of Parent Education Materials." *University of Iowa Studies in Child Welfare*, Vol. 8, pp. 11-32 and 251-272.

University of Iowa Studies in Child Welfare

Vol. VIII

1934

The Reading Ability of Parents and Factors Associated with Reading Difficulty of Parent Education Materials

By Ralph H. Ojemann

READING is a very important factor in the process of educating parents just as it is in many other types of education. The recognition of this fact has resulted in numerous books, pamphlets, and magazines written primarily for parents. As one surveys this material the important question arises: How much of it falls within the reading comprehension of parents? Is the material too easy, too difficult, or of approximately the right degree of difficulty?

The importance of these questions requires little emphasis. If parents find their reading profitable and enjoyable in the sense that they derive significant ideas without too great an expenditure of time and effort, they will make extensive use of reading materials. On the other hand, if they find that they cannot understand what they read and that reading is a decidedly distasteful task, they will tend to make little use of this valuable means of education.

To determine the type of material that can be read by parents, two sets of facts are needed: First, what is the distribution of reading ability among parents? Second, what factors are most closely related to reading difficulty and what are the characteristics in terms of these factors of materials at various levels of difficulty? This investigation was undertaken to answer these questions.

SUBJECTS

The subjects used for studying the distribution of reading ability were 209 parents (mainly mothers) attending study groups in a city in Iowa having a population of approximately 30,000. In an effort to secure a ran-

dom sample, groups were selected in such a way that all parts of the city were represented. To interpret the data, it is important to know how this sample compares with parents in general. Two sets of facts, which were called for on the first page of the reading test (Appendix I), gave information concerning the extent of education and the occupation of husbands of the 209 subjects. The distribution of the parents as to extent of education is shown in the following tabulation:

Extent of Education	Number	Per Cent
Eighth grade or less	51	24.50
Ninth to twelfth grade	95	45.45
Above high school	53	30.14

In order to compare this sample with the population as a whole, the 1925 Iowa state census was studied as the best data available. Roughly 50 per cent of adults between the ages of nineteen and forty-five in 1925 had not extended their formal schooling beyond the eighth grade.⁵² It would seem, therefore, that the members of the study groups were drawn largely from the higher educational levels, as indicated by the fact that only one-fourth had a partial or complete elementary school education.

The distribution given below is based upon the occupation of husbands of the subjects studied:

Occupational Group*	Number	Per Cent
Professionals	18	8.61
Semi-professional and business	105	50.24
Skilled labor	55	26.32
Semi-skilled labor	14	6.70
Unskilled labor	14	6.70
Unknown	3	1.44

*Tausings classification was used

Comparisons with the 1920 census are of interest and value if they are

⁵² Actual calculation indicated that 482,603 of the 1,490,062 persons who had not extended their education beyond the eighth grade rural or grammar school were below nineteen years of age. This left a total of 1,007,459 persons nineteen years of age or above who had eight years or less in school. Considering the total number of persons (1,570,468) above the age of eighteen in Iowa, approximately two-thirds of the population over eighteen years of age in 1925 had not extended their education beyond the eighth grade. Of the total number of persons above eighteen in Iowa, there were 972,476 between the ages of nineteen and forty-five inclusive. Assuming that 90 per cent of the 597,792 individuals above the age of forty-five had only an eighth grade education, we find that 469,266 persons or approximately one-half the total number between the ages of nineteen and forty-five had not extended their education beyond the eighth grade.

carefully interpreted. (The 1930 census data are not yet available.) The data above show that 13.40 per cent of the families were classified in the semi-skilled or common labor groups and that 8.61 per cent belonged to the professional group. According to the 1920 census of the city in which this investigation was undertaken, approximately one-fourth (25.07 per cent) of the male working population may be classified as semi-skilled or common laborers and one-twentieth (4.97 per cent) as members of the professional class. From the decade 1920 to 1930 no unusual occupational changes were in evidence in this city; it is, therefore, reasonable to assume that the 1930 census figures will not differ greatly from those just presented. If this assumption is correct, there are slightly less than twice as many husbands in the professional group and about half as many in the semi-skilled and unskilled labor groups among the parents used as subjects of this study as in the total population.⁵³

In the study of factors associated with reading difficulty, the subjects were parents in an Iowa city of 18,000 population and mothers in the maternity ward of the State University of Iowa Hospital.⁵⁴ A total of 156 subjects was used. Thirty-seven subjects had a partial or complete elementary school education; sixty-four had spent some time in high school; and fifty-five had extended their education beyond high school.

DISTRIBUTION OF READING ABILITY

METHOD OF INVESTIGATION

To obtain an estimate of the variation in ability, a reading-comprehension test constructed from representative parent education materials was administered to 209 parents attending child study groups. The test itself may be found in Appendix I. The three selections of the test proper are sets of consecutive paragraphs, one set taken from each of three books: *Everyday Problems of the Everyday Child* by Thorn, *Child Guidance* by Blanton and Blanton, *Parents and the Pre-School Child* by Blatz and Bott.⁵⁵ Each set includes fifteen questions for the subject to answer;

⁵³ Comparisons In the semi-professional, business, and skilled labor groups are somewhat more difficult to make and will not be attempted since the classifications are not entirely satisfactory.

⁵⁴ The writer wishes to express his appreciation to Dr. E. D. Plass for permission to use the subjects in the State University of Iowa Hospital.

⁵⁵ Blatz, William E., and Bott, Helen: *Parents and the Pre-School Child*. New York: William Morrow, [c.1929] Pp. xii, 340. (p. 52-54)

Thorn, Douglas A.: *Everyday Problems of the Everyday Child*. New York: D. Appleton,

the score consists of the number of questions answered correctly. These three selections were taken from a group of nine prepared and submitted to several competent judges who examined the material critically to assure that each question had only one answer. Two of the nine selections proved unsatisfactory and were eliminated; three selections were chosen from the remaining seven.

From the above description it may be inferred that the test is primarily one of reading comprehension. Thirteen minutes were allowed for each selection. This time limit appeared ample for practically all subjects.

An analysis of the results indicated that the test had a high degree of reliability and that it might possibly be shortened to two selections. A study was made, therefore, of the distribution of reading ability in the group of 209 subjects and in an additional group of 156 parents in terms of the abridged reading test, that is, a test consisting of only the first two selections. The abridged test had a satisfactory reliability and was used in the study of the factors associated with reading difficulty.

RESULTS

In discussing the results of this part of the study two questions arise: (1) How reliable is the test used? (2) What scores were made by the subjects and how are these scores to be interpreted?

The reliability of the reading test is .96 as measured by correlating odd and even scores and applying the Spearman-Brown formula.

Another method that gives a valuable indication of the way in which a test functions is to examine the behavior of individual test items. By test item is meant each of the forty-five questions comprising the test; for example, question 1 in the first selection constitutes one test item, question 2 in the same selection another, and so on. One criterion of a good test item is that its difficulty should decrease in passing from a group of subjects of low ability to a group of significantly higher ability. In the data available, there is an opportunity to apply this criterion. The group having eight years or less of schooling made an average score of 18.98 ± 1.02 . The group having a partial or complete high school education made an average score of $29.25 \pm .68$. The subjects who had extended their edu-

1927. Pp. xiv. 349. (p. 182-184)

Blanton, Smiley, and Blanton, Margaret Gray: *Child Guidance*. New York: Century, [c.1927] Pp. xviii, 301. (p. 31-32)

tion beyond high school achieved an average score of $34.97 \pm .83$. The per cent of errors for each group on the individual test items is given in Table 1.

An analysis of the data reveals that each item shows a decrease in the per cent of errors when the group having an educational level of eighth grade or less is compared with the ninth to twelfth grade range and with the above high school group. Even though the difference in ability between the ninth to twelfth grade group and the above high school group is only slightly more than four times its probable error, each item shows the same or lower percentage of errors except item five in Selection 1, item six in Selection 2, and item two in Selection 3. The data presented thus far indicate that the test has a satisfactory reliability when chance halves are correlated, and that individual items show a satisfactory behavior when groups of significantly different abilities are compared.

Table 1
Per Cent of Errors Made on Each Test Item

Question	Selection								
	1			2			3		
	Extent of Education								
	Eighth Grade or Less	Ninth to Twelfth Grade	Above High School	Eighth Grade or Less	Ninth to Twelfth Grade	Above High School	Eighth Grade or Less	Ninth to Twelfth Grade	Above High School
1	69	53	29	43	21	13	45	15	11
2	73	66	38	55	34	21	.35	12	14
3	71	56	49	59	42	21	.82	57	38
4	45	14	11	29	11	11	.69	32	27
5	61	38	40	31	20	13	.76	46	29
6	51	26	14	49	19	22	.41	18	11
7	73	31	22	65	33	17	.71	29	22
8	82	52	41	49	34	29	.73	47	27
9	43	11	6	35	11	10	.82	52	35
10	63	31	21	84	53	37	.67	32	14
11	69	29	16	39	17	11	.73	55	35
12	78	42	25	80	61	37	.78	48	33
13	49	24	13	67	29	22	.67	41	27
14	76	42	25	59	18	10	.57	36	16
15	86	64	40	63	24	16	.76	58	29

Distribution of Scores.—Distribution of scores for the entire group and for the groups classified on the basis of grade reached in school is presented in Table 2. The scores range over the entire scale. Two parents made scores of 0 or 1; eleven made a score of 44 or 45; the average for the entire group was $28.75 \pm .56$.

Interpretation of the Scores.—In interpreting the scores the question arises, "What score should a parent make to warrant the statement that he can read the material satisfactorily?" Several considerations may guide us in attempting to answer this question. First, in constructing the test an effort was made to select questions involving the important ideas expressed in the paragraphs. The judgments of three people, including the writer, were used for this purpose. Second, consideration may be given to what the results would mean in terms of a whole book. Thirteen minutes were allowed for each selection, a selection being equivalent to about one and one-half pages of the average book. At the rate of thirteen minutes for each one and one-half pages, it would take a parent making the average score of the group, 28.75, forty-four hours to read a book of this grade of difficulty. According to the method used in the test, he would be able to identify only about 63 per cent of the ideas called for. It is the rare individual who would enjoy reading under such conditions!

If we assume that a parent should make a score of at least 30 (two-thirds of a perfect score), it appears that there is a considerable proportion of parents who have difficulty with these materials. Fifty-one per cent of the subjects in this study scored below 30. The average score of subjects having a partial or complete elementary school education was 18.98 ± 1.02 , 24.40 per cent of the total subjects falling in this group. As previously stated, the proportion in the general population is much larger than this. If, then, the problem of providing reading materials is conceived in terms of all parents, the number of adults having difficulty with the type of material used in this study becomes very large and the problem takes on tremendous significance.

Distribution on the Short Form of the Reading Test.—For the purposes of the second part of this investigation, it was desirable to shorten the test. The reliability of the first two parts for the 209 subjects, as measured by the correlation of even and odd scores, is $.87 \pm .01$; with the application of the Spearman-Brown formula this becomes .93. This reliability is sufficiently high to make possible the abridgment of the test to the first two selections which require a total time of only twenty-six minutes.

Table 2
Distribution of Reading Scores

Score	Total Group	Eighth Grade or Less	Ninth to Twelfth Grade	Above High School
44	11	0	0	11
42	13	0	6	7
40	16	1	8	7
38	22	1	12	9
36	11	1	6	4
34	11	2	7	2
32	6	2	4	0
30	12	0	6	6
28	17	3	11	3
26	10	3	3	4
24	13	4	8	1
22	9	3	5	1
20	8	4	Z	2
18	11	7	3	1
16	13	5	6	2
14	1	1	0	0
12	6	2	3	1
10	0	0	0	0
8	6	3	2	1
6	3	2	1	0
4	5	3	1	1
2	3	2	1	0
0	2	2	0	0
Mean	28.75	18.98	29.25	34.97
Probable Error of Mean	±.56	±1.02	±.68	±.83

Since 156 subjects were used in the second part of the investigation, data showing the distribution of scores on the short reading test are available from 365 subjects. The mean score on the test for the three sub-groups, classified according to extent of education, is presented in the following tabulation:

Extent of Education	Subjects	Mean
Eighth grade or less	88	14.50 ± .48
Ninth to twelfth, grades	159	21.36 ± .26
Above high school	118	25.10 ± .36

The mean of the subjects having more than a high school education is slightly too low, since the test is not sufficiently complex to take care of

the highest levels of reading ability. However, it was considered unnecessary to increase the difficulty of the test since the 2 reading problem was mainly one for parents whose comprehension falls below the college level.

FACTORS ASSOCIATED WITH READING DIFFICULTY

METHOD OF INVESTIGATION

To carry out this part of the investigation, the reading difficulty of sixteen selections was determined experimentally and the variation of numerous factors was analyzed as regards the difficulty of the selection. The general method consisted in finding the average reading ability of subjects who achieved a specified score on a selection, and expressing the difficulty in terms of this average reading ability.

Sixteen selections, each consisting of approximately 500 words, were chosen at random from fifteen different parent education books. Each selection was prepared in reading test form. Considerable care was exercised in the preparation, of the questions. An attempt was made to use short sentences with a vocabulary-similar in degree of difficulty to the respective selections. Five judges examined the questions and took the tests to remove special difficulties and ambiguities. Fifteen minutes were allowed for each test. The purpose of the ample time limit was to obtain a measure of comprehension rather than of speed. The score consisted of the number of questions answered correctly.

The general method for each selection involved finding the mean reading ability of subjects who achieved a specified score. The selections were submitted, therefore, to parents whose scores on the reading test were obtained at the same time. A trained examiner administered the tests to individuals or small groups where conditions could be carefully controlled. The materials were given in random order and each parent took approximately six tests. Each selection was included in the series until fifteen to twenty cases were secured whose scores fell between nine and twelve inclusive. This range of scores was assumed to indicate satisfactory comprehension and at the same time to be sufficiently wide to avoid making the gathering of the data an unduly long process.

When the order of difficulty had been determined, a quantitative and qualitative analysis of the materials was made and the relation of various factors to difficulty studied. The quantitative study included six measures

of vocabulary difficulty and eight factors of composition and sentence structure. The details of measuring each factor can best be explained in the presentation of results.

In the qualitative analysis consideration was given to such factors as concreteness or abstractness of the relationships discussed as distinguished from the individual words used, obscurity in expression, and incoherency in expression. In similar studies that have been carried out for the most part -with school children, qualitative factors have been overlooked in general. Their importance may be made clearer by considering an extreme example. If in a set of paragraphs the sentences were arranged in random order, the number of simple sentences, the vocabulary difficulty, etc. would remain constant, but there is considerable possibility that comprehension would be interfered with. The details of the qualitative analysis will be presented in the discussion of the results.

RESULTS

Relative difficulty was measured in terms of the average score achieved on the abridged reading test by subjects who comprehended the material in a given selection.

Table 3

Mean Score on Reading Test Required for Satisfactory Comprehension

Selection	Subject	Mean Score on Experimental Selection	Mean Score on Beading Teat
17	18	10.94	27.94
18	16	10.56	27.13
12	31	10.90	24.13
10	18	11.05	23.83
8	19	11.31	23.37
13	19	10.89	22.00
9	18	10.94	21.56
15	19	11.09	21.51
10	19	10.95	21.16
7	15	10.80	18.58
14	16	11.05	18.56
6	15	11.13.	18.47
1	15	11.11	17.20
5	16	11.00	14.50
4	17	10.92	12.94
3	17	11.09	12.35

Table 3 gives the number of subjects making a score between 9 and 12,

the average score achieved by these subjects on the experimental selections, and the average score on the abridged reading test. The selections are arranged in order of difficulty. From column three it will be seen that the average score on the experimental selections varied about three-fourths of a point. The range in difficulty is indicated by the figures in the last column.

The relative difficulty of the materials is presented graphically in Figure 1. It is interesting to note that most of the selections fall at or above the average comprehension ability of subjects having a partial or complete high school education. Only three are at or below the level of subjects having a partial or complete elementary school education. Four selections fall in the neighborhood of 17 and 18, four between 21 and 22, three in the interval 23 to 25, and the two most difficult selections fall slightly above 27. Since all of these selections, excepting perhaps the two most difficult ones which were taken from two books written by Gesell, were selected at random from books prepared primarily for parents, it is not difficult to see that most of them are pitched above the average comprehension ability of the large group of parents who have not extended their education above the eighth grade. About half of the selections are quite far above this average ability.

Quantitative Analysis of Factors.—In the quantitative analysis⁵⁶ of the experimental materials fourteen factors were studied. These include eight factors of composition and sentence structure: number of simple sentences, number of complex sentences, number of compound sentences, number of dependent clauses, average length of dependent clauses, ratio of total words in dependent clauses to total words in selection, number of prepositions, and number of prepositions plus infinitive signs. Six factors of vocabulary difficulty were studied: the per cent of words in Thorndike's first 1000 using the new Thorndike word list published in 1931, the per cent of words in Thorndike's first 2000, and four factors of word difficulty which require some explanation.

It was realized that the frequency of use of a word may not be very closely related to its difficulty, and that the same word may have different degrees of difficulty depending upon whether the more concrete or the more abstract meaning is used. It was desired, therefore, to obtain a

⁵⁶ The quantitative analysis of the experimental selections were made by two independent observers.

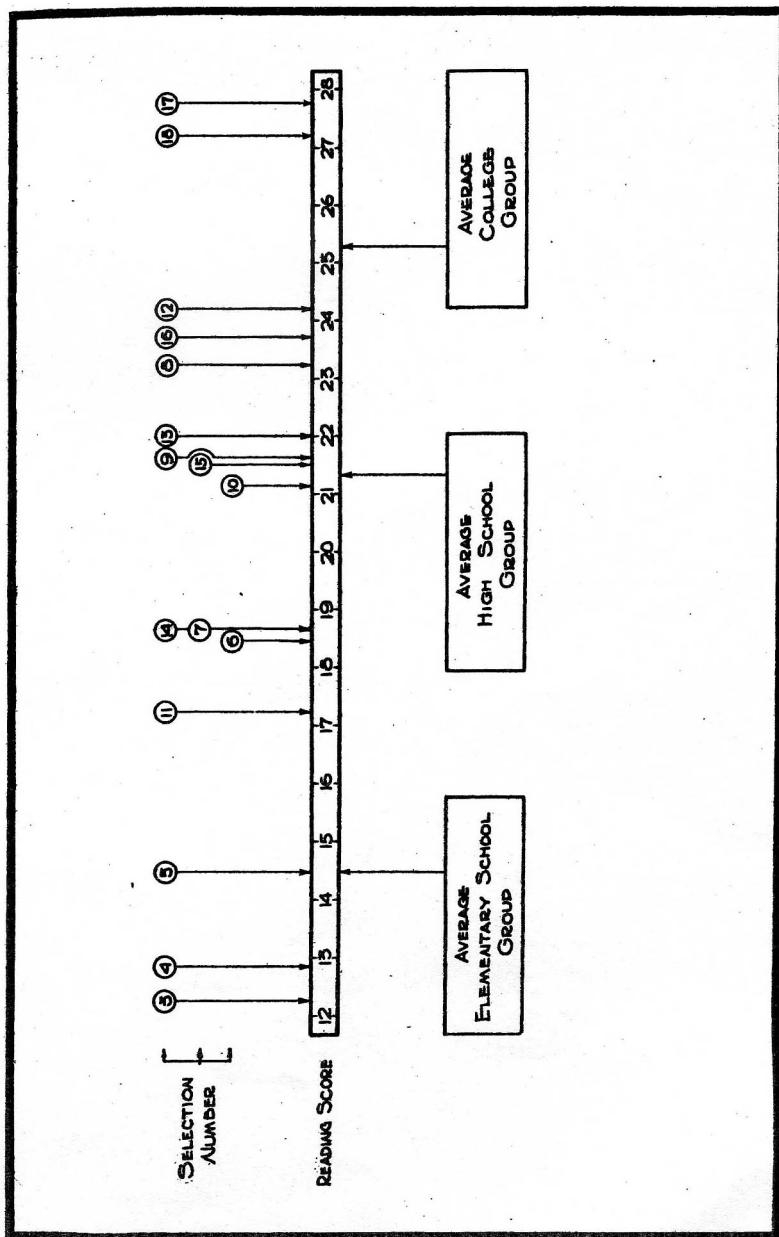


Figure 1. Relative Difficulty of Materials

measure of word difficulty. The most feasible plan that presented itself was to obtain the relative difficulty in terms of some standard group such as pupils in the sixth grade. Through the courtesy of Dr. Edgar Dale of

the Bureau of Educational Research at Ohio State University, a tentative list of the familiarity of 8000 common words to pupils in the fourth, sixth, and eighth grades was obtained. In the preparation of this list, it was assumed that approximately 2000 words appearing in the Thorndike list of 10,000 and in several vocabularies of kindergarten and primary children were known by at least 90 per cent of the children in grade four and beyond and, therefore, were not tested. Several words with two or more distinct meanings were tested in phrase or sentence context.

From this list the average difficulty per word in terms of the per cent of sixth grade pupils familiar with a given meaning, the average difficulty of different words, the percentage of words known by 90 per cent of sixth grade pupils, and the percentage known by 70 per cent of the pupils were calculated. There were 168 words in the sixteen selections that are not found in Thorndike's 10,000. Judgments of twenty-one carefully selected sixth grade teachers were averaged to obtain an estimate of the proportion of sixth grade pupils familiar with these words. This procedure was adopted because it was assumed that if this list of unusual words was presented to sixth grade pupils, the results would not be as reliable as those obtained by presenting a list containing only a small proportion of unusual words. The correlations of the fourteen factors considered in the quantitative analysis of difficulty are given below:

Factor Correlated With Difficulty	Correlation
Number of simple sentences	—.61
Number of complex sentences	.13
Number of compound sentences	Small range
Number of dependent clauses	.02
Mean length of dependent clauses	.40
Ratio of total words in dependent clauses total words in selection	.50
Number of prepositions plus infinitive signs	.66
Number of prepositions	.67
Mean difficulty of different words	—.74
Mean difficulty per word	—.72
Per cent of words known by 70 per cent of pupils	—.73
Per cent of words known by 90 per cent of pupils	—.73;
Per cent of words in Thorndike's first 1000	—.64
Per cent of words in Thorndike's first 2000	—.73

The above tabulation shows that there are three sets of factors which correlate closer than .60: number of simple sentences, number of prepositions including and excluding infinitive signs, and all of the six vocabulary factors. The correlation of number of simple sentences with difficulty is $-.61$. The two factors, number of prepositions plus infinitive signs and number of prepositions, correlate $.66$ and $.67$ respectively. The next highest factor dealing with sentence structure is ratio of total words in dependent clauses to total words in selections which correlates $.50$.

The correlation of the four vocabulary factors—per cent of words in Thorndike's first 2000, per cent of words known by 70 per cent and by 90 per cent of pupils, average difficulty per word, and average difficulty of different words—are all closer than $.70$. The factor showing the closest relationship is average difficulty of different words ($-.74$).

An inspection of the data upon which the tabulation on page 22 [120] is based shows that the relationship is not linear in all cases. A more useful analysis of the variation of a given factor with difficulty can be obtained, therefore, by studying directly a graphic presentation.

The curves in Figures 2, 3, 4, and 5 show the relation to difficulty of the four factors—number of simple sentences, number of prepositions, average difficulty of different words, and ratio of total words in dependent clauses to total words. In preparing the figures the sixteen selections were divided into five groups: Group I comprises the three selections whose difficulty in terms of the score on the reading test was 12.35, 12.94, and 14.50 with an average of 13.36. Group II is composed of the four selections ranging in scores from 17.20 to 18.58 and averaging 18.20. In Group III there are four selections ranging from 21.16 to 22.00 and averaging 21.56. The three selections that make up Group IV range in scores from 23.16 to 24.13 and average 23.77. In Group V are found the two hardest selections for which the difficulty scores are 27.13 and 27.94, an average of 27.53. The data upon which the curves are based are presented in Table 4.

Figure 2 shows the variation of average difficulty of different words with reading difficulty. Except for Group III, which is on the whole slightly easier in vocabulary than Group II, each group shows a greater vocabulary difficulty than the immediately preceding one. Figure 3 shows the variation of number of prepositions. The trend is distinctly upward, only one group deviating slightly from this trend.

Table 4

Analysis of the Five Groups of Selections

Group	Mean Score	Mean Difficulty of Words	Prepositions*	Simple Sentences	Proportion of Words in Dependent Clauses
I	13.26	94.4	43.7	16.0	.19
II	18.20	89.8	59.5	10.3	.33
III	21.56	91.5	55.5	5.5	.37
IV	23.77	85.3	62.6	7.6	.33
V	37.53	74.2	68.0	9.0	.30

* Exclusive of infinitive signs

The variation of the number of simple sentences with difficulty is represented in Figure 4. In the very simple selections the number of simple sentences tends to be high, but the curve soon approaches a region where the changes are not great. The curve for the ratio of total words in dependent clauses to total words in selection shows just the inverse behavior (Figure 5). A decrease in the number of simple sentences is accompanied, on the average, by an increase in the proportion of words in dependent clauses.

Qualitative Analysis of Factors.—Such factors as the concreteness or abstractness of the relations discussed as distinguished from individual words used, obscurity in expression, and incoherence in expression are involved in reading comprehension in addition to the quantitative factors of vocabulary and sentence structure. It is difficult to obtain an objective measure of such qualitative factors apart from reading comprehension scores. A subjective estimate, however, has the advantage of making the investigator aware of their existence.

In the writer's opinion, an analysis of the sixteen selections does not show any outstanding cases of lack of clearness in expression, or of abstractness in relationships other than those accounted for by the use of a more abstract vocabulary which, in turn, would affect the average difficulty of different words. There are two possible exceptions. Selection 3 contains a large number of relatively discrete ideas that could be more thoroughly integrated. Selection 13 tends to emphasize rather abstract relationships. The discussion deals with children's fears,—their relation to ignorance and curiosity, and the differences between fear and caution. The following paragraph from the selection is an example of difficult abstract relationships.

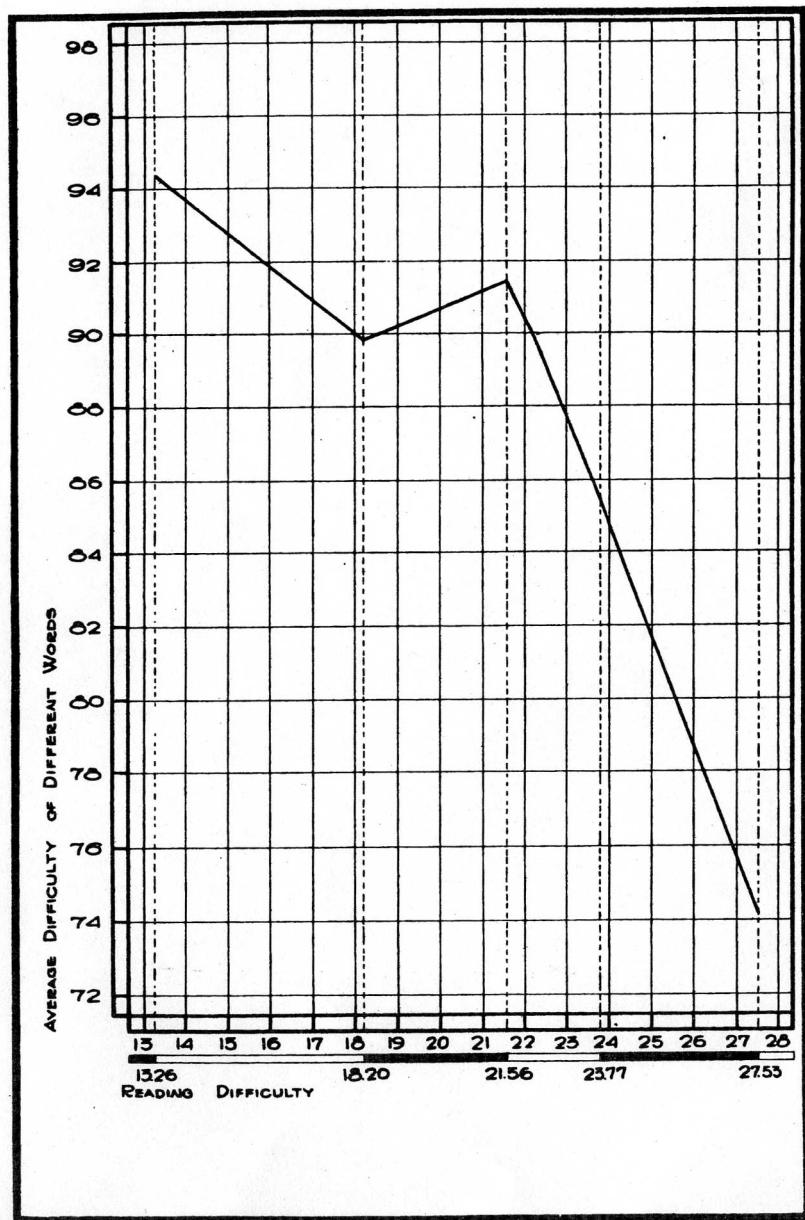


Figure 2. Relation of Reading Difficulty to Average Difficulty of Different Words

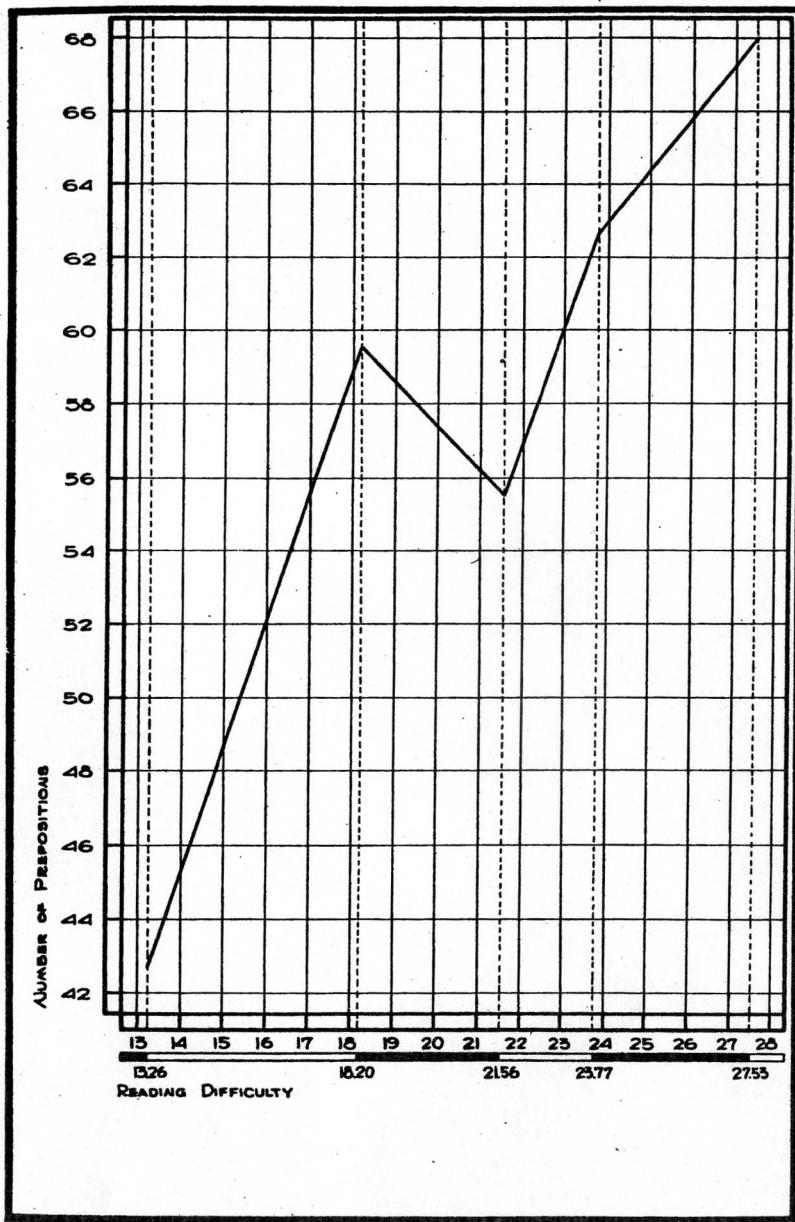


Figure 3. Relation of Reading Difficulty to Number of Prepositions

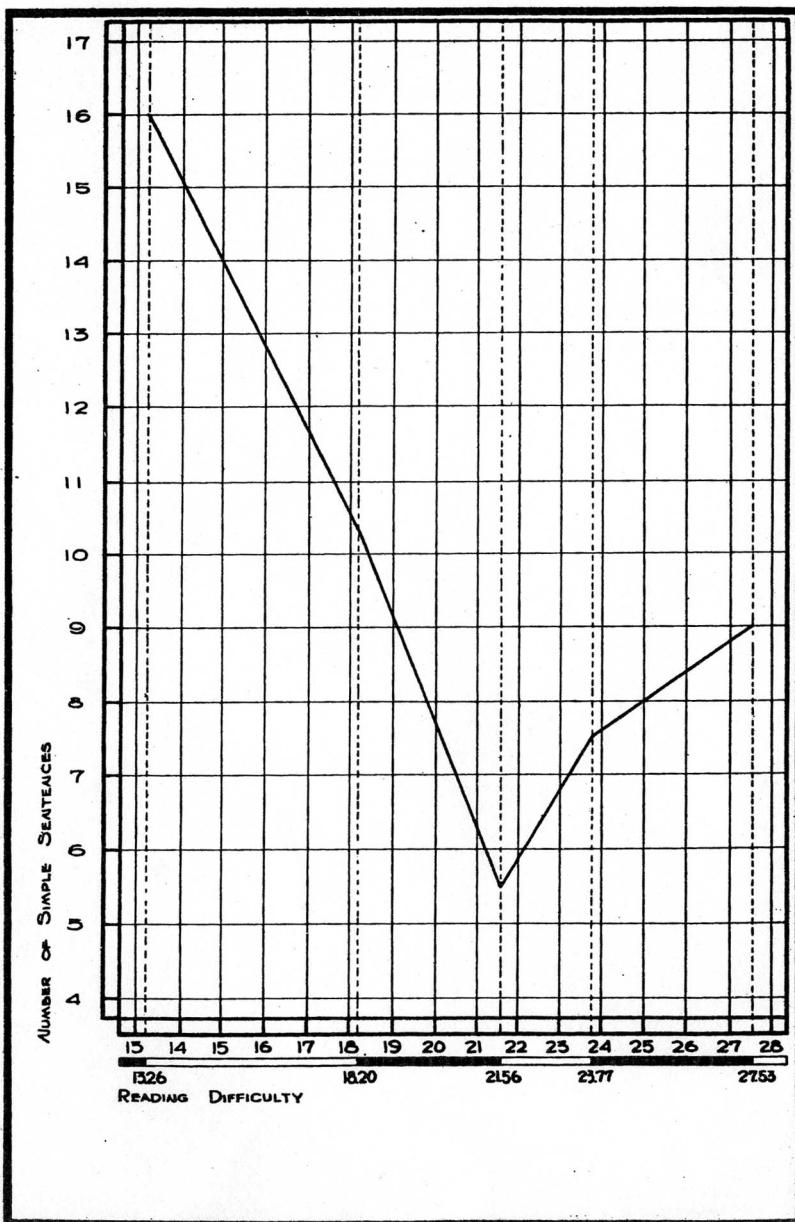


Figure 4. Relation of Reading Difficulty to Number of Simple Sentences

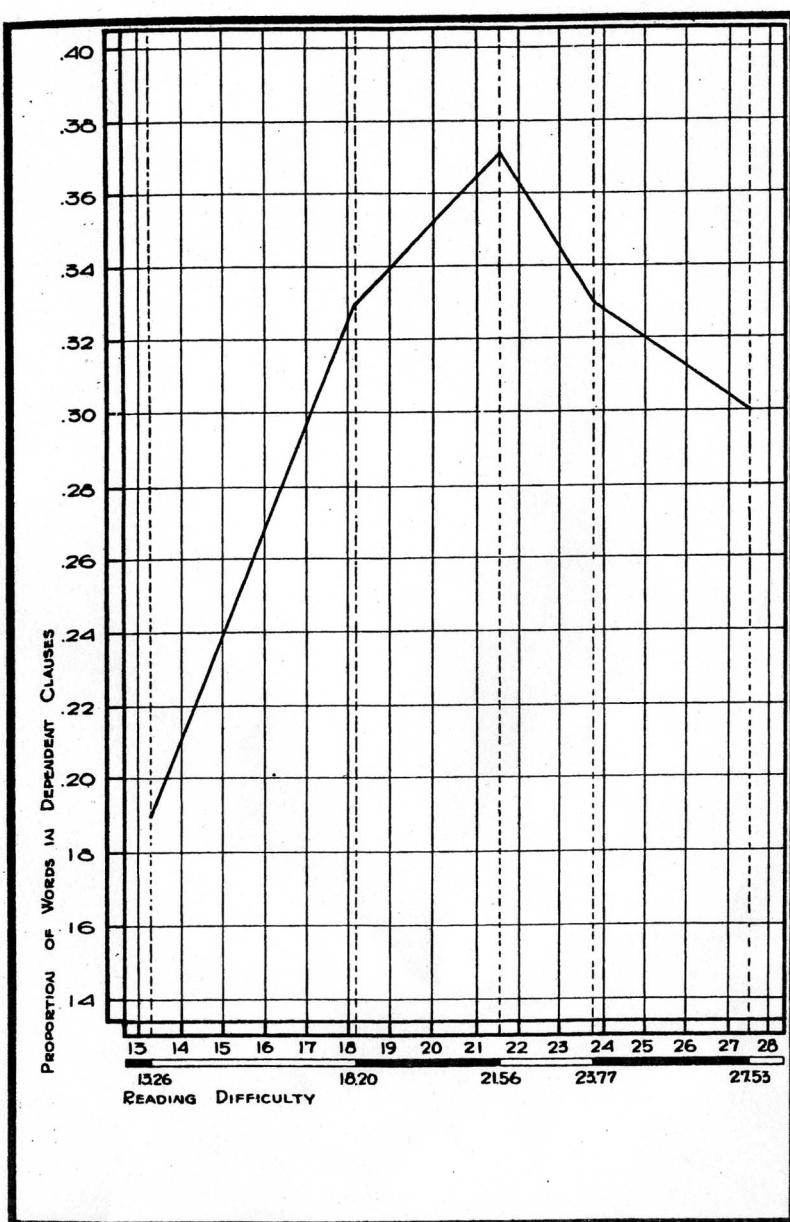


Figure 5. Relation of Reading Difficulty to Proportion of Words in Dependent Clauses

"Since fear rests so largely on ignorance, curiosity is its enemy, because it dissipates ignorance. A little boy who had a certain fear of the figures in the mirror that were so vivid and yet so unreal, used to try to come into a room in which there was a large mirror, and steal upon the objects of his curiosity unawares. His double was always there as soon as he, and caught his eye; but the child lost his fear only after he became familiar with the character in the looking-glass. In the same way curiosity will often help the child to become so well acquainted with the source of his fears as to drive the latter quite out of his experience."

Characteristics of Materials at Lower Levels of Difficulty

Writers of parent education materials are interested in the characteristics of materials that can be read by parents of limited ability. As we have stated above, about one-half of the adult population has only a partial or complete elementary school education. What types of materials can be read by this group? An analysis of the three selections, which the data show are at or below the average score made by parents who have not extended their education beyond the eighth grade, will supply some information on this problem (Table 5). The least difficult of the three selections is characterized by a large number of simple sentences, 23 per 500 word sample; a low proportion of total words in dependent clauses, .11; a small number of prepositions, 32 per 500 word sample; and a very simple vocabulary, average difficulty per different word equaling 94.9. Selection 4, next in order of difficulty, stands approximately .6 of a point above Selection 3.

Table 5

Analysis of Three Selections at the Lowest Level of Difficulty

Selection Number	Reading Difficulty	Simple Sentences	Mean Difficulty of Words	Prepositions*	Proportion of Words in Dependent Clauses
3	12.35	23	94.9	32	.11
4	12.94	20	94.4	49	.16
5	14.50	5	93.9	47	.31

* Exclusive of infinitive signs

As compared with the latter, it has approximately the same number of simple sentences, 20 as against 23; approximately the same ratio of total words in dependent clauses to total words in selection, .16 as against .11; approximately the same vocabulary difficulty, 94.4 as compared with 94.9; but a larger number of prepositions, 49 as compared with

32. Selection 5 is 1.5 points higher in difficulty than Selection 4. As compared with Selection 4 it has approximately the same vocabulary difficulty, 93.9 as against 94.4; approximately the same number of prepositions, 47 as compared with 49; but a very small number of simple sentences, 5 instead of 20; and a significant increase in the ratio of total words in dependent clauses, .31 as compared with .16.

An analysis of the content of the three selections shows that they deal with relatively concrete ideas—feeding and bathing the baby, providing play space and play equipment, and supplying fresh air and exercise.

These materials are characterized, therefore, by easy vocabulary and simplified sentence structure as evidenced either by a relatively large proportion of simple sentences or a small proportion of words in dependent clauses, and by a small number of prepositional phrases. An increase in the number of complex sentences or in the number of prepositional phrases tends to increase difficulty even when vocabulary is held constant.

In these simplified selections few abstract ideas were involved. The requirements of materials low in reading difficulty discussing concrete phases of child development are thus fairly clear. A more difficult problem is presented in the preparation of materials discussing the more abstract phases such as the principles of discipline or control. An analysis of one of the selections which ranks only a few points above the mean of the elementary school group in difficulty gives an indication of how this problem may be solved. Selection 6 discusses the principles of guiding the child in learning to put away his toys. Following a general discussion of the factors complicating the situation, three principles are set forth and are explained in relatively concrete terms. For example, the principle that picking up toys should be treated as a learning situation is made clear by enumerating several concrete acts which are involved in carrying out the principle: providing accessible places for toys, showing the child where to place them, helping the child to put them away, and being patient with the early fumbling efforts. This selection is slightly above the average ability of the elementary school group, but by simplifying the vocabulary and the sentence structure (through the use of more simple sentences and fewer complex sentences) and by supplementing the other abstract principles with concrete examples, it appears that the discussion could be brought within the desired comprehension range.

JUDGING THE DIFFICULTY OF READING MATERIALS

In Appendix I the sixteen selections are arranged in order of difficulty. At the beginning of each section a table presents the difficulty score and an analysis in terms of vocabulary difficulty as measured by the average difficulty of different words, number of prepositions, number of simple sentences, and proportion of words in dependent clauses. This series may be used to judge the difficulty of new materials by comparing the latter with the selection it most nearly resembles in quantitative and qualitative analysis. This method has been used successfully by the writer in a preliminary test with several different judges.

SUMMARY

In this investigation an analysis was made of the distribution of reading ability among adults, the factors most closely associated with reading difficulty, and the characteristics of materials at the various levels of difficulty in terms of these factors. These data are valuable in determining the type of material that can be read by parents at different levels of ability.

To study the distribution of reading ability a comprehension test was constructed from representative parent education materials and administered to 209 parents attending study groups. An abridged form of the test was administered to an additional group of 156 parents. The relative difficulty of sixteen selections of 500 words each was determined experimentally and the variations of quantitative and qualitative factors with difficulty analyzed.

Finally, the characteristics of materials in terms of important factors at the lower levels of difficulty were analyzed to assist writers in preparing simplified materials for parents.

The important findings are as follows:

1. The reliability of the original reading test as measured by the correlation of even and odd scores and the application of the Spearman-Brown formula is .96. The reliability of the abridged reading test, consisting of the first two of the three parts of the original test, is .93.
2. Parents attending study groups vary considerably in reading ability. Subjects having a partial or complete elementary school education made an average score on the original reading test of 18.98 ± 1.02 ; those having a partial or complete high school education, $29.25 \pm .68$; and those having

more than a high school education, $34.97 \pm .83$. On the abridged reading test the results for 365 subjects were $14.50 \pm .48$; $21.36 \pm .26$; and $25.10 \pm .36$ respectively. A perfect score on the original test is 45; on the abridged form, 30.

3. As measured by the technic used in this study, all but three of the sixteen selections taken at random from parent education materials required a greater reading ability than the average of subjects who have not continued their education beyond eighth grade. Nine required an ability equal to or above that of the average of subjects who have not advanced beyond high school.

4. Of the fourteen factors studied quantitatively in their relation to reading difficulty, the number of simple sentences, number of prepositions (inclusive and exclusive of infinitive signs), and all factors of vocabulary difficulty correlated above .60.

5. The materials having a difficulty equal to or below the average reading ability of subjects who had not extended their education beyond the elementary school are characterized by (1) low vocabulary difficulty (average difficulty 94.4), (2) simplified sentence structure as evidenced either by a relatively large proportion of simple sentences (average 16 per 500 word samples), or by a small proportion of words in dependent clauses (average .19), and (3) by a small number of prepositional phrases (average 42.7 per 500 word sample). An increase in the number of prepositional phrases or in the number of complex sentences tends to increase difficulty even when vocabulary is held constant.

Appendix 1

TEST OF PARENT EDUCATION MATERIALS AND SELECTIONS USED TO STUDY FACTORS ASSOCIATED WITH DIFFICULTY

PARENT EDUCATION READING MATERIALS

Iowa Child Welfare Research Station
State University of Iowa
Iowa City, Iowa

Name..... Age.....
Years Months

If you live in a town or city put the name of your town or city and your street address here

If you live *in* the country put your rural route number and your post-office address here

Occupation (of husband if wife is writing).....

Number of children..... Ages of children.....

Did you attend parent education groups before this year?.....

If so, where did you attend and how many years did you attend?.....

Put a circle around the highest grade reached in school:

Elementary School1.....2.....3.....4.....5.....6.....7.....8.....

High School1.....2.....3.....4.....

College1.....2.....3.....4.....

Directions

1. Read these directions and do not turn the pages until the leader directs you to do so.
2. Inside this folder there are three selections of parent education reading materials which you will be asked to read. Each selection occupies one page. You will be allowed thirteen minutes for each page.
3. On the right-hand side of each page there is a list of questions which are to be answered from the selection printed on the same page. The selection on the lower half of this page is a sample for practice. Look at it.
4. You are to answer by writing on the dotted line before the number of the question the number of the bracketed passage in the selection which contains the correct answer. Begin by reading the passage completely through, then turn to the questions. (In the original test, the phrases were marked with brackets and numbered. The number of the correct phrase was placed before the question number.)

The first two questions in the sample below are answered correctly by the numbers on the dotted lines which refer to passages in the selection. Read the passage, then answer the next two questions. If there is anything about this that you do not understand, be sure to ask the leader to explain it to you during this practice period.

Sample for Practice

Answer

The diet of the nursing mother must be generous. It should contain each day at least one quart of milk and an abundance of vegetables—at least two, besides potatoes. One of these should be preferably some form of greens—spinach, cabbage, chard, lettuce, etc. The daily diet should also include at least one egg; one serving of meat or fish, some cereals; some butter; and three servings of fruit, one of which should be a serving of uncooked fruit.

1. How many eggs should be included in the daily diet of the nursing mother?
2. What examples of greens are given?
3. How much milk should be included in the daily diet of the nursing mother?
4. How many vegetables should mother's diet contain?

Work on this section until time is called.

The conditions surrounding the meal are of great importance. We recognize this in adult life, and take pains to make our tables attractive with dainty linen, glassware, fine china, and flowers. We should not think that because the child is young he is not susceptible to such esthetic influences.

The child should preferably have his own small table and a comfortable low chair of the proper height—a situation he can recognize as his. The equipment of the table should be as attractive as possible; plain oil cloths in colors, blue, rose, mauve, and yellow, vary the appearance of the table, are inexpensive, sanitary, and interesting to the child by reason of their color. His own dishes in an attractive pattern will further increase the pleasant associations of the meal. Without being fussy or elaborate, all the surroundings should be as pleasing as possible.

The same care should be taken in the serving of food. Small portions attractively served will do more to stimulate the child's appetite than any amount of direct persuasion. Here Aldrich's principle that cultivation of the appetite is of first importance should determine one's procedure.

A regular time for meals should always be observed. It is unfair to make the little child wait over the protracted periods that are sometimes made nec-

Answer

1. How do adults improve the surroundings at their own meals?
2. What desirable method of stimulating the child's appetite is mentioned?
3. Should a young child be left entirely to his own devices during the meal?
4. Who is referred to as being the author of the statement that, in training children to eat, the stimulation of the appetite is highly important?
5. How is the practice of having an ordinary maid supervise the child's meal described?
6. Why is it often desirable in a large family to postpone the time when the young child comes regularly to the family table until the child is five or six years of age?
7. What condition should determine the age at which the child comes regularly to the family table?
8. What rule should be followed *in* determining a time-schedule for children's

essary by the interruptions of adult life and it is bad training as well.

Mealtimes should be observed with scrupulous regularity. On few occasions should the child be interrupted in his play activities, . but mealtimes are one of the few occasions.

To assume a businesslike attitude, neither to bolt his food nor dawdle over a meal, to take proper-sized portions, to eat willingly what is set before him, to use table utensils correctly, and to keep a tidy place are all desirable habits of eating. Success in teaching these habits depends largely on the skill and judgment of the person in charge. That the young child eats alone should never mean that he eats without supervision. The mother or a suitable substitute should always be present to preside over the meal. To leave the child to eat off the kitchen table in the tutelage of an untrained maid is the worst possible procedure for forming good eating habits.

A certain author says that children should not come regularly to the family table till they are eight or nine years of age. Most people would prefer not to postpone this so long, but the special situation of the family rather than any arbitrary rule must determine this matter. Probably no arbitrary time should be set for making this transition from the individual to the communal meal, In a small family it may come early; in a larger family, where there are other children or several adults who may pay

meals?

9. What kind of dishes should the child have?
10. Is there anything said about training the child in the-use of the spoon, pusher, or other implements?-
11. Should the child eat what is set before him?
12. How frequently should the child be interrupted in his play?
13. What kind of cover should the child's table have?
14. Why should the child's dishes have a pleasing
15. What undesirable method of stimulating the child's appetite is mentioned?

too much attention to the child, it is better to defer the change even till the child is five or six years of age.

Work on this section until time is called.

Answer

During his early years, the child has no sense of values. His little brown mug means more to him than the most costly piece of china. The oriental rug can in no way compete in interest with the highly colored linoleum on the kitchen floor; yet what a fuss is made when something is spilled in the living room, and what a calamity if he happens to drop a piece of the family china.

The child's activity, although poorly coordinated and very uncertain at times, is not purposeless. There is a plan behind it and an end in view. As he pulls, twists, crosses, breaks, tears, cuts, and defaces it is seldom with the idea of being malicious. It simply happened, sometimes with intention, at times accidentally. He pulls at the table cover to help him get up from the floor; he twists the cat's tail because it results in noise and action; he cuts his stockings to demonstrate his ability to use the scissors; he crushes the flowers to show he is pleased with them; with chalk or pencil he has discovered he can leave his imprint on the wall or woodwork. This gives him a sense of power which he has demonstrated and from which he gets much enjoyment. It

1. Why do young children sometimes pull at the table cover?
2. How does the young child feel when his parents do not like the way he knocks things around?
3. What undesirable tendency does the parent develop who must continually ask the child to refrain from doing things?
4. Which is more interesting to the child, an oriental rug or highly colored linoleum?
5. Does a child prefer costly china to cheap china?
6. What should a mother do before punishing a young child who out holes in her dress with scissors¹?
7. When is the child lacking in knowledge of the value of things?
8. How could much of the trouble children make, by breaking things be avoided?

does not occur to him that all these newly acquired accomplishments are inflicting damage which is painful and annoying to grown-ups. He is surprised and usually sorry to find his actions have not met with approval. He is grieved by the reprimands he receives and feels keenly the injustice of his punishment. Necessary as it may be to protect the child from inflicting such damage, habitually, it is often even more important that all the circumstances and conditions that led up to the unhappy event should be carefully investigated and taken into consideration. Much of the annoyance that parents have to experience from what we term destructive tendencies would be avoided if the child could have a corner of his own in which to play.

The surroundings in which the adult lives furnish too many alluring attractions for the child. The temptation to handle and explore is irresistible to many children; and the necessity for the parents to be constantly imploring the child to inhibit his activity soon develops into chronic nagging, resulting in irritation and often anger on the part of the parent, and defiance and open rebellion on the part of the child. If the child has his own domain, whether it be a play room or merely a corner where he can carry on unmolested, much of this friction is eliminated.

It should be kept in mind that much of the activity which is looked upon by

9. What does a child often do with a pencil?
10. Do parents like to have children break things?
11. Why do children often crush pretty flowers?
12. Do children break things with the idea of making trouble for their parents?
13. Do children like to handle the things about them?
14. What is the child's reaction to constant nagging by the parent?
15. In what way is the tearing down and the breaking of things by the child often useful to the child?

the adult as being of the destructive type is to the child essentially constructive; that is, it, furnishes an opportunity to the child to inform himself regarding the physical laws that have to do with the things with which he comes in contact.

Work on this section until time is called.

Answer

There is a great difference in the way in which infants react to the necessity of taking the breast. Some babies take it at once, avidly; some take it calmly; some are trained to take the breast only with great difficulty. Some few refuse to nurse at all, but these infants usually have defective nervous systems and are "sloughed off" by the process of nature, unless they are fed by the elaborate methods made possible by modern science.

Disinclination to feeding may be due to a poor milk supply in the mother's breasts or to physical anomalies of the breasts. Sometimes a nervous child nurses poorly because of the anxiety of the mother or nurse. In general, it may be assumed that the manner in which the infant takes the breast is due to the type of nervous system with which he is endowed rather than to external causes. One is led to this conclusion not only through the study of infants but also through the study of the early life of adolescents suffering from emotional difficulties.

1. If an infant refuses to nurse at all what is usually the cause?
2. What type of child is mentioned as being affected in its nursing by the anxiety of the mother?
3. What emphasis has been placed in the past upon the second stage in learning to eat?
4. What is the last stage in learning to eat that is mentioned?
5. Are there any normal babies who must be trained to nurse from the breast?
6. When should the first additions be made to the diet of a normal breast-fed infant?
7. Why is the second stage in learning to eat of especial importance?

The early feeding response is fairly indicative of the type of general response to be expected of the child in most of his activities, and it is one of the earliest clues to his personality type and to his type of original endowment. The child may be studied at this earliest stage with a view to modifying as much as is necessary and possible his general temperament.

Considered from the psychological standpoint, there are three stages in learning to eat, the infantile, intermediate, and modified adult. These three diet periods are, of course, not clear-cut. They overlap, and with the introduction of new foods one diet period merges into the other.

During the infantile stage, the infant sucks from the bottle or the breast. Under normal conditions he has nothing at first but the mother's milk. Most pediatricians enlarge the child's diet, at about three months of age, by giving orange juice, tomato juice, sometimes prune juice, and cod-liver oil.

The weaning period from the infantile to the intermediate stage occurs usually when the baby is from eight to ten months old. This intermediate stage has not previously been considered as a definite period. By the time the child is a year old, he should be in the second stage. Whether it is a period of actual weaning from the bottle or not, it certainly must be a period of diet enlargement. Cereals, mashed potatoes,

8. What may be the result if the mother has an insufficient supply of milk?
9. What reaction of the child gives us the earliest indication of his personality?
10. Are the three stages learning to eat distinct from one another?
11. What is the chief food for the normal child in the infantile stage?
12. How are infants who cannot be taught to nurse treated in order to survive?
13. When does the infant usually pass from the first to the second stage in learning to eat?
14. What foods are added to the infant's diet during the intermediate stage?
15. Does the diet of a normal infant during the first diet stage contain any other article of food aside from milk?

scraped beef, eggs, toast, and also such vegetables as carrots and spinach are now introduced. This is a very important period in the child's life because "food peculiarities" are often developed at this time.

Often infants who adapt themselves to the diet of this period refuse to adapt themselves to its customs. They will take their milk or their water only from a bottle. This capriciousness sometimes continues far into the adult stage. Children even four or five years old sometimes refuse to eat with an implement other than a spoon, or will refuse to feed themselves at all. These are the beginning symptoms of a behavior difficulty.

Work on this section until time is called.

SELECTIONS FROM PARENT EDUCATION
READING MATERIALS

Difficulty Score	12.35
Vocabulary Difficulty	94.9
Number of Simple Sentences	23
Number of Prepositions	32
Proportion of Words in Dependent Clauses	.11

Once upon a time there was a mother whose name was Mrs. Lee. She had a sweet new baby, and it was well and strong. Mrs. Lee was a wise woman. She had talked with her doctor. She had read books about babies, and had gone to the city nurse for help. The nurse told her what clothes to make for baby. They were to be simple and loose. And not very many at first. Babies outgrow their clothes quickly. The nurse showed Mrs. Lee how to bathe babies. They talked about food and fresh air. So Mrs. Lee knew what to do for her new baby.

She nursed her baby, because mother's milk is best. It makes baby strong and healthy. Cow's milk can be used, but it must be changed. It is not sweet enough for babies, and it makes big curds in the stomach. Cow's milk must be fresh, and bottles should be put in a cool place. Spoiled milk makes baby sick.

At one month the baby had orange juice, but sometimes Mrs. Lee gave tomato juice. She used a clean teaspoon, and gave it full once a day. The doctor told her to give cod liver oil; ten drops a day. At six months the baby had strained vegetables. Spinach, lettuce, and carrots were cut up, and then they were cooked together. The nurse showed Mrs. Lee how to prepare the vegetables.

Every day the baby had a bath, and it was given at the same time each day. Mrs. Lee put four or six inches of warm water in the tub, and then she undressed the baby. She tested the water with her elbow. Water that feels right for the elbow is right for baby's tender skin. Mrs. Lee kept the baby's clothes near by. She wore a washable dress, and her hands were always clean.

The baby slept in a strong clothes basket. Babies should always sleep alone. She set the basket on two kitchen chairs, where it could not be

overturned very easily. Sometimes baby slept on the porch. The porch was screened, and no flies or mosquitoes could touch baby.

Little babies need fresh air every day, and they need it as much as sleep and food. Fresh air is cheap, while bad air may cause sickness. Mrs. Lee's baby had fresh air on the porch. Then too, she took the baby out in his carriage. This she did in the afternoon, after her work was done. She was careful not to let the sun shine in baby's eyes. The face was left uncovered, and the baby could breathe freely. Fresh air is strengthening. Babies left in hot, stuffy rooms become sick and weak. Fresh air helps prevent colds.

Every day the baby got exercise. Mrs. Lee placed him in the middle of a big bed, and there he could not roll off. He played for twenty minutes. He swung his arms, he kicked, and he rolled his head a bit. Sometimes he cried a little. Crying is good exercise, for it fills the lungs with air.

Difficulty Score	12.94
Vocabulary Difficulty	94.4
Number of Simple Sentences	20
Number of Prepositions	49
Proportion of Words in Dependent Clauses	.16

Every child needs a playground, and this is true wherever the child lives. In crowded cities it is hard to find a good place to play, for there are few private yards. Grassy places are scarce, streets are dangerous, traffic is noisy and dirt and smoke add to the confusion. Public playgrounds take the place of home yards. These playgrounds contain the most modern playthings, among which are slides, swings, horizontal bars, see-saws, jungle gyms, and other things. Some public playgrounds have a space fenced off for the little children where may be found smaller slides, sand piles, chair-swings, and wading pool.'

Roof space in cities can also be used for play, and a roof playground does not have the dust and dirt of the street. The sun shines all day on many roof playgrounds. Little tots can play there without the danger of crossing streets and they are near their homes and their mothers. Mothers living in one apartment house can take turns in looking after the children on the roof.

On a roof the children must be protected against falling off. A strong

fence of heavy wire netting should be used. The roof floor should be smooth and the skylights should be protected with wire netting. Flower boxes, awnings, and porch furniture help to make the roof homelike. A roof playground is smaller than a park playground, but most of the park playthings can be used on the roof.

For very young children a screened porch in city or town makes a good place to play. Their toys and playthings take up little room. When they are on a porch the children are near their mothers.

Not all children live in cities. In small towns, the houses usually have a yard about them, where space can be fenced in. There the small children can play in safety and freedom. A low wire fence is best. In a large yard, space should be set aside for gardens in which the children should plant both vegetables and flowers. Children learn patience and industry in caring for growing things. A child feels pride when he sees his flowers upon the family table and when he eats spinach or carrots he has grown. In the yard there should be space for a cage for rabbits or guinea pigs. A dog or cat makes an additional pet. Trees are useful for shade, for swings, and for climbing. Bird houses bring birds and cheerful song. Feeding and caring for the pets help to make children thoughtful and unselfish.

Sometimes two or three families may unite and make a common playground for their children. Back yards can be used together, and a larger playground is the result. The cost of playthings could be divided among the families. In such a group the children learn to give and take, and to share with others, and to be agreeable companions. A home playground is usually a happy place. It offsets some of the "outside attractions," and it centers the play life of the child about the home.

Difficulty Score	14.50
Vocabulary Difficulty	93.9
Number of Simple Sentences	5
Number of Prepositions	47
Proportion of Words in Dependent Clauses	.31

When a new baby came to your grandmother's home, she, perhaps, could keep all the doors and windows closed because the open fireplace and the manner in which houses were built in those days allowed free circulation of the air.

Today, however, steam heat, close-fitting windows, and tightly built houses prevent good circulation of the air unless the windows are opened.

Fresh air is needed all the time to keep the baby well. It is as important as sleep and food. There is nothing cheaper than fresh air, nor more costly than bad air. Fresh air will help to keep the baby from getting sick and will help to cure him should he be sick. Cold air will never hurt the baby if he is warmly dressed. After the baby is several weeks old, and for a month or two thereafter, if the baby cannot be taken out-of doors in his carriage, he should be taken out in the attendant's arms every day for at least a few minutes, unless he is ill or the weather is very bad.

If the weather is below the freezing point, the baby should be carefully watched to see that the nose, cheeks, chin, and hands are not frost-bitten. He should be kept out for only a few minutes at first and for a longer time as he becomes used to the outdoors. By the time the child is a few months old, he should be out-of-doors two hours daily, and he may well stay in the fresh outdoor air most of the day if the weather is suitable.

In severe or stormy weather, the baby should get his airing indoors. To do this, dress him in his outdoor clothes, place him in the crib or buggy, and open wide all of the windows in the room, closing the door to keep the rest of the house warm. When the indoor airing is over, take the baby out of the cold room, and, in a warm place, remove the outdoor clothing. Close the window, and, when the room has become warm enough, take the baby back into it. Then open the windows just enough to keep the air fresh. Delicate babies require more protection against chilling than strong ones and should therefore have a hot-water bag in the bed under the blanket at the feet in cold weather. Great care should be taken to prevent burning.

During very hot weather, the child should be taken out, both early in the morning and late in the afternoon, but should be kept indoors in the coolest room of the house during the very hot period of the day—except for the time he is getting his sun bath.

The baby should be properly protected against flies, mosquitoes, and other insects at all times. The surroundings of the home should be free from uncovered garbage, rubbish, and manure, as these things attract disease-carrying insects.

The carriage should have easy springs so as to ride smoothly, and an

adjustable hood, lined with dark material that will protect the eyes against the glare of sunshine from the sidewalk or roadway.

Difficulty Score	17.20
Vocabulary Difficulty	90.2
Number of Simple Sentences	14
Number of Prepositions	64
Proportion of Words in Dependent Clauses	.32

The young child spends a great deal of time in play. At one time grave adults looked upon this as an appalling loss of time. The expression "child's play" was used as a synonym for what was a mere useless waste of energy. A deeper knowledge of the psychology of childhood, however, has led to a juster view. We now realize that play is an essential element in growth. The child who does not play is to that extent handicapped in his after life.

The parent should plan for the child's play with as much care as is exercised in planning for his education. Play is, in fact, an element in education, and the child who has a rich play life in his early years has reason for gratitude.

There are some other aspects of the question which cannot be neglected. Street play is extremely dangerous. Mr. Earl J. Reeder of the National Safety Council has reported that street play is responsible for most of the traffic accidents in which children are involved. The child who is sent on an errand or who is on his way to school is likely to be careful about crossing the street, but the child who runs after a lost ball will be less careful. Merely as a safety measure, therefore, it pays to have plenty of facilities for home play.

The moral side also calls for some consideration. The young child needs rather close supervision. As long as he is happily playing in the house, in the yard, or in the immediate neighborhood, the mother can keep him more or less under observation in the midst of her housework. If, however, he is forced to leave the neighborhood in search of play facilities, the consequences may be unfortunate.

Every family should provide, if possible, both indoor and outdoor play space for the young child. The ideal arrangement to take care of indoor play is, of course, a room set aside as a playroom. For this purpose an at-

tic may be quite satisfactory. In the modern home, especially those using gas or oil furnaces with a consequent increased cleanliness, it may be possible to equip a playroom in the basement. Where a separate playroom is impossible, at least a play corner may be provided somewhere about the house.

A little ingenuity plus a very small outlay of money may create a very attractive playroom. The floor should be covered with linoleum or wash rugs. This is important because the floor plays a much more important role in the child's world than in the adult's, many of the former's activities being carried on just there. The furniture should be durable and of a proper size. From the child's standpoint, the ordinary chair is a table upon which one climbs at considerable effort. The wall paper and the paint should be soft-colored and not too reflective of light. Decorations which stimulate the child's imagination should be chosen. A low cupboard for playthings is a welcome convenience.

Difficulty Score	18.47
Vocabulary Difficulty	90.2
Number of Simple Sentences	5
Number of Prepositions	50
Proportion of Words in Dependent Clauses	.41

The- subject of putting away toys is one of the unfailing bugbears of parents i» training small children- It occasions many pitched battles between parent and child and rebellion in the child with emotional tension and a feeling of annoyance and frustration on the part of the parent are not infrequent concomitants.

It seems a good place to try out a type of control based on insight, a definite plan of action and satisfaction of the child's needs. First, then, let us ask why the child does not want to put away his toys. Conceivably for the same reason that we do not answer our letters, make up our accounts, or keep our bureau drawers in order. It is a bother to do so; tidiness is a social virtue—an end acquired after much pains and certainly not a primitive impulse. Yet this is one point where the freedom of the play impulse would seem to call for regulation and training.

It is plain also that in many cases the putting away of toys is not the simple matter of the child's complying with adult wishes; it has been

erected into a test of wills by bad handling of early situations. It thus becomes an anticipated battle ground on which the child tries out his persistence against that of the parent. The child usually gauges accurately the parent's resistance, as in the case of one little girl who was heard to mutter, "I'll pick up one block, and she'll do the rest." It is highly important that such a situation should not be allowed to arise, where the child consciously tries to get control over the adult, for its effect on general discipline is bound to be unhealthy, and radiate to other settings.

What, then, should be our plan? First, not to allow ourselves to become emotional over the situation. What does it matter that the child storms and protests so long as his behavior does not cause the parent to lose control of herself and therefore of the situation.

Secondly, the picking up of toys should be treated as a learning situation. There should be proper and easily accessible places for putting the toys away; the child should be shown where and how to put them, and he should be helped in doing so. Patience must be shown with his fumbling efforts and slow pace as well as firmness in seeing that the work is actually accomplished. No new worth while accomplishment can be acquired quickly, and the process is doubly hard when it does not make any special appeal to the child's natural interest.

As regards satisfaction, the parent should be content at first if she can avoid emotional conflicts and secure some amount of pleasant cooperation from her child. Rewards are taboo since they usually only complicate the issue. By degrees the child should begin to take a pride in his accomplishment if he puts his possessions in order.

Difficulty Score	18.56
Vocabulary Difficulty	87.8
Number of Simple Sentences	10
Number of Prepositions	62
Proportion of Words in Dependent Clauses	.24

It sometimes happens that feeding problems become acute in the pre-school period. Prevention of food fads should begin in the first year, when new foods are gradually being introduced, a half teaspoonful at first. Each new food should be very carefully cooked to prevent disagreeable associations. Spinach, for example, should be washed free of sand, cooked ten

or fifteen minutes finely chopped, seasoned with salt and butter, and garnished with a spoonful of crumbled egg yolk on top, if desired. Any new food should be introduced on a day when the child is feeling reasonably jolly. The rest of the family should openly show their enthusiasm for all the good foods. A reasonable portion of the food should be put on the child's plate, and no direct reference to his eating should be made. If such a course has been persistently pursued the chances are that the child will eat his food as the others do. Well-cooked food, a cheerful atmosphere, other people obviously enjoying the food, and no attention focused on the child himself are conducive to good eating habits.

In overcoming bad feeding habits, more difficulty is encountered when bad eating habits have gained headway. One child teases for tastes of all the foods the adults are eating. Some one at one time began giving him a taste of this and a taste of that. This practice weakened the association between sitting at the table with adults and eating only the food put on his own dishes. Probably adults would be healthier if they ate the same food as the children but in most cases, they do not. If the child, therefore, eats at the grown-ups' table, he must have the habit of being content with the food that is placed before him. It is probably better for preschool children to eat at a little table by themselves, except on special occasions.

Sometimes children refuse to eat at mealtime—no cereal, no eggs, and above all no vegetables. Coaxing is ineffective. Scolding is bad physiologically as well as psychologically. To correct this bad habit Dr. Frank H. Richardson advocates the following procedure: Small portions of food suitable for the child are put on his plate. He is required to sit down at the table promptly each meal. There is to be no word of reproach or command; not the slightest effort should be made by any of his elders to induce him to eat. At the end of thirty minutes he is to leave the table "whether he has eaten much, little, or nothing." The length of time before he begins to eat under this new regime varies with the "determination of the child, his knowledge of the perseverance of his parents, and the degree of appetite perversion that has been wrought by the course previously pursued. It may safely be stated, however, that the cure is usually 'well on the way long before the conclusion of the second day.'" Aldrich gives a sane, honest, and highly practical analysis of the psychologic and physiologic factors in this common disorder, lack of appetite.

Difficulty Score	18.58
Vocabulary Difficulty	91.0
Number of Simple Sentences	12
Number of Prepositions	62
Proportion of Words in Dependent Clauses	.35

Each child should sleep in a bed alone. It is desirable that each individual have a room of his own, but this is not possible in the average household.

A bed not too soft and yielding is preferable. The covers should be warm but light. Some children get up in the morning exhausted from the weight of the heavy comforters imposed on them.

Children should not sleep in the same room with the parents after early babyhood. Neither should they sleep in the same bed with a person of the other sex. This is a matter of importance, frequently unnoticed.

Proper ventilation is an important element in the child's sleep. Without wide-open windows and plenty of fresh air, sleep is of less benefit to the child. What goes on during sleep is not precisely understood, but it is known that certain poisons, manufactured in the body as a result of fatigue, are being thrown off. One has only to recall the heavy, sluggish feeling that pervades the body after one has fallen asleep in a warm close room, to realize that without circulating air, the body poisons cannot be neutralized.

So-called "laziness" in children is often found to be the result of lack of sleep, or of sleep under unfavorable conditions. Since the quality of sleep has as much to do with the individual's welfare as the quality, the need will be seen for good ventilation and proper bedding.

There are good reasons why the young child should be in bed at an early hour. The evenings should be a time of relaxation and recreation for older members of the family. Things which are relaxing for parents are not so for children. Music, conversation, laughter serve as a necessary outlet for the adult, but act as stimulants to the child, who can stand stimulation only for limited periods.

The young child who is nervously stimulated by entering into the family life at dinner and afterward is not in a mood to fall asleep quietly. Parents often say, "Bobby might just as well stay up. He lies awake for an

hour or more, and gets more and more restless as he is kept in bed." This is one of the points at which the parent may ask himself, "What is the stimulus which is bringing about this undesirable behavior?" For without a cause, no habit can take root. Perhaps the child has had his supper too early or has had too light a meal. Perhaps he has had no outdoor exercise since his afternoon nap. Perhaps he was allowed to sleep too long at that time and is actually not sleepy at his proper bedtime. This is sometimes the case when the mother is negligent about having the nap hour come early enough in the day. A child who has already formed a habit of not falling asleep until the older members of the family go to bed must necessarily sleep poorly during the time when a change of habit is being set up.

Difficulty Score	21.16
Vocabulary Difficulty	91.5
Number of Simple Sentences	6
Number of Prepositions	60
Proportion of Words in Dependent Clauses	.35

The learning of property rights is very difficult in our homes as they are arranged at present. Most property in the homes is held in common, and is replaced from a common fund. If a child assumes that conditions are the same outside the home, the result is fatal. Crimes against property are the commonest of all crimes, and the discrepancy between the organization of the home and the organization of society is certainly one of the things that is greatly responsible for them.

The child must be taught the theory of personal ownership. In order to do this, he must have toys of his own and his own personal toilet articles, and he must be paid for some of his small jobs and purchase some of his own toys. He must be taught, from the beginning, absolute restraint with regard to the property of other members of the family, and also in the houses of other people.

Adults who are not members of the family often help to confuse the issue. A child will pick up some toy or some trinket that has no value, and when the parent tells him not to take it, the other person will say, "He may have that," so that the child does not perceive the line between his possessions and those of another. The child values many things which to the adult seem worthless objects, and the adult often again confuses the issue by destroying things which are valuable to the child. He may have a col-

lection of his own which is to the adult of the veriest worthlessness—a few tacks, nails, strings, marbles, and old broken things; but for the child, it marks the beginning of property consciousness; if the adult does not protect his interests and preserve his property, then he cannot be expected to respect those of the adult in return.

The privileges of the youngest child in the house often are confusing in this respect. The baby is given almost anything that he wants and if he reaches for the toy of an older child, one of the parents is likely to say, "Oh, give that to the baby; he's just a baby and he wants it." This attitude is fatal not only in the teaching of property rights to both children, but to the later relationship between the two. Much unhappiness and many unfortunate relationships in the family can be traced to just such early scenes as this.

People are often praised for being unusually unselfish. To be called "unselfish to a fault" is supposed to be a compliment. "He would give away the coat off his back" is another compliment that is of doubtful value, for what is usually not told is that the individual who would give the coat off his back is more than likely to give away the other fellow's as well and is very likely to give the coat to some one who does not need it, while some deserving or dependent person is left coatless. Promiscuous giving is just as much a sign of maladjustment as promiscuous keeping.

Difficulty Score	21.51
Vocabulary Difficulty	92.0
Number of Simple Sentences	6
Number of Prepositions	54
Proportion of Words in Dependent Clauses	.42

Saving should be regular, that is to say, the child should put aside a portion of his money weekly. It does not train the child if he saves two cents this week, five cents next week, and nothing the next. It would be far wiser to save three cents each of these weeks, then the habit of regularity in saving would begin, to be set up. Savings should be only a part of the total, since a child gets no pleasure out of money which must all be saved. He gets experience in spending only by spending. He will not enjoy his allowance which you give him unless he has this experience. It is of no value to him to save unless he also knows how to spend. The money which is saved should be put in the bank by the child himself. If it is to be

put in a savings bank, the child should go with his mother and father to deposit the money and should see the book in which the deposit is recorded. Parents should never use the money which the child has saved except in an emergency, and then only for the child himself.

Help the child to be specific about spending his money, except that part of it which is allowed, for free spending money. A mother who had given her child an allowance of 35 cents on her ninth birthday without any instruction as to the use of this, found the 35 cents on the window sill some time later. She returned it to the child only to find it on a chair the next day. This child had never been trained in the use of money and had never spent money herself. What the child saves should be saved for some immediate purpose—a doll, a piece of doll furniture, a present for father or mother, or, with the older child, a strong truck or a bicycle are perfectly legitimate—things to save for. The child gains much pleasure from the object which his pennies have bought. Such saving teaches him far more than the mere collection of pennies and dimes in a small bank, the contents of which in time are transferred to a larger bank. Money so collected means nothing.

A little girl, who had always to save her Christmas and birthday money, oil being asked on her tenth birthday what she wanted, said, "I want just enough money so that I won't have to put it in the bank." To her, money put in the bank was just money lost sight of.

One thing which must always be kept in mind in teaching the value of money is that the child should understand the relation, between, cost and the object itself. All children begin to trade long before they begin to realize actual coat in dollars and cents. To trade a kiddy kar worth 4 dollars for a pen knife worth 25 cents is bad business. The child should have the exchange explained to him.

Difficulty Score	21.56
Vocabulary Difficulty	50.2
Number of Simple Sentences	4
Number of Prepositions	66
Proportion of Words in Dependent Clauses	.46

When our little children of two or three years begin to show an interest in books, we are overjoyed, and gladly spend part of every day looking at

pictures in story books with them, telling them stories, and reading simple things to them. We find that the little ones are delighted with cumulative tales, such as that of the pig who would not go over the stile. Other stories that have the same element of repetition, *Chicken Little*, for instance, or *The Three Pigs*, we eliminate from our repertory because of their emphasis on fear. *Red Riding Hood* we throw overboard for this reason.

Jingles and rhymes appeal to the small children by reason of their cadence. *Mother Goose* is especially good, provided we omit the unpleasant verses. Animal stories and stories about children are favorites, but we have to make sure that there is no sadness in the tales. There are many delightful primers today, prepared for use in teaching a child to read, which are just right to read to a little one of three or four.

The poems of Stevenson come next. There are many sweet little poems by authors great and small suitable for the years from three to six. Do not decide that a little tot will not like a certain poem until you have read it to him. Often, the very poem or portion of a poem that you were about to discard because it seemed to you too difficult for the child turns out to be the stanza he enjoys most, and perhaps tries to repeat.

Myths, legends, and fairy tales have an important mission in helping to arouse the dormant imagination; but they should not be allowed to become the exclusive fare of the hungry child-mind, as their common theme of the overthrow of the strong and powerful by the poor and weak fosters daydreaming. The child of six to twelve years sometimes prefers to read fairy tales, reveling in the adventures of the weakling or the youngest son who becomes the prince of the land, rather than to go outdoors and meet his playmates in rough and ready contest. At their hands he might be worsted, while in the book he may identify himself with his chosen hero, and reap the rewards of hazards he has not run.

It is the girl, oftener than the boy, who thus withdraws from the world of hard, everyday, facts and seeks refuge *in* fairyland. By no means forbid her the books of fairy tales, but encourage her to play outdoors more, and lead her to interest herself in practical activities. The world of fancy is a winged horse that must be tamed with the golden bridle of application to the affairs of life today, if this Pegasus is to be a useful servant of its young admirer. To win a child away from indulging only his play of fancy in reading, choose simply told biography, history, and natural science, or colorful stories of travel.

Difficulty Score	22.00
Vocabulary Difficulty	92.3
Number of Simple Sentences	6
Number of Prepositions	42
Proportion of Words in Dependent Clauses	.36

The period of most intense fears is between the ages of five and seven. Apparently, the feeling which the infant has attached to a frightening experience can be recalled in the dark by the fancy running wild. It is imaginative children who naturally suffer most, but they can call up bright fancies to cheer them. Robert Louis Stevenson must have had a lovely time in the dark, seeing circuses and things, as he tells us in his poem which begins:

All night long and every night
When my mamma puts out the light,
I see the circus passing by
As plain as day before my eye.

Although fear is a common human emotion, it is not universal, and once in a while we find a child who has no instinctive fear. If such a child is not frightened, he may remain quite ignorant of the feeling for many years. I know a boy who, at the age of five, was unacquainted with the sensation of fear, and, never having been frightened, also did not know the meaning of the word "fear." He had heard it used by other children and knew that it was something unpleasant. One day at dinner he said to his mother, "You know, I think I am afraid of spinach," meaning that he did not like it. It was evident that the feeling of fear was quite foreign to him.

Since fear rests so largely on ignorance, curiosity is its enemy, because it dissipates ignorance. A little boy who had a certain fear of the figures in the mirror that were so vivid and yet so unreal used to try to come into a room in which there was a large mirror, and steal upon the objects of his curiosity unawares. His double was always there as soon as he, and caught his eye; but the child lost his fear only after he became familiar with the characters in the looking glass. In the same way curiosity will often help the child to become so well acquainted with the source of his fears as to drive the latter quite out of his experience.

We must be careful to avoid confusing fear and caution. Fear is not

necessarily related to any real danger. Caution is a direct outcome of the knowledge of danger. Two little boys were watching a young man shooting off firecrackers. Whenever a bunch was lit the older boy stepped away, while the younger one held his ground. Some one taunted the older boy, saying, "You see, Robert is not afraid, and you are," To which he very sensibly replied, "I ain't afraid neither, but Robert doesn't know that lie might get hurt, and I do."

While we do not wish our children to be cowards, neither do we want them to be reckless. Caution and courage may well go together in the child's character. But constantly warning the child against possible danger does not develop caution; it is more likely to destroy all spontaneous action.

Difficulty Score	23.37
Vocabulary Difficulty	87.9
Number of Simple Sentences	7
Number of Prepositions	57
Proportion of Words in Dependent Clauses	.30

It seems inconceivable that parents should be willing to hand over the sex instruction of their children to an outsider. By so doing they are building up a barrier between the children and themselves, making it a topic not to be broached between them. Yet many parents seem to feel that they have discharged their responsibility if they send their children to a school where a master is reputed to give his pupils sex instruction. They seldom know either what or how he teaches. No. one would wish to belittle the important function of the school in the sex education of children; a teacher who is alive to the indications of developing sex consciousness in his pupils can often make a significant if incidental contribution to the sex training of older children. This, however, can never take the place of the fundamental instruction in the early stages of the child's life which only the parents could give; it should be supplementary to the home instruction. Parents cannot evade this without doing a grave injustice to the child.

Again, parents sometimes give the child a book on sex to read, and then feel that their duty has been done. This is a poor substitute for the frank discussion between parent and child of the child's first questions. A book 'written for the child, if carefully prepared, may, however, be a use-

ful ally to the parent's efforts. By the time the child can read he will derive from the book impressions more precise than those which he gains from conversation alone. This is especially the case if the book is well illustrated—indeed, even a young child will find illustrations fascinating and will learn much from them. Karl de Schweinitz' *Growing Up* is a model for use with children. It is clear, simple, and explicit in its exposition, and the illustrations are delightful.

If a child has been allowed to reach an age approaching puberty without any plan of instruction in regard to sex, the best that a parent can do is probably to give a carefully selected book to be read, telling the child to come and ask any question after digesting what is contained in the book. This is distinctly a second best, due to neglect of the matter at the proper time.

Care should be taken that any books the child is allowed to read are accurate and free from sentimentality. Simple direct statements of fact are desirably without any glaze of romantic fancy. The child wants essential facts and a fanciful setting will only distract him from his purpose or make his mind a fruitful soil for a noxious crop of half truths. The fairy tale atmosphere is out of place in the realm of plain matters of fact and behavior.

The same rule applies to the manner in which instruction about sex should be given. Parents should be as casual and matter-of-fact in manner and as plain spoken in regard to what they have to tell, as the child himself is.

Difficulty Score	23.83
Vocabulary Difficulty	83.4
Number of Simple Sentences	7
Number of Prepositions	62
Proportion of Words in Dependent Clauses	.8

To the young child, the approbation of the parent is his highest reward for good behavior. As the child grows, we must be careful that his ideas of why he is being "good" grow too—that his conduct must be satisfactory not just because it is pleasing mother or father, but because such conduct is desirable socially. We must beware of using material rewards, or the child will come to rely on them, rather than on the inner faculty for

discrimination which is becoming his habitual guide.

At each stage of the child's development, we must appeal to the highest motive possible to his understanding.

The very young child can appreciate only those satisfactions which are immediate. He cannot look far ahead, so his rewards to induce certain behavior must come at once. As he grows older, more remote aims and objects toward which to work should be steadily in line. Ways must be found of keeping up his interest in the immediate undertaking as a means to an end. The school bank, for example, will not promote the habit of thrift, unless the child has something definitely in mind for which to save.

Success, even in small things, proves a powerful stimulant toward higher things. The encouragement engendered by praise, if not unwisely depended on, can be an incentive toward greater endeavor. Constant blame, on the other hand, drags the child down, his attitude becomes despondent, and the material reward is only a sop.

We should keep in mind the fact that adults, as well as children, work for rewards. "Joy in one's work," "the satisfaction of a task well done" are too intangible and vague to represent reward to the young child. Only by degrees may we substitute rewards on a higher level, and work toward a sublimation of desires. The degree in which the discipline will be constructive is determined largely by the relation between parent and child. One characteristic peculiar to childhood, when the brain is ready and open to learning of all types, proves a double-edged sword—namely, that the child is just as plastic for poor responses as for good responses.

The child at this period reveres his parents. He has come to love them as the source of his pleasures; they are the patterns which he copies in Ma behavior. The parent who is interested in helping the child build up desirable habits will take advantage of this early suggestibility by himself exhibiting the desired behavior.

We do not want the child to be too suggestible to any type of stimulus, else he may in the end exhibit "lack of purpose and acceptance of suggestions as the easiest line to follow."

When the child is made the butt of incessant suggestion and persuasion, he may gradually become stubborn and negative in his attitude. Children, as well as adults, like best to work with those who leave some-

thing to their initiative and ability, where suggestions and comments are not constant.

Difficulty Score	24.13
Vocabulary Difficulty	84.6
Number of Simple Sentences	9
Number of Prepositions	69
Proportion of Words in Dependent Clauses	.29

Obedience implies submission to the control of others, and the "others" to whom we refer are parents or those who are in authority. It is not instinctive, like hunger. It is something to be acquired by experience and training. There are, however, certain innate tendencies, such as imitation, love of approbation, and plasticity, that can be utilized in developing habits leading to obedience. Obedience should not be looked upon, as an end in itself. Mere submission to parental authority may prove to be an attitude harmful to the child in later life. It may soon deteriorate into a willingness to conform to the wishes and desires of any one who has a strong will. Obedience is a means to an end. That end is self-control and restraint.

This means not only conformity to social laws and customs as well as family rules' and regulations, but also obedience to principles and standards which go to make up the moral fabric of personality.

Obedience in itself cannot be taken as a criterion of character, and the¹ ease with which it is developed is not an indication of his ability to make a satisfactory social adjustment later in life. The plasticity of a child's mind is the greatest factor to be reckoned with in the development of this particular trait. "'Plasticity,' says William James, 'means the possession of a structure weak enough to yield and strong enough not to yield all at once....the phenomena of habits in human beings are due to plasticity.'"

The parental attitude that demands obedience at all costs leads to the use of such drastic and unfair measures that, in gratifying their desire for power and authority, the parents may lose all those finer feelings that should exist in the parent child relationship. How often mothers say, "I can't do a thing with him, but it takes just one look from his father." This tells the story of » child whose sole motive for desirable conduct is fear. He has not been taught how much pleasure and satisfaction can be de-

rived from honest effort, cooperation, and team play. He is indifferent to approbation. Uninfluenced by either praise or blame, he directs his course, or rather has it directed for him, toward a future filled with difficult social situations.

Disobedience may be looked upon as a form of self-assertion that is misplaced. It may be brought about by physical illness, when the child's sense of well-being is at low ebb. Forceful and uncompromising measures on the part of the parents that tend to push rather than to lead the child are common causes of stubbornness, which is often construed as disobedience. Doubts in-decisions, jealousy, and fear all tend to disturb the emotional balance of the child and lead to conflicts with the environment that appear as disobedience.

The negativistic phase of the child's life is recognized as being a normal reaction occurring in the process of growing up.

Difficulty Score	27.13
Vocabulary Difficulty	74.0
Number of Simple Sentences	9
Number of Prepositions	68
Proportion of Words in Dependent Clauses.	.22

Since the development of adaptive behavior is a genetic problem, it is of fundamental importance to record chronological age and to build up data with systematic reference to agedness. But it is questionable whether mental age equivalents can be of any great service. Enough sins have been committed in the name of mental age to justify extreme caution in affixing such ages to the behavior complexes of different species.

A limited psychometric point of view would make us content to measure separate capacities and degrees of difficulty, and conclude that the superficial equivalences bespeak a common behavior equipment. As a matter of fact, such similarities are but the starting point for the reconstruction of the mental life of the infrahuman creature. The actual "meaning" of any behavior in any individual can be appreciated only in terms of its genetic configuration, that is, its historic relation to his past and future career. The relation of the measured behavior to the total economy of the individual's life cycle is significant. The determination of ontogenetic ages of representative abilities in different species becomes significant when

these ages are used to define comparatively the characteristic growth patterns and developmental cycles peculiar to each species.

That a growing orang-utan will at a given nascent age reach for a cube is an interesting fact and establishes a certain similarity with the six months human infant; but the similarity is offset by a mass of contextual differences. The similarities are apparent; the differences are hidden and profound. Yet in these differences resides the uniqueness of human infancy; a uniqueness which makes itself manifest very early. At no phase of the entire life cycle are infant and simian the same. The human characteristics are not added as an installment upon a lower primitive stage, but they inhere in the beginnings of the infant's behavior.

The preeminence of human infancy lies in the prolongation and deepening of plasticity. There is specific maturation of behavior patterns as in subhuman creatures; but this proceeds less rigidly and the total behavior complex is suspended in a state of greater formativeness. This increased modifiability is extremely sensitive to the social milieu and is constantly transforming the context of adaptive behavior. In the impersonal aspects of adaptive behavior of the nonlanguage type (general practical intelligence) there is a high degree of early correspondence between man and other primates. This correspondence may prove to be so consistent in some of its elements as to suggest evolutionary and even recapitulatory explanations.

But transcending, pervading, and dynamically altering that strand of similarity is a generalized conditionability and a responsiveness to other personalities, to which man is special heir. This preeminent sociality exists even through the prelanguage period, long before the child has framed a single word. Herein lies his humanity. This humanity does not wait for upright posture and speech. It is present at birth. It came, to be sure, late in the history of the race, but it arrives early in the history of the individual.

Difficulty Score	27.94
Vocabulary Difficulty	74.3
Number of Simple Sentences	9
Number of Prepositions	68
Proportion of Words in Dependent Clauses	.39

Acceleration of development, likewise, is typically an inherent biological characteristic of the individual, most probably hereditary in nature. There is no convincing evidence that fundamental acceleration of development can be readily induced by either pernicious or enlightened methods of stimulation. Through sheer conditioning and training it is possible to teach both infants and animals prodigious tricks. It is possible, also, that certain kinds of conditioning may exert a deep augmenting effect upon the dynamics of individual growth, reaching the endocrine constitution. This would be a secondary, derived kind of acceleration comparable to the reduction of development in secondary amentia. It is a theoretical possibility rather than a frequent clinical manifestation. Abnormal forms of precocity encountered clinically are likely to be partial and unsymmetrical. They constitute atypical deviations; they may be associated with infantilism and with unusual or pathological glandular conditions. The wholesome variety of acceleration found with superior endowment is really part and parcel of that endowment, a symptom of intensified growth, a fundamental individual difference, characteristic of, because necessary to, the developmental mechanics of certain kinds of ability. If the methods of biochemical measurement were available, it might be possible to determine certain differences in the energetics or dynamics of these rapidly growing infants, even in the first months of existence, when we could scarcely attribute their precocity to special educational or environmental stimulation.

If the superior individual as a rule grows mentally not only faster but for a longer time, this lengthened span may be regarded as primarily a manifestation of inherent endowment. That secondary, derived factors also come into play will be presently noted. The interspecies and interracial differences in the duration of plasticity are doubtless correlated with differences in organic constitution. Within limits, comparable individual differences in the growth cycle of man may be presumed to have a similar basis.

This does not, however, exclude the operation of extrinsic influences. Whether one regards thyroid extract as an article of diet or as a biochemical activator, it is certain that in some instances it affects the metabolism of the body so profoundly as to have a demonstrable effect upon both physical and mental growth. Here, then, an extrinsic factor modifies the tempo and trend of development. It is also possible that future insight into endocrine physiology will actually lead to a postponement and ameliora-

tion of senility. Then again the dominance of hereditary determination would give way to environmental regulation. For similar reasons it is probable that superior physical hygiene will continue to have a favorable effect upon growth, particularly in cases of previous neglect or partial deprivation. In one or two of the "atypical" growth cases reported in the foregoing section, it is possible that obscure but genuine alteration of the physiological economy were responsible for the psycho-developmental improvement. On the whole, the stability of the developmental trend and tempo is more conspicuous than its sensitiveness to "external" influences.

1935—Gray and Leary: What Makes a Book Readable

IN 1935, Dr. William S. Gray, Director of Research in Reading at the Graduate School of Education, University of Chicago and Bernice Leary of St. Xavier College, Chicago, published their landmark work in reading research, *What Makes a Book Readable*. Like Dale and Tyler's work, it attempted to discover what makes a book readable for adults of limited reading ability.

Dr. Gray was responsible for many of the *Dick and Jane* books, which attempted to relate to the real world of young readers. More than any single individual, he influenced the teaching of reading around the world. He advocated the “whole language” approach to reading. He taught that students learn reading and writing as they had learned how to speak and listen, as methods for coping with their real world.

The criterion used by Gray and Leary included 48 selections of about 100 words each, half of them fiction, taken from the books, magazines, and newspapers most widely read by adults. They established the difficulty of these selections by a reading-comprehension test given to about 800 adults designed to test their ability to get the main idea of the passage.

No subsequent work has examined readability so thoroughly or investigated so many style elements or the relationships between them. The authors first identified 228 elements that affect readability and grouped them under these four headings:

1. Content
2. Style
3. Format
4. Features of Organization

The authors found that content, with a slight margin over style, was most important. Third in importance was format, and almost equal to it, “features of organization,” referring to the chapters, sections, headings, and paragraphs that show the organization of ideas (See Table I on the next page).

They found they could not measure content, format, or organization statistically, though many would later try (See below, “The measurement of content”). While not ignoring the other three causes, Gray and Leary

concentrated on 80 variables of style, 64 of which they could reliably count. They gave several tests to about a thousand people. Each test included several passages and questions to show how well the subjects understood them.

TABLE I

SUMMARY OF JUDGMENT CONCERNING THE RELATIVE INFLUENCE ON READABILITY OF THE FOUR MAJOR CATEGORIES

MAJOR CATEGORY	ALL PERSONS			LIBRARIANS			PUBLISHERS			OTHERS INTERESTED IN ADULT EDUCATION		
	M.	σ	Range	M.	σ	Range	M.	σ	Range	M.	σ	Range
I. Format.....	20.26	7.68	45-5	24.13	7.64	45-7	17.08	5.20	25-10	17.92	6.27	30-5
II. General Features of Organization.....	15.38	7.04	40-3	15.71	6.59	26-3	15.42	6.27	30-5	15.20	7.24	40-5
III. Style of Expression and Presentation.....	30.71	9.17	50-0	32.74	8.39	50-20	32.92	8.27	50-25	27.50	9.24	40-5
IV. Content.....	33.64	13.11	75-7	27.42	9.95	50-7	34.58	12.83	50-10	39.37	12.54	75-20

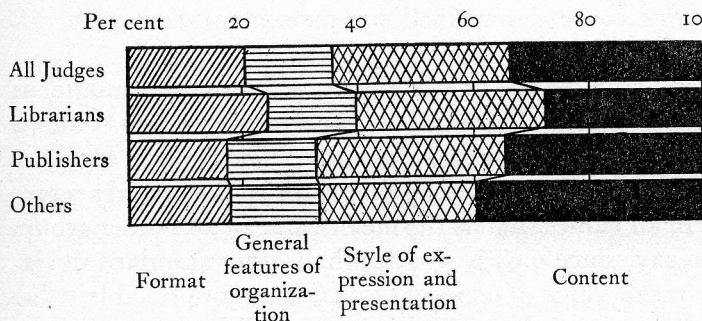


FIG. 1.—Opinion concerning the influence of classified factors on readability

Fig. 1. The four major categories of readability (Gray and Leary, p. 31).

Having a measure, now, of the difficulty of each passage, they were able to see what style variables changed as the passage got harder. They used correlation coefficients to show those relationships.

Of the 64 countable variables related to reading difficulty, those with correlations of .35 or above were the following (p.115):

1. Average sentence length in words: -.52 (a negative correlation, that is, the longer the sentence the more difficult it is).
2. Percentage of easy words: .52 (the larger the number of easy words

the easier the material).

3. Number of words not known to 90% of sixth-grade students: -.51
4. Number of “easy” words: .51
5. Number of different “hard” words: -.50
6. Minimum syllabic sentence length: -.49
7. Number of explicit sentences: .48
8. Number of first, second, and third-person pronouns: .48
9. Maximum syllabic sentence length, -.47
10. Average sentence length in syllables, -.47
11. Percentage of monosyllables: .43
12. Number of sentences per paragraph: .43
13. Percentage of different words not known to 90% of sixth-grade students: -.40
14. Number of simple sentences: .39
15. Percentage of different words: -.38
16. Percentage of polysyllables: -.38
17. Number of prepositional phrases: -.35

Although none of the variables studied had a higher correlation than .52, the authors knew by combining variables, they could reach higher levels of correlation. Because combining variables that were tightly related to each other did not raise the correlation coefficient, they needed to find which elements were highly predictive but not related to each other.

Gray and Leary used five of the above variables, numbers 1, 5, 8, 15, and 17, to create a formula, which has a correlation of .645 with reading-difficulty scores. An important characteristic of readability formulas is that one that uses more variables may be only minutely more accurate but much more difficult to measure and apply. Later formulas that use fewer variables may have higher correlations.

Gray and Leary’s work stimulated an enormous effort to find the perfect formula, using different combinations of the style variables. In 1954, Klare and Buck listed 25 formulas for children and another 14 for adult

readers. By 1981, Klare noted there were over 200 published formulas.

Research eventually established that the two variables commonly used in readability formulas—a semantic (meaning) measure such as difficulty of vocabulary and a syntactic (sentence structure) measure such as average sentence length—are the best predictors of textual difficulty.

Some experts consider the number of morphemes for each 100 words to be a major contributor to semantic (meaning) difficulty and the number of Yngve word depths (branches) in each sentence to be a major contributor to syntactic (sentence) difficulty. One study (Coleman 1971) showed that Flesch's index of syllables for each 100 words correlates .95 with morpheme counts. Another study (Bormuth 1966) found that the number of words in each sentence correlates .86 with counts of Yngve word depths. Measuring the average number of syllables per word and the number of words in each sentence is a much easier method and almost as accurate as measuring morphemes and word depths.

—WHD

1944—The Lorge Readability Index

Introduction

IRVING LORGE was interested in psychological studies of language and human learning. At Columbia University's Teachers College, he came under the influence of Lyman Bryson. In 1938, he published *The Semantic Count of the 570 Commonest English Words*, a frequency count of the meaning of words rather than the words themselves. In 1944, he was co-author of E. L. Thorndike's last book, *The Teacher's Word Book of 30,000 Words*.

Lorge wanted a simpler formula for predicting the difficulty of children's books in terms of grade scores.

In a 1939 article, "Predicting Reading Difficulty of Selections for Children,"¹ he demonstrated that new combinations of variables gave predictions of higher accuracy than the Gray-Leary formula. Lorge again established that "vocabulary load is the most important concomitant of difficulty."

In 1944, Lorge published his new Lorge Index in the *Teachers College Record* in an article entitled, "Predicting Readability," reprinted here. In 1948, Lorge² published corrections to his formula, which are given here in the footnote on p. 182.

Though created for children's reading, Lorge's Index was soon widely used for adult material as well. Where Gray and Leary's formula had five elements, Lorge's had these three, setting a trend for simplifying the formulas that was to follow:

- Average sentence length in words
- Number of prepositional phrases per 100 words
- Number of hard words not on the Dale list of 769 easy words.

Lorge's use of the McCall-Crabbs *Standard Test Lessons in Reading* as a criterion of difficulty greatly simplified the problem of matching readers

¹ Lorge, I. 1939. "Predicting Reading Difficulty of Selections for Children." *The Elementary English Review*. Vol. 16, 229-233.

² Lorge, I. 1948. "The Lorge and Flesch Readability Formulae: A Correction." *School and Society*, Vol. 67, pp. 141-142.

to texts. Although these passages were far from ideal, they remained the standard criteria for readability studies until the studies published by John Bormuth of the University of Chicago in 1969.

During and after World War II, the government bureaus and the Armed Services of the U.S. searched for efficient ways of assessing the readability of their materials. Lorge's formula was one of the best available, and it came into wide use.

Lorge's work established the principles for the readability research that would follow and set the stage for the Dale-Chall and Flesch Reading Ease formulas, which were introduced in 1948.

—WHD

Teachers College Record

Vol. XLV

March, 1944

Predicting Readability

IRVING LORGE

Associate Professor of Education, Teachers College

WHAT a person understands of the material he reads depends upon his general reading ability and the readability of the text he is reading.. His reading ability, moreover, depends upon his intelligence, education, environment, and upon his interest and purpose in reading. The readability of a text depends upon the kind and number of ideas it expresses, the vocabulary and its style, and upon format and typography.

Reading comprehension must be viewed as the interaction between reading ability and readability. Reading ability can usually be estimated by a person's success with an adequate reading test. Readability, however, must be measured in terms of the success that large numbers of persons have in comprehending the text. In measuring the readability of texts, the material is presented to a random sample of persons whose reading, ability is known. The readability of the text is assigned the average reading ability score of the sample. In assigning the average reading ability score as an estimate of the readability of a text, one must assume, of course, that the variations in people's interests and purposes in reading are balanced.

THE CRITERION OF READABILITY

Research in readability originated in the desire to grade textbooks and other materials for use in the elementary grades. Subsequently, the research activities were extended not only to demonstrate the lack of adequate reading materials for adults, but also to suggest how more adequate materials might be prepared. The research in readability became a search for a relationship between structural elements of the text and some meas-

ure of success with that text by large groups of readers. The literature of readability is concerned with the criterion for readability as well as with predictors of readability. In terms of the definition of readability, the criterion must be a measure of success that a large number of readers would have with the text. Such a criterion may be obtained by judgment or by more objective methods of appraisal. The method of judgment utilizes ratings of estimated difficulty of texts. Recently, Flesch¹, using the method of judgment, assumed that the text in magazines like *The American Scholar*, *Foreign Affairs*, and *The Yale Review*, was more difficult (less comprehensible to a random sampling of readers) than the text in magazines' like *True Confessions*, *Modern Screen*, and *Romantic Story*.

Therefore, on the assumption that magazines are written on different levels of readability, he assigned criterion level scores to groups of magazines. More objective measures of readability, however, have been used, Vogel and Washburne's criterion for the readability of a book was the average paragraph meaning score on the Stanford Achievement Test of children who had read and liked that book. Gray and Leary used the criterion of the average reading comprehension test score of a group of adults as an estimate of readability.

VARIABLES USED TO PREDICT READABILITY

The variables used to predict readability are aspects of the text, e.g., vocabulary load, sentence structure and style, and interest. One or more measures of vocabulary load is used as a predictor in every study of readability. The more usual measures are the following:

- (a) Number of running words.
- (b) Percentage of different words.
- (c) Percentage of different infrequent, uncommon, or hard words,
- (d) Percentage of polysyllabic words.
- (e) Some weighted measure of vocabulary difficulty.
- (f) Vocabulary diversity (related to b).
- (g) Number of abstract words.
- (h) Number of affixed morphemes (prefixes, inflectional endings, etc.).

¹ Studies referred to in this article, together with other pertinent references, are listed in the Bibliography, page [186].

Most studies also predict readability on the basis of one or more measures of sentence structure or style, e.g.,

- (i) Percentage of prepositional phrases.
- (j) Percentage of indeterminate clauses.
- (k) Number of simple sentences.
- (l) Average sentence length.

Less frequently, the prediction of readability is based on some measure of human interest, e.g.

- (m) Number of personal pronouns.
- (n) Number of words expressing human interest.
- (o) Percentage of colorful words.
- (p) Number of words representing fundamental life experiences.
- (q) Number of words usually learned early in life (related to b).

Essentially, the prediction of readability requires calculation by means of an empirical formula relating specific variables of readability to the criterion for readability. Vogel and Washburne developed their equation predicting the average grade level equivalent of the paragraph meaning score of those children who read and *liked* specified books from four predictors: percentage of different uncommon words, number of prepositional phrases, and the relative number of simple sentences.²

Gray and Leary, after relating more than forty different predictors to their criterion, empirically chose five variables to predict readability: the number of different words, the percentage of uncommon words, the relative number of personal pronouns, the relative number of prepositional phrases, and the average sentence length.³

Gray and Leary's predicted readability score was a number which was transmuted into, a letter representing areas of difficulty of readability from A (very easy) to E (very difficult): Lorge, basing his work on that of Gray and Leary, tried to obtain a prediction in terms of grade level of reading. The sample of materials chosen for analysis was the 376 passages in the four books of McCall and Crabbs' *Standard Test Lessons in*

² The multiple correlations between the criterion and the weighted composite of the predictors was .845. Subsequently, Washburne and Vogel reported a multiple correlation of .869 on the basis of certain modifications.

³ The multiple correlations between the five predictors and the criterion used by Gray and Leary was .644.

Reading. The criterion was the grade level score, equivalent for a group of readers who would get half of the test questions right on each passage. The predictors studied by Lorge were the five used: by Gray and Leary: a weighted score for-vocabulary based on Thorndike's 20,000 word list, and four elements used by Morriss and Holversen (percentage of elemental words, percentage of simple localisms, percentage of concrete word-labels, and percentage of abstract word-labels). Later, Flesch's two factors, (affixed morphemes and human interest) were also used.⁴

The simple three-factor prediction equation of Lorge's was modified by the addition of a constant to give an estimate of the grade level score equivalent to passing three-quarters of the questions on a given passage. The formula is given on the, work sheet. (see page 210) for computing the readability index. A reasonably good prediction of readability can be obtained by using a weighted composite of vocabulary and sentence structure. Of these, the most important is some measure of vocabulary load. It should be recognized, however, that such elements as the number of abstract words, the number of uncommon words, the number of polysyllabic words, and the weighted index of difficulty of vocabulary are all inter-correlated. Any one of them could be used in place of any other, provided suitable adjustment were made in the empirical formula. Certainly some aspect of vocabulary load must be used as a predictor.

Structural elements of the passage provide the second most important basis for estimating the readability of text. As in measures of vocabulary, most measures of sentence structure are interrelated, so that little additional information is yielded by several measures of sentence structure.

Lorge's formula, as described in the following pages, uses as predictors the factor of uncommon words (vocabulary) and the factors of average sentence length, and the relative number of prepositional phrases (sentence structure.)

⁴ The multiple correlations were obtained predicting the criterion from various combinations of these factors. Empirically, the best prediction using the fewest factors was obtained with three factors (also used by Gray and Leary): the average sentence length, the relative number of prepositional phrases, and the relative number of different words not common to Dale's list of 769 words. The multiple correlation coefficient between the average grade score on the Thorndike-McCall Reading Test and the three predictors was .77. Adding as predictors the weighted index for word frequency and/or the four factors of Morriss and Holversen, separately or in combination, and/or the two factors of Flesch, separately or in combination, did not increase the multiple correlation significantly.

FORMULA FOR JUDGING READABILITY

The Lorge formula, therefore, is a means of judging the relative difficulty or readability of either read or spoken passages. Readability is based upon the comprehension of passages by school children. Comprehension is judged by the correctness and completeness of responses to questions about a passage. Such questions usually deal with specific details, general import, appreciation, knowledge of vocabulary, and understanding of concepts.

It is obvious that the purpose of the reader in reading and the kinds of questions asked in estimating reading comprehension will influence greatly the estimate of reading difficulty. Since the Lorge formula is based on a criterion derived from responses to questions of the five types listed above, it tends to overestimate the difficulty of passages to be read primarily for appreciation or for general import and to underestimate the difficulty of passages to be read primarily for specific details or for following directions. Nevertheless, the formula provides an overall estimate which should be useful in grading reading materials. As an estimate, it should not be considered definitive nor used blindly.

As developed in the work sheet, the readability index is an estimate of the reading grade at which the average school child will be able to answer with adequate completeness and correctness about three-fourths of the questions concerning detail, appreciation, import, vocabulary, and concept. The reading grade so obtained may be thought of in terms of reading grade scores on a test of reading comprehension. A readability index of 5.2 for a passage may be considered indicative of the material of the fifth grade; it may be thought in terms of placement of the material as within the reading comprehension of average fifth grade children. Such placement, however, should consider the interest of pupils, the suitability of subject matter, and other factors. The readability index is an estimate and not a rigorous determination.

The Lorge formula, in addition to its use in estimating the reading difficulty of passages for children, may be used to advantage in estimating the difficulty of silent and oral passages for adults. It yields a readability index which places materials in relative order; that is, a reading passage with an index of 7.1, etc. Moreover, the suitability of texts for adults can be interpreted in terms of the reading grade scores of adults on acceptable reading tests.

Teachers of adults, or indeed, any person choosing tests for specific audiences, might give a reading test to a sample of adults to determine the average reading grade score (as well as the range of such scores). They then could choose texts within the demonstrated range of comprehension of such adults.

THE READABILITY INDEX

The Lorge Readability Index, in addition to its utility in grading text materials, may also be used for passage simplification. If the text for children is, let us say, designed for grade level 6.0 and on the basis of the formula has a reading index of 7.6, then the text may be revised by simplifying sentence structure, by substituting simple sentences for prepositional phrases, and by an adequate choice of vocabulary. Since vocabulary is the most important factor in passage difficulty, care must be taken to indicate the meaning of more difficult words by definition, example, or context. Choice of vocabulary, furthermore, may be controlled by use of *The Teacher's Word Book of 30,000 Words*, a new word book compiled by Thorndike and Lorge, in which every word is given a value according to relative frequency in the English language. The value of AA indicates words that occurred a hundred or more times per million words; the value of A indicates words that occurred from fifty to ninety-nine times per million words; the values 49, 48, 47, etc., indicate the number of times the word occurred-per million words. In selecting vocabularies for the revision of texts, a safe rule is to utilize, in addition to the information given by the index, these values recommended by Thorndike and Lorge.

In actual practice, the formula has proved, to be very serviceable in the simplification of texts for adult use. The grade placement of the text may be compared with the average highest grade reached by adults for whom it is designed. The median highest grade reached, for adults, twenty years and over is reported by the Bureau of the Census for the year 1940. For the adult population "20 years old and over" the median highest grade (number of years of school completed) was 8.8. In writing for such an average population, it may safely be assumed that the reading ability as measured by grade score on a reading test will be somewhat lower, let us say, about eight-tenths of a school year. Hence, in writing for a population with an assumed grade level score or a reading rest score of 8.0, steps should be taken to select vocabulary, simplify sentence structure, and reduce the number of prepositional phrases. Again, *The Teacher's Word Book of 30,000 Words* should be of considerable help, since it gives sepa-

rate evaluation for vocabularies found in adult magazines, e.g., *Saturday Evening Post*, *Ladies' Home Journal*, *Woman's Home Companion*, *True Story*, and *Reader's Digest*.

COMPUTING THE READABILITY INDEX

The following are directions for computing the readability index.

A. Selecting the sample:

1. Short passages of 100 words or less.

When a short passage is to be appraised, it is advisable to analyze the entire passage.

2. Longer passages.

When longer passages are to be appraised, it is advisable to analyze samples of the material. Select a sample near the beginning, another sample, near the middle, and another sample near the end of the passage. Each of these samples should be approximately one hundred words in length.

A good procedure might be to number the lines of text serially and then count the number of words per line (about ten lines) to get an estimate of the number of words. For instance, a passage has 141 lines; ten lines chosen at random have 11, 12, 13, 13, 12, 12, 12, 12, 16, and 16 words, or an average of 13 words to the line. The passage thus has approximately 1,833 words. A sample of 100 words would then be approximately eight lines in length. The three samples could be chosen in a variety of ways: e.g., beginning at or near line 3 through line 11; at or near line 53 through line 61; and at or near line 103 through line 111. In this way, a sample is chosen in each third of the passage.

It should be noted, moreover, that each sample should start with the beginning of a sentence and should stop at the end of a sentence. When, the samples have been located with beginning and end points, the remainder of the analysis can be made.

3. Books

When books are to be appraised, it would be advisable to analyze samples of the book, say, from 5 per cent to 10 per cent of the book (but never less than five samples). These samples should be

chosen throughout the book. For instance, a book has 92 pages of text with an average of 195 words per page. This indicates an approximate wordage of 18,000 words. A 5 per cent sample would be 900 words; a 10 per cent sample would be 1,800 words. For the 5 per cent sample this would require approximately five pages; for the 10 per cent sample, approximately nine pages. Thus every eighteenth page should be chosen for the 5 per cent sample; every tenth page, for the 10 per cent sample. Thus the sample might be pages 3, 21, 39, 57, 75 in the one instance; or 4, 14, 24, 34, 44, 54, 64, 74, 84 in the other. Of course, a sample must start with the beginning of a sentence and stop at the end of a sentence.

B. Labeling the work sheet.

1. Fill out the information about the title, author, edition, publisher, and date, of publication (latest copyright year listed).
2. Carefully identify the location of the sample, thus: "p. 14, line 2, The answer...p. 14 line 26, ever after."

C. Counting the number of words.

1. Begin with the beginning of the sample and count (or number serially) each word in the sample. Observe the following rules:
 - (a) Hyphenated words are counted as one word. When in doubt about uncommon hyphenations, follow Webster's Unabridged Dictionary (2nd edition); if listed in dictionary as hyphenated, count as one word; if not listed, count as two words.
 - (b) Words separated at the end of a line to the beginning of the next line are counted as one word.
 - (c) Numbers are counted as words, e.g., in "January 3, 1940" 3 is counted as one word and interpreted as the word *three*, 1940 is counted as one word and interpreted as *nineteen-forty*.
 - (d) Compound words like place names or persons' names are counted as one word, e.g., New York, United States, van Loon, Santa Claus, St. Nicholas.
 - (e) Contractions are counted as one word; e.g., don't, he's, they'll, they'd, etc., are each counted as one word.
2. Record the count under Basic Data, number 1.

D. Counting the number of sentences.

1. Begin at the beginning of the sample and count the number of complete sentences.
2. Record the count under Basic Data, number 2.

E. Counting Prepositional Phrases.

1. Count each prepositional phrase in the sample. Observe the following rules:

(f) A phrase is made up of a preposition and a noun, or a preposition and a pronoun, or a preposition and a gerund, e.g., to the house (noun), for him (pronoun), in skating (gerund).

(g) Some common prepositions are:

about	from
above	in
across	inside
after	into
along	of
among	off
at	on
before	onto
behind	outside
below	till
beneath	to
beside	under
beyond	up
by	upon
during	with
except	within
for	without

(h) Less common prepositions are:

despite (the opinion), concerning (the idea), notwithstanding (the opposition).

(i) Infinitive phrases are *not* to be counted. An infinitive phrase is made up of the word *to* and a verb, e.g., to swim, to sing, to an-

swer.

- (j) If a preposition word is followed by a clause, it is a conjunction, and hence is *not* counted, e.g., "After the storm had passed" is *not* counted.
2. Record the count under Basic Data, number 3.
- F. Counting hard words.
1. Use the Dale list⁵ to cross out in the sample every word on the Dale list, regardless of its meaning.⁶ The list is given on pages 183 to 186.
 2. Since the count is the number of different hard words, each hard word is counted only once. For instance, if in the passage *reliability* occurred three times, it still would be counted only once.

Observe the following rules:

(k) Nouns.

Separate counts are not made of plurals and possessives in *s*, plurals in *es*, or plurals in which *y* is replaced by *ies*: e.g., *boys*, *churches*, *berries* are counted with *boy*, *church*, *berry*; however, *knife* and *knives*, *goose* and *geese*, *man* and *men* are all counted as different words.

(l) Special cases.

An *s* added to a word in the text not forming a plural or possessive forms a different word from the root form: e.g., *Robert* and *Roberts* are two different words.

Proper nouns which seem to be composed of root and derived forms are not tabulated with the root form: e.g., *Wheeling*, the proper name is not counted with *wheel*. *Browning*, the proper name, is not counted with *brown*. Nouns formed by adding *r* or *er* to the other nouns or to verbs are not counted with the original word: e.g., *own* and *owner* are two different words.

(m) Adverbs.

Separate counts are not made of adverbs formed by adding *ly*: e.g., *badly*, *sadly* are counted with *bad*, *sad*.

⁵ The list is reproduced by permission of the author, Dr. Edgar Dale.

⁶ That is, *spring*, meaning season, jump, water, or steel coil, is counted as one word.

Adverbs formed from an adjective ending in *e*, as *gently* from *gentle*, *truly* from *true*, are counted as different words.

(n) Adjectives

Separate counts are not made of adjectives formed by adding *n* to proper nouns: e.g., *Austrian*, *Bavarian* are counted with *Austria*, *Bavaria*.

(o) Special cases.

An adjective formed by adding *ly* to a noun is counted as a different word from the noun: e.g., *home* and *homely* are two different words.

(p) Comparatives and superlatives of adjectives and adverbs.

Special counts are not made of comparatives and superlatives formed by adding *er* or *r* and *est* or by changing *y* to *ier* and *iest*: e.g., *longer*, *prettier*, *bravest* are counted with *long*, *pretty*, *brave*.

(q) Special cases.

The rule applies to adjectives doubling the final consonant and adding *er* and *est*: e.g., *red*, *redder*, *reddest* are counted as one word.

(r) Verbs.

Special counts are not made of verb forms ending in *ing* and in *s*, *d*, *ed*, or of forms changing *y* to *ies* and *ied* or of past participles formed by adding *n*: e.g., *plays*, *playing*, *played* are counted with *play*.

(s) Special cases.

Verb forms which drop the final *e* and add *ing* are counted with the root form: e.g., *pace* and *pacing* are counted as one word.

Verb forms which double the final consonant and add *ing* or *ed* are counted as one word: e.g., *drip*, *dripped*, and *dripping* are counted as one word.

Past participles formed by adding *en* to a verb are counted as different from the verb: e.g., *eat* and *eaten* are two different words.

(t) Hyphenated words.

In case of uncommon hyphenated words, follow Webster's Unabridged Dictionary (2nd edition). Any hyphenated word is considered as one word if it is listed thus in the dictionary; otherwise it is counted as two words.

(u) Compound words.

Compound names of persons or places, like *New York*, *United States*, *St. Louis*, *Santa Claus*, and *Van Dyke*, count as single words.

(v) Contractions

Count contractions as different words from those from which they are derived: e.g., *because* and *'cause* are two different words. *He's* is not counted with *he* or with *is*.

(w) Words which may be both common and proper.

In the case of words which may be both common and proper nouns, count the proper noun as being the same word as the common: e.g., *Jack* and *jack* are the same word.

(x) Miscellaneous special cases.

Words formed by adding *y* to a word in the list are counted as different from the root word: e.g., *snow* and *snowy* are different words. *German* and *Germany* are different words.

Words of different spelling listed in the dictionary as one word are counted as the same word: e.g., *honor* and *honour* are the same word. *Frankfort* and *Frankfurt* are the same word.

If a word is formed by adding two or more suffixes to a listed word, one of which when added to the listed word is counted with it, that word is different from the root word: e.g., *happen* and *happening* are the same word but *happenings* is a different word. *Excite* and *excited* are the same word, but *excitedly* is a different word.

Words formed by adding *en* are counted as different from the original word: e.g., *wool* and *woolen* are two different words, *bit* and *bitten* are two different words.

3. Record the count under Basic Data, number 4.

G. Proceed to computation. Watch decimal points carefully. Check all computations.

- H. Record on the work sheet the index (R. I.) to one decimal place.
- I. Make sure that the analyst, computer, and checker have signed the record blank and dated their entries.
- J. If a book or a long passage has had several samples selected from it, the average of the R. I.'s is the rating for the passage of the book.

The Dale list of easy words is made up of words which are common to Thorndike's first thousand most frequent English words and the first thousand most frequent words known by children entering the first grade. It is a list of words that are likely to be known by all children and adults. The Dale list, therefore, can be used to estimate ease of vocabulary; or, if the easy words are eliminated, an estimate of vocabulary difficulty can be made.

The passage chosen to illustrate the mechanics of estimating the readability index is the first revision of the Gettysburg Address.

LINCOLN'S GETTYSBURG ADDRESS

Four score and seven years ago our fathers brought forth on this continent a new nation, conceived in Liberty and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battlefield of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives, that that nation might live. It is altogether fitting and proper that we should do this. But, in a larger sense, we cannot dedicate—we cannot consecrate—we cannot hallow—this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us, the living, rather to be dedicated here to the unfinished work which they who fought here, have thus far, so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us—that from these honored dead we take increased devotion to that cause for which they here gave the last full measure of devotion—that we here highly resolve that these dead shall not have died in vain—that this nation, under God, shall have a new birth of freedom—and that, government of the people, by the people, for the people, shall not perish from the earth.

The usual procedure is to cross out all words on the Dale list, to encircle all prepositions, then to list the words not on the Dale list. An example of the listing of the hard words is given on the next page.

ILLUSTRATION OF LISTING OF HARD WORDS

A	altogether	add	advanced				3
B	battlefield	birth					2
C	continent	con- ceived /	created	civil	cannot //	consecrate /	6
D	dedicated ////	detract	devotion /				3
E	equal	engaged	endure				3
F	forth	final	fought	freedom			4
G	government						1
H	hallow	honored					2
I	increased						1
J							
K							
L	Liberty						1
M							
N	nation ////	nobly					2
O							
P	proposition	portion	proper	power	perish		5
Q							
R	remaining	resolve					2
S	score	sense	struggled				3
T	testing	thus	task				3
U	unfinished						1
V	vain						1
WXYZ							

43

After the number of sentences has been counted, the work sheet may be completed as shown on page 210.

The Lorge, Readability Index was developed after an analysis of the relationship between the score of readability for each of 376 passages and three internal measures of vocabulary and sentence structure. The resulting formula predicts readability well.

Teachers will find the directions for applying the formula simple and direct. The time required to analyze a passage is relatively short.

Teachers will find that the expenditure of time and effort in grading materials is easily justified in terms of the increased understanding of, and the possible reduction of, difficulties in communication.

The meaning of the index is simply the school grade at which the passage can be understood. The index, therefore, can be used to place texts and other books in appropriate grades: further, it should indicate ways in which passages may be rewritten to be appropriately placed for designated readers.

Unlocking Language

FORMULA FOR ESTIMATING GRADE PLACEMENT OF READING MATERIAL

WORK SHEET

Title of article: Gettysburg Address

Edition: first revision

Name of author: Abraham Lincoln

Publisher:

Date of Publication: Nov. 19, 1863

Location of sample in text: Complete

R. I. = 6.5

BASIC DATA

1.	The number of words in the sample.....	269
2.	The number of sentences in the sample	10
3.	The number of prepositional phrases in the sample.....	26
4.	The number of hard words in the sample.....	43

COMPUTATION⁷

Item 6, average sentence length: Divide 1 by 2 = 26.90 x .07 = 1.8830

Item 8, ratio of propositional phrases: Divide 3 by 1 = .0967 x 13.01 = 1.2581

Item 9, ratio of hard words: Divide 4 by 1 = .1599 x 10.73 = 1.7151

Constant = 1.6126

Add 6, 8, 9, and C

Readability Index: 6.4694

NOTES

lives, n. called easy

Last sentence, although long, is broken up by adequate punctuation

Name of Analyst: I. D. L

Date of analysis: Nov 23, 1943

Name of Computer: I. D. L.

Date of computing: Nov. 23, 1943

Name of checker: J. C.

⁷ 1948 Lorge corrections: In Item 6, change "x .07" to "x .06"

In Item 8, change "x 13.01" to "x .10"

In Item 9, change "x 10.37" to "x .10".

and for the Constant, change "= 1.6126" to "= 1.99" —WHD

THE DALE LIST OF 769 EASY WORDS

A	bag	both	cent	crowd	easy
a	ball	bottom	center	crown	eat
about	band	bow	chair	cry	edge
above	bank	box	chance	cup	egg
across	basket	boy	change	cut	eight
act	be	branch	chief		either
afraid	bear	brave	child	D	else
after	beat	bread	children	dance	end
afternoon	beautiful	break	choose	dark	England
again	because	breakfast	Christmas	day	English
against	bed	bridge	church	dead	enough
ago	bee	bright	circle	dear	even
air	been	bring	city	deep	evening
all	before	broken	class	did	ever
almost	began	brother	clean	die	every
alone	begin	brought	clear	different	everything
along	behind	brown	clock	dinner	except
already	being	build	close	do	expect
also	believe	building	cloth	doctor	eye
always	bell	built	clothes	does	
am	belong	burn	cloud	dog	F
American	beside	busy	coal	done	face
an	best	but	coat	don't	fair
and	better	butter	cold	door	fall
animal	between	buy	color	double	family
another	big	by	come	down	fancy
answer	bill		coming	draw	far
any	bird	C	company	dream	farm
anything	bit	cake	cook	dress	farmer
apple	black	call	cool	drink	fast
are	bless	came	corn	drive	fat
arm	blind	can	corner	drop	father
around	blood	cap	cost	dry	feed
as	blow	captain	could	dust	feel
ask	blue	car	count		feet
at	board	care	country	E	fell
away	boat	careful	course	each	fellow
	body	carry	cover	ear	felt
B	bone	case	cow	early	fence
baby	book	catch	cried	earth	few
back	born	cause	cross	east	field

Unlocking Language

fight	God	him	L	many	near
fill	going	himself	lady	march	neck
find	gold	his	laid	mark	need
fine	golden	hold	lake	market	neighbor
finger	gone	hole	land	matter	neither
finish	good	home	large	may	nest
fire	got	hope	last	me	never
first	grain	horse	late	mean	new
fish	grass	hot	laugh	measure	New York
fit	gray	house	lay	meat	next
five	great	how	lead	meet	nice
fix	green	hundred	learn	men	night
floor	grew	hunt	leave	met	nine
flower	ground	hurry	left	middle	no
fly	grow	hurt	leg	might	noise
follow	guess	I	lesson	mile	none
food		I	let	milk	noon
foot	H	ice	letter	mill	nor
for	had	if	lie	mind	north
forget	hair	in	lift	mine	nose
fourth	half	Indian	light	minute	not
found	hall	instead	like	miss	note
four	hand	into	line	money	nothing
fresh	hang	iron	lion	month	now
friend	happy	is	lips	moon	number
from	hard	it	listen	more	
front	has	its	little	morning	O
fruit	hat		live	most	oak
full	have	J	lead	mother	ocean
	he	jump	long	mountain	of
G	head	just	look	mouth	off
game	hear		lost	move	office
garden	heard	K	lot	Mr.	often
gate	heart	keep	loud	Mrs.	old
gave	heavy	kept	love	much	on
get	help	kill	low	music	once
gift	her	kind		must	one
girl	here	king	M	my	only
give	herself	kiss	made	myself	open
glad	hide	knee	mail		or
glass	high	knew	make	N	other
go	hill	know	man	name	our

out	R	second	small	summer	till
outside	race	see	smile	sun	time
over	rain	seed	smoke	suppose	tire (d)
own	ran	seem	snow	sure	to
	rather	seen	so	surprise	today
P	reach	self	soft	sweet	together
page	read	sell	sold		told
paint	ready	send	soldier	T	tomorrow
pair	real	sent	some	table	tongue
paper	reason	serve	something	tail	too
part	red	set	sometime	take	took
party	remember	seven	song	talk	top
pass	rest	several	soon	tall	touch
path	rich	shake	sound	taste	town
pay	ride	shall	south	teach	trade
pen	right	shape	space	teacher	train
people	ring	she	speak	tear	tree
pick	river	sheep	spot	tell	true
picture	road	shine	spread	ten	try
piece	rock	ship	spring	than	turn
place	roll	shoe	square	thank	twelve
plain	roof	shop	stand	that	twenty
plant	room	short	star	the	two
play	rose	should	start	their	
please	round	shoulder	station	them	U
point	row	show	stay	then	uncle
poor	run	shut	step	there	under
post		sick	stick	these	until
pound	S	side	still	they	up
present	said	sign	stone	thick	upon
press	sail	silk	stood	thin	us
pretty	salt	silver	stop	thing	use
pull	same	sing	store	think	
put	sand	sir	storm	this	V
	sat	sister	story	those	valley
Q	save	sit	straight	though	very
quarter	saw	six	street	thought	visit
queen	say	size	strike	thousand	
quick	school	skin	strong	three	W
quiet	sea	sky	such	through	wait
quite	season	sleep	sugar	throw	walk

	seat	slow	suit	tie	wall
want	wear	where	wide	without	Y
war	weather	whether	wild	woman	yard
warn	week	which	will	wonder	year
was	well	while	win	wood	yellow
wash	went	white	wind	word	yes
waste	were	who	window	work	yesterday
watch	west	whole	wing	world	yet
water	what	whom	winter	would	you
wave	wheat	whose	wish	write	young
way	wheel	why	with	wrong	your
we	when				

BIBLIOGRAPHY

- BROWN, ROBERT. "Vocabularies of History and Reading Textbooks." *Bulletin of the Department of Elementary School Principals*, Vol. 10, pp. 408-411, April, 1931.
- BUREAU OF THE CENSUS, U.W. Department of Commerce. Educational Characteristics of the Population of the United States, by Age: 1940. Series P-19, No. 4, December 17, 1943
- DALE, EDGAR. "A Comparison of Two Word Lists." *Educational Research Bulletin*, Vol. 10, pp. 484-489, December, 1941.
- DALE, EDGAR AND TYLER, RALPH W. "A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability." *Library Quarterly*, Vol. 4, pp. 384-412, July, 1934.
- DE LONG, VAUGHN R. "Primary Promotion by Reading Levels." *Elementary School Journal*, Vol. 38, pp. 663-671, May 1938.
- DOLCH, E. W. "Sampling of Reading Matter," *Journal of Educational Research*, Vol. 22, pp. 213-215, October, 1930.
- FLESCH, RUDOLF. *Marks of a Readable Style*. Contributions to Education, No. 897. Bureau of Publications, Teachers College, Columbia University, New York, 1941.
- FLESCH, RUDOLF. "Estimating the Comprehension Difficulty of Magazine Articles." *Journal of General Psychology*, Vol. 28, pp. 61-80, January 1941.
- GRAY, WILLIAM S. AND LEARY, BERNICE E. *What Makes a Book Readable*. University of Chicago Press, Chicago, 1935.
- HOLLAND, B. F. "The Effect of the Length and Structure of the Sentence on the Silent Reading Process." *Psychological Bulletin*, Vol. 30., pp. 668-669. November 1933.
- JOHNSON, GEORGE R. "An Objective Method of Determining Reading Difficulty." *Journal of Educational Research*, Vol. 21, pp. 283-287, April, 1930.
- KEBOCH, F. D. "Variability of Word Difficulty in Five American History Textbooks." *Journal of Educational Research*, Vol. 15, pp. 22-26, January, 1917.

- LEWERENZ, ALFRED S. "Measurement of the Difficulty of the Difficulty of Reading Materials." *Los Angeles Educational Research Bulletin*. Vol. 8, pp. 11-16, March 1919.
- LEWERENZ, ALFRED S. A Vocabulary Grade Placement Formula. *Journal of Experimental Education*, Vol. 3, p. 236, March 1935.
- LIVELY, BERTHA, A. AND PRESSEY, S. L. "A Method for Measuring the Vocabulary Burden of Textbooks." *Educational Administration and Supervision*, Vol. 9, pp. 389-398, October, 1923.
- LORGE, IRVING. "Predicting Reading Difficulty of Selections for Children." *Elementary English Review*, Vol. 16, pp. 229-233., October, 1939
- MCCALL, WILLIAM A. AND CRABBS, LELAH M. *Standard Test Lessons in Reading*. Books II, III, IV, and V. Bureau of Publications, Teachers College, Columbia University, New York, 1938.
- MCCLUSKY, HOWARD Y. "A Quantitative Analysis of the Difficulty of Reading Materials. *Journal of Educational Research*, Vol. 28, pp. 276-282, December, 1934.
- MORRISS, ELIZABETH C. AND HOLVERSEN, DOROTHY. "Idea Analysis Technique." Unpublished Ms., Teachers College, Columbia University, New York, 1938.
- OJEMANN, RALPH. "The Reading Ability of Parents and Factors associated with Reading Difficulty of Parent-Education Materials." *Researches in Parent Education*, II. *University of Iowa Studies in Child Welfare*, Vol. 8, pp. 11-32, March 1, 1934.
- PATTY, W. W. AND PAINTER, W. I. "Improving Our Method of Selecting High School Textbooks." *Journal of Educational Research*, Vol. 24, pp. 23-32, June, 1931.
- THORNDIKE, EDWARD L. *A Teacher's Word Book of Twenty Thousand Words*. Bureau of Publications, Teachers College, Columbia University, New York, 1932 (revised).
- THORNDIKE, EDWARD L. AND LORGE, IRVING. *The Teacher's Word Book of 30,000 Words*. Bureau of Publications, Teachers College, Columbia University, New York, 1943.
- TILLEY, HARVEY C. "Technique for Determining the Relative Difficulty of Word Meanings among Elementary School Children." *Journal of Experimental Education*, Vol. 5, pp. 61-64, September, 1936.
- VOGEL, MABEL AND WASHBURN, CARLETON. "An Objective method of Determining Grade Placement of Children's Reading Material." *Elementary School Journal*, Vol. 18, pp. 273-281, January, 1928.
- WASHBURN, CARLETON AND MOPHETT, MABEL V. "Grade Placement of Children's Books." *Elementary School Journal*, Vol. 38, pp. 355-364, January, 1938.
- YOAKUM, GERARD ALAN. "A Technique for Determining the Difficulty of Reading Materials." Unpublished Study, University of Pittsburgh, 1939

1948—The Dale-Chall Readability Formula

Introduction

THE DALE-CHALL readability formula arrived at the end of a national research effort that began in the 1920s and that eventually resulted in over a thousand published studies on the readability formulas.

In the beginning, the purpose was to develop reading materials for first-generation immigrants coming into high school.

This research intensified in the Second World War as the U.S. war effort focused on the need for clear and expressive writing. After the war, scholars harvested those hard-won lessons, and they gave us a new set of readability formulas for creating written materials for adults with limited reading ability.

The formulas created at that time, including the Dale-Chall formula, the Flesch Reading Ease formula, and the Gunning Fog Index, have remained the workhorses of many sectors of commerce, education, the military, and government.

Of all the readability formulas, the Dale-Chall formula has consistently been the most reliable. It has a correlation coefficient of .92 with comprehension as measured by reading tests.

Most of the readability formulas use a word variable and a sentence-length variable. Unlike most other modern formulas, the Dale-Chall formula uses a list of 3,000 easy words. Using the formula requires counting the number of “hard” words—those not on the list. Doing this manually becomes easy with practice. There are also a few computer programs available that apply the formula for you.

For 25 years a professor of education at Ohio State University, **Edgar Dale** was a respected authority on communications. He worked his whole life to improve the readability of books, pamphlets, and newsletters—the stuff of everyday reading.

Dale was one of the first critics of the Thorndike vocabulary lists. He claimed it failed to measure the familiarity of words accurately. He subsequently developed new lists that were later used in readability formulas.

Of major importance was *The Living Word Vocabulary: A National Vocabulary Inventory*, which he wrote with Joseph O'Rourke. This work, published by the publishers of *World Book Encyclopedia* in 1981, lists the grade levels of 40,000 words.

In 1948, Dale published the formula he developed with **Jeanne Chall**¹. She later was the founder and director for 20 years of the Harvard Reading Laboratory. She also led the battle for teaching early reading systematically with phonics. Her 1967 book *Learning to Read: The Great Debate*, brought research to the forefront of the debate. For many years, she also was the reading consultant for TV's *Sesame Street* and *The Electric Company*.

Dale and Chall introduced their readability formula in two issues of the *Educational Research Bulletin*. They included this simple disclaimer, "We do not claim the formula developed here is definitive. The nature of the multiple-correlation coefficient makes this point rather obvious. We do believe, however, that it is a short cut in judging the difficulty of written materials."

Millions of readers, young and old, throughout the world have benefited immensely from the work of Edgar Dale and Jeanne S. Chall. No small measure of these benefits has resulted from the use of their easy-to-use and reliable readability formula.

—WHD

¹ Dale, E. and J. S. Chall. 1948 "A Formula for Predicting Readability." *Educational Research Bulletin*, Vol. 27, No.1, pp. 11-20, 28.

Dale, E. and J. S. Chall. 1948. "A Formula for Predicting Readability: Instructions" *Educational Research Bulletin*, Vol. 27, No. 2, pp. 37-54.

Educational Research Bulletin

Vol. XVII

January 21, 1948

No. 1

A Formula for Predicting Readability

By EDGAR DALE and JEANNE S.CHALL

SEVERAL MONTHS AGO the editor of the Wall Street Journal ran a full-page advertisement, in one of the leading literary magazines, announcing two honors recently awarded to it. One of these honors was a statement made by Robert P. Gunning that the Wall Street Journal had "the most readable front page in the country."

How did Mr. Gunning come to this conclusion? Did he actually sample a cross section of the readers, have them read the front pages of leading, newspapers, and then compare their ability to read and understand the various front pages? No. He used an accepted short cut. He predicted the reading difficulty of the various front pages by using a readability formula and found that the Wall Street Journal was the "most readable."

This recognition by leading journalists that readability is an important selling point for their newspapers is an event that is still quite new on the publishing horizon. Although some objective techniques for measuring readability have been known for at least twenty-five years, they have been neatly buried in educational and psychological journals, doctoral dissertations, Masters' theses, and the like. If the techniques were used at all, they were confined to children's textbooks. What has taken the dust off the technical journals and made readability a household word in the writing and publishing field?

As in the rise in popularity of any technique, there was a critical need for this one. The war period made us realize more than ever the importance of reaching large audiences. More, people had to fill out tax forms; more people had to be appealed to to buy war bonds; more people had to co-operate in numerous activities to help win the war. Because a larger audience had to be reached, the writers had to use a style that could be understood by more persons than the usual book readers. They could no longer afford to hit or miss with printed materials.

ALONG with the growing need for more scientific means of verbal communication, there was a growing fund of practical objective measurement of readability. The Lorge formula was one of the first easy-to-apply readability formulas.¹ By the use of this formula we could predict in a fairly short time how difficult a piece of written material was to read and understand. It was no longer necessary to guess. By counting the relative number of different uncommon words, the average sentence length, and the relative number of prepositional phrases, we could get a good index of readability in terms of grade scores.

In 1943 Rudolf Flesch produced his readability formula.² He presented a very convincing argument for the superiority of his formula over the previous ones, especially for use with materials for adult readers. With numerous correlation tables he showed that the Lorge formula, in its use of the Dale List of 769 Easy Words as a measure of vocabulary difficulty, failed to discriminate satisfactorily between materials that were above the eighth-grade level in difficulty. Since the average adult has approximately eighth- or ninth-grade reading ability, he thought that another technique was needed to predict the readability of materials for adult readers. In his formula, Mr. Flesch used three factors: average sentence length, relative number of affixed morphemes (prefixes, suffixes, inflectional endings) and relative number of personal references.

WHEN the Flesch formula was first released, we were evaluating the educational materials published by the National Tuberculosis Association. It was our job to analyze the pamphlets already published and to find ways of writing them so that they could be understood by the average adult. We used the Flesch formula to help us evaluate the reading difficulty of the pamphlets.

On the whole, we found the formula adequate. However, we also found some shortcomings. The most serious shortcoming was the count of affixes, which we found to be rather arbitrary, in the sense that two people making a count on the same sample would usually come out with a different number of affixes. If we were extremely careful and consulted a

¹ Lorge, Irving. "Predicting Reading Difficulty of Selections for Children," *Elementary English Review*, XVI (October, 1939), pp. 229-33, and "Predicting Readability," *Teachers College Record*, XL (March, 1944), pp. 404-19.

² Flesch, Rudolf. *Marks of a Readable Style*. New York: Teachers College, Columbia University, 1943. (Teachers College, Columbia University, Contributions to Education, No. 897).

dictionary to be certain that all affixes were included and that no non-affixes were included, we found that the work was too time-consuming.

Mr. Flesch's reasons for using affixes as a count of difficulty are very well stated, with statistical evidence, in both his books.³ His logic was that word recognition, although an important factor in reading for beginning or poor readers, is of practically no importance for more mature readers. For the better readers, it is the relationship between the words and abstractness of the words that contribute to difficulty. He actually computed the affixes (as a measure of verbal relationship) and the abstract words contained in five levels of magazines and found that both of these factors were a good measure of difficulty. He dropped the count of abstract words in his formula because the magazine experiment "had shown that the count of affixes was a practically equivalent measure of abstractness ($r = .7849$) and the latter method was far less cumbersome."⁴ In fact, in another section of his book, he refers to the count of affixes as "a simple short cut to the count of abstractions."⁵

If Mr. Flesch used a correlation of .7849 to justify his calling the affixes a "simple short cut to the count of abstractions," could we not also call the Dale List of 769 Easy Words a short cut to the count of abstractions, since Mr. Lorge found a high correlation between affixed morphemes and words outside this list? Or could we not argue that Mr. Flesch's count of affixes is just another way of counting hard words?

In his article, "Predicting Readability," Mr. Lorge makes the following statement about measuring vocabulary load:

It should be recognized that such elements as the number of abstract words, the number of uncommon words, the number of polysyllabic words, find the weighted index of difficulty of vocabulary are all inter-correlated. Any one of them could be used in place of any other, provided suitable adjustments were made in the empirical formula.⁶

If all counts of vocabulary load, whether abstract words, affixed morphemes, or number, of uncommon words, are inter-related, why use a less exact and more cumbersome method when a simpler one can be used?

³ Flesch, op. cit. and *The Art of Plain Talk* (New York: Harper and Brothers, 1946),

⁴ Flesch, Marks of Readable Style, p. 32.

⁵ Ibid., p. 24.

⁶ Lorge, *loc. cit.*, 406.

FROM the evidence given, we believed that there was value in using a word list to measure vocabulary load. Mr. Flesch's main objection to the use of the Dale list of 769 words was that it did not differentiate between the higher levels of difficulty. What would happen if a larger word list were used?

Such a list would not be a discriminating instrument at the easy levels of writing since it would contain words not known to some of the readers. But by using a list which included most of the words well known to fourth-grade readers, a more discriminating instrument would be devised for the upper levels of reading ability.

The second shortcoming of the Flesch formula was the count of personal references. In our numerous analyses we found that the personal-reference count was not a reliable index of difficulty. For example, when we speak of *John* and *Mary* and *he* and *she*, referring to John and Mary, there is a justification for subtracting from difficulty. This is because in writing about John and Mary we usually say things that are not abstract or general. However, subtracting from difficulty for personal references such as *R. J. Thomas* of the automobile industry, or *Senator Austin*, when we are writing about atomic energy or the United Nations, does seem to us a bit inaccurate. If the reader does not know these persons, the difficulty of the written material is not decreased. In fact, these individuals are no longer personal, they are abstractions. Flooding printed materials with personal references to these “abstract” persons will add little to “human interest” and ease of comprehension.

A recent article in the *American Psychologist* by S. S. Stevens and Gerldine Stone reported that Koffka's *Principles of Gestalt Psychology* had predicted a Flesch score much lower than had been expected. In fact, it came out only a little higher than the elementary textbooks in psychology. It was startling news for them. They wrote:

The Harvard graduate students don't believe it, because they read Koffka and sweat.

Now how can Koffka, the students' choice for unreadability, score so low? Opinion around Harvard seems divided on this question, but this opinion is based on mere casual introspection, not on the result of careful analysis. A few things appear evident, however. For one thing, Koffka helps his score by peppering his passages with personal pronouns: 5.8 per hundred words. But his “I,” “we,” and “you” are rhetorical devices—he is

actually very rarely talking about us or about himself. He is talking about abstractions and complicated relations and *he* and *we* get into it as mere guinea pigs in an experiment.⁷

Here is the sample they quote from Koffka:

In the first cases, real moving objects present in the field, the shift of the retinal pattern leads to the behavioral motion of objects, whether I fixate a non-moving object or follow a moving one with my regard; in the second case, when my eyes roam over stationary objects, such a shift will *not* have this result. Although the two facts belong closely together, the second one will be fully discussed in Chapter IX, after we have introduced the ego. Here we concentrate mainly on the first, even if we cannot entirely avoid referring to the second. Thus we turn now to the theory of perceived motion.⁸

This passage has 7 personal references per hundred words. According to Flesch's Quick Reference Chart,⁹ a similar number of personal references characterizes materials that in difficulty are standard and are comparable to digest magazines.

IN VIEW of the shortcomings of the Flesch counts of affixes and personal references, we undertook to find a more efficient means of predicting readability. Our hypotheses were:

First, a larger word list would predict as well as, if not better than, the count of affixes. It would avoid the pitfalls of lack of discrimination at the upper levels of difficulty.

Second, a count of personal references does not add very much to the prediction of readability.

Third, a shorter, more efficient formula could be evolved with the use of a word factor and a factor of sentence structure.

For our sample passages, we used the McCall-Crabbs *Standard Test Lessons in Reading*,¹⁰ the same passages used by Mr. Lorge and Mr.

⁷ *American Psychologist*, II (July, 1947), p. 233.

⁸ Koffka, K., *Principles of Gestalt Psychology*. New York: Harcourt, Brace and Company, 1935. p. 280.

⁹ Flesch, *The Art of Plain Talk*, p. 205.

¹⁰ McCall, W. A., and Crabbs, Lelah. New York: Bureau of Publications, Teachers College,

Flesch. These are a series of 376 passages of children's readings, already graded in difficulty on the basis of comprehensibility of questions at the end of each passage. This material, it should be noted, has serious deficiencies as a criterion, but it is the best we have at the present time. The writers, however, checked their findings against other passages as noted later. Following these authors, our criterion was the grade-level score equivalent for a group of readers who would get half of the test questions right on each passage. Mr. Lorge made his data-sheets available to us.¹¹ These data sheets also included the Flesch counts of affixed morphemes and personal references.

Our word count was based on the Dale list of approximately three thousand words. This list was constructed several years ago by testing fourth-graders on their knowledge in reading of a list of approximately ten thousand words. This larger list included the most common words in the Thorndike,¹² Buckingham and Dolch,¹³ and other word lists. Words such as *milkman*, *carrot*, *candlestick*, *catbird*, and so on, which appeared in the high thousands, on the Thorndike list, were also tested with fourth-graders to see whether they knew them. An attempt was made to include all words that fourth-graders would possibly know. A word was considered as known when at least 80 per cent of the fourth-graders checked it as known.

This list differs from the Thorndike-word lists in that it is a measure of familiarity in reading rather than a measure of frequency of appearance in printed materials. Words such as *bracelet*, *watermelon*, and *cabbage*, appearing in the high thousands in the Thorndike lists, are included in the Dale three thousand list. In that respect it is less artificial than the Thorndike lists. No claim is made that all the words actually known in reading by at least 80 per cent of fourth-graders are on this list. Some may have been left out. The testing method used is crude. But it does present a fairly complete list of familiar and simple words.

Columbia University, 1926.

¹¹ The authors wish to thank Mr. Lorge for making the data sheets available and for permission to publish the intercorrelations of his factors.

¹² Thorndike, Edward L. *A Teacher's Word Book of Twenty Thousand Words*. New York: Teachers College, Columbia University, 1931.

¹³ Buckingham, B. R., and Dolch, E. W. *A Combined Word List*. Boston: Ginn and Company, 1936

WE WENT through the 376 passages in Books II to V of the McCall-Crabbs test lessons. In each passage, we counted the relative number of words not on the Dale list of 3,600 words.¹⁴

We punched this information on Hollerith cards, along with the information made available by Mr. Lorge. The intercorrelations appear in Table I.¹⁵

From Table I the reader can see that the highest correlation with the criterion is the relative number of words outside the Dale list of 3,000 words. The correlation is .6833. The two next highest factors are the Lorge hard-word count (based on the Dale list of 769 words) and the Flesch affixed-morphemes count. The intercorrelations among these three factors are high; between the Dale score and the Flesch morphemes, .7932; between the Dale score and the Lorge hard-word count, .7988. This table corroborates Mr. Lorge's findings that a measure of vocabulary load is the most important factor in reading difficulty, and that all the measures of vocabulary are highly intercorrelated.

The next highest measure of difficulty is average sentence length—which correlates .4681 with the criterion.

After making several combinations of factors, we found that the following two, plus a constant, gave the most efficient empirical formula:

$$X_{C_{50}} = .1579X_1 + .0496X_2 + 3.6365$$

When:

$X_{C_{50}}$ = reading-grade score of a pupil who could answer one-half of the test questions correctly

X_1 = Dale score (relative number of words outside Dale list of 3000 words)

X_2 = average sentence length

3.6365 = constant

The multiple-correlation coefficient of these two factors with the criterion is .70. Adding the factor of human interest (personal reference) of

¹⁴ The specific instructions for counting will be included in the instructions which will appear in next month's issue.

¹⁵ We wish to thank Mr. Flesch for permission to use his factors and to publish the intercorrelations of his factors, and Harold A. Edgerton for invaluable statistical help.

Mr. Flesch raises the multiple-correlation coefficient to .7025, an insignificant increase.

Because of the correction in the sentence-length factor, we recomputed the multiple-correlation coefficients on the Lorge and Flesch formulas. The corrected Lorge formula also has a multiple correlation of .66. We see that the one factor, words outside the Dale list of 3,000 words, alone, has a greater prediction than the three-factor Flesch and Lorge formulas.

TABLE I
Intercorrelations between Four Style elements and grade Score of a Pupil Who Answered One-Half of the Questions on McCall and Crabbs

	Dale Score (3,000 List)	Flesch Affixed Morphemes	Flesch Personal References	Lorge Hard Words (Dale 769)	Criterion C ₅₀	Mean	Standard Deviation
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average sentence length	.5108	.4428	-.2201	.4913	.4681*	16.8037	5.3813
Dale score (words outside 3,000 list)7932	-.4033	.7988	.6833	8.1011	6.3056
Flesch affixed morphemes†	-.3254	.7441	.6017	25.2819	11.0668
Flesch personal references†	-.3422	-.3675	7.8245	5.5439
Lorge hard words (outside Dale 769 list)6148	17.4165	7.1659
Criterion (C ₅₀)	5.7492	1.6565

* The correlation coefficient reported here between the average sentence length and the criterion is much lower than the one reported by Mr. Lorge and later by Mr. Flesch. They reported a correlation coefficient of .6174. We checked this with Mr. Lorge. He went over his data and found that an error had been made in the computation. He is publishing the correlation in an article which will appear in *School and Society*, February 21, 1948.

† The intercorrelations of the two Flesch factors here reported are slightly different from those presented by Flesch in *Marks of Readable Style*. These differences are not significant and were probably caused by our using gross scores on Hollerith cards while Flesch used grouped data for his correlations.

DOES this new two-factor work in predicting the difficulty of reading materials other than the McCall-Crabbs reading passages? We conducted several experiments comparing the formula predictions with the judgments of experienced teachers, the judgment of readability "experts," and the actual comprehension scores of readers on

passages.

On fifty-five passages of health-education materials, we found that our two-factor formula predictions correlated .92 with the judgments of readability experts, and .90 with the reading grades of children and adults who were able to answer at least three questions out of four on thirty of these passages. They ranged from the extremely easy to the very difficult.

On 78 passages on foreign affairs from current-events magazines, government pamphlets, and newspapers, the correlation between the predictions of the formula and judgments of difficulty by expert teachers in the social studies was .90.

As a result of these various experiments, we set up the following table of estimated corrected grade levels:

Formula Score	Corrected Grade Levels
4.9 and below	Grade 4 and below
5.0 to 5.9	Grades 5-6
6.0 to 6.9	Grades 7-8
7.0 to 7.9	Grades 9-10
8.0 to 8.9	Grades 11-12
9.0 to 9.9	Grades 13-15 (college)
10 and above	Grades 16 and above (college graduate)

The formula developed by the writers is a simple, two-factor formula that is easy to apply. With the use of a factor of vocabulary load (relative number of words outside the Dale list of 3,000 words) and a factor of sentence structure (average sentence length), we have a good prediction of readability. The additional validation on health and social-studies materials shows that it compares favorably with judgments of experts and with actual reader comprehension.

The corrected grade levels help interpret the scores obtained by the formula and give a more usable means of placing materials within the comprehension of the various grades. For example, a given piece of material having a formula score of 5.2 (corrected grade level of Grades V-VI) should be within the comprehension of children who have fifth-to sixth-grade reading abilities. By this we mean that these children will be able to answer approximately one-half to three-fourths of the questions asked on the material, concerning specific details, general import, appreciation,

knowledge of vocabulary, and so on.

For adults, the corrected grade levels may be interpreted to mean the number of years of schooling required to read the material with ease and understanding. For example, if an article or book has a formula score of 6.3 (corrected grade level of Grades VII-VIII), it would be within the comprehension of the average adult who has had about eight and one-half years of schooling.

WE DO not claim that the formula developed here is definitive. The nature of the multiple-correlation coefficient makes this point rather obvious. We do believe, however, that it is a short cut in judging the difficulty of written materials.

The formula can also be used as an aid to text simplification. When a text has an undesirably high score according to the prediction of the formula, it may be simplified by substituting more concrete, familiar words for the unfamiliar and abstract words. Perhaps sentences can be shortened and made clearer. Writing should not be any harder to read and understand because the ideas are hard and complicated. It may be impossible to simplify this type of writing. On the other hand, a good deal of writing is hard because the words used are unnecessarily abstract and the sentence and paragraph structure needlessly complex. A later article will discuss these problems. But we must be cautious about "writing for a readability formula." We must remember at all times that a formula is a statistical device. It means that, on the whole, longer sentences make comprehension more difficult. This does not mean that all long sentences are hard to read and understand. There are some very short sentences that may be harder to comprehend than longer ones. The same holds true for the use of familiar words. On the whole, the more unfamiliar the words used, the harder the material will be to understand. But sometimes familiar words are used in a symbolic or metaphoric sense. "To be or not to be" is not an easy idea although the sentence is short and the separate words used would usually be called simple and familiar ones. Readability formulas are not sensitive to such subtle variations in meaning.

Furthermore, the nature of the difficulty of a given piece of writing depends to a great extent upon what we expect a reader to get out of the material. If we ask difficult questions on a passage, even if the passage is fairly simple, the reader may not be able to answer the questions asked and therefore will not understand it by our set criterion.

The reader's purpose in reading and his interest and background in the subject-matter must also be considered by anyone using a readability formula. To say that a given article on chemistry is comfortable reading for average adults because it has a predicted grade level of VII-VIII, is giving an incomplete picture. For those readers who have no interest or no background in chemistry, the article will probably not be comfortable reading and they may get very little meaning from it. For others who are interested in chemistry and do considerable reading in the subject, the same article will probably be most comfortable reading. This difference in ease of reading and comprehension may exist even though both groups of readers have completed approximately eight and one-half years of schooling and have the same general reading ability on a standardized reading test.

Taking account of differences in background is especially important in writing and selecting materials for persons who have a specialized understanding of the field. Thus, in material written for farmers, the inclusion of such words as *barley*, *flax*, *hybrid*, *husk*, *fertilizer*, *mulch* will increase the predicted grade level of the material. But if these words are in the common vocabulary of the farmer, they may not offer any special difficulty in comprehension. This factor, therefore, must be taken into account in dealing with materials having a specialized vocabulary. Thus the direction, "Hand me that Stillson," is perfectly clear to any mechanic but not very meaningful to the layman.

Keeping these cautions in mind, we have found that this formula can be a useful tool in selecting and preparing reading materials that can be understood by specified audiences.

[Vol. XXVII, No. 1]

Educational Research Bulletin

Vol. XVII

February 17, 1948

No. 2

A Formula for Predicting Readability: Instructions

By EDGAR DALE AND JEANNE S. CHALL

AN ARTICLE in the January issue of the EDUCATIONAL RESEARCH BULLETIN discussed the way in which a formula for testing the grade-level difficulty of reading materials was developed.¹ The limitations of the formula, the circumstances under which it is properly applied, and specific examples for its use were given. This article, a continuation of the one just mentioned, gives specific information concerning the technique of using the formula.

The formula is based on two counts—average sentence length and percentage of unfamiliar words (words outside the Dale list of 3000 words). Rules for selecting samples of a text to be analyzed and for computing the average sentence length and percentage of unfamiliar words are presented in this article. As each count is made, it is recorded on a work sheet² where detailed steps are given for arriving at the grade-level of reading difficulty. To illustrate the mechanics of using the formula, we analyzed three samples from a pamphlet, *Your Baby*.³ The various counts and computations are given in the work sheet. The directions to guide the various steps in filling out the work sheet follow.

I. Selecting Samples:

Take approximately 100 words about every tenth page for books.⁴

¹ Dale, Edgar, and Chall, Jeanne S. "A Formula for Predicting Readability," *Educational Research Bulletin*, XXVII (January 21, 1948), pp. 11-20, 28.

² See page 43 [210] of this issue. Mimeographed copies of the work sheet may be obtained from Edgar Dale, Bureau of Educational Research, Ohio State University.

³ See pages 42 [208] and 44 [209].

⁴ When a more exact grading of books is desired, 100-word samples every tenth page will probably give a more reliable measure. See Leifeste, Bertha V., "An Investigation of the Reliability of the Sampling of Reading Material," *Journal of Educational Research*,

For articles, select about four 100-word samples per 2,000 words. Space these samples evenly. For passages of about 200 to 300 words, analyze the entire passage. Never begin or end a sample in the middle of a sentence.

II. Labeling-Work Sheet:

Enter such information as title, author, publisher, date of publication, etc., regarding the sample to be appraised.

III. Counting the Number of Words:

- A. Count the total number of words in the sample.
- B. Count hyphenated words and contractions as one word.
- C. Count numbers as words.

10 is one word.

1947 is one word.

- D. Count compound names of persons and places as one word.
St. John, Van Buren, del Rio, Le Brun, and so on are each counted as one word.
- E. Do not count initials which are part of a name as separate words.
John F.W. St. John is counted as two words—John and F.W. St. John.
- F. Record the number of words under No. 1 of the work sheet.

IV. Counting; the Number of Sentences

- A. Count the number of complete sentences in the sample.
- B. Record this under No. 2 of the work sheet.

V. Counting the Number of Unfamiliar Words:

Words which do not appear on the Dale list⁵ are considered unfamiliar. Underline all unfamiliar words, even if they appear more than once.

⁵ XXXVII (February, 1944), pp. 441-50.

⁵ See the Dale list on pages 25-34 [212-221]

In making this count, special rules are necessary for common and proper nouns, verbs, and other parts of speech. These are given in the section which follows.

A. Common Nouns:

1. Consider familiar all regular plurals and possessives of words on the list.

boy's is familiar because boy is on the list (possessive).

girls is familiar because girl is on the list (plural by adding s). *churches* is familiar because church is on the list (plural by adding es).

armies is familiar because army is on the list (plural by changing y to ies).

2. Count irregular plurals as unfamiliar, even if the singular form appears on the list.

oxen is unfamiliar, although *ox* is on the list.

Several irregular plurals, however, are listed in the word list. When the plural appears as a separate word, or is indicated by the ending in parentheses next to the word, it is considered familiar.

goose-and *geese* both appear on the list and are both considered familiar.

3. Count as unfamiliar a noun that is formed by adding *er* or *r* to a noun or verb appearing on the word list (unless this *er* or *r* form is indicated on the list).

burner is counted as unfamiliar, although *burn* is on the list. *'owner* is considered familiar because it appears on the list, as follows—*own(er)*.

B. Proper Nouns:

1. Names of persons and places are considered familiar.

Japan, *Smith*, and so on, are familiar, even though they do not appear on the word list.

2. Names of organizations, laws, documents, titles of books, movies, and so on generally comprise several words.

- a. When determining the number of words in a sample, count all the words in the name of an organization, law, and the like. *Chicago Building Association* should be counted three words.

Declaration of Independence should be counted three words.

- b. For the unfamiliar word count, consider unfamiliar only words which do not appear on the Dale list, except names of persons or places.

SPECIAL RULE: When the title of an organization, law, and so on is used several times within a sample of 100 words, all the words in the title are counted, no matter how many times they are repeated.

Chicago Building Association is counted one unfamiliar word — *Association*. *Building* and *Chicago* are familiar. *Declaration of Independence* is counted as two unfamiliar words — *of* is on the list.

SPECIAL RULE: When the name of an organization, law, document, and so on is used several times within a sample of 100 words, count it only twice when making the unfamiliar word count.

Security Council, if repeated more than twice within a 100-word sample, is counted as four unfamiliar words.

3. Abbreviations:

- a. In counting the words in a sample, an abbreviation is counted as one word. *Y.M.C.A.* is counted one word.

Nov. is counted one word. *U.S.* is considered one word. *A.M.* and *P.M.* are each counted as one word.

- b. In making the unfamiliar word count, an abbreviation is counted as one unfamiliar word only. *Y.M.C.A.* is considered one unfamiliar word. *Nov.* is considered familiar because the names of the months are on the word list. *U.S.* is considered familiar. *A.M.* and *P.M.* are each considered familiar.

SPECIAL RULE: An abbreviation which is used several

times within a 100-word sample is counted as two unfamiliar words only.

C.I.O. repeated five times in a 100-word sample is counted two unfamiliar words.

C. Verbs:

1. Consider familiar the third-person, singular forms (*s* or *ies* from *y*), present-participle forms (*ing*), past-participle forms (*n*), and past-tense forms (*ed* or *ted* from *y*), when these are added to verbs appearing on the list. The same rule applies when a consonant is doubled before adding *ing* or *ed*.

asks, asking, asked are considered familiar, although only the word *ask* appears on the word list.

dropped and *dropping* are familiar because *drop* is on the list.

D. Adjectives:

1. Comparatives and superlatives of adjectives appearing on the list are considered familiar. The same rule applies if the consonant is doubled before adding *er* or *est*.

longer, prettier, and bravest are familiar because *long, pretty, and brave* are on the list. *red, redder, reddest* are all familiar.

2. Adjectives formed by adding *n* to a proper noun are familiar. For example, *American, Austrian*.
3. Count as unfamiliar an adjective that is formed by adding *y* to a word that appears on the list. But consider the word familiar if it appears in parentheses following the word.

woolly is unfamiliar although *wool* is on the list.

sandy is familiar because it appears on the list, as *sand (y)*.

E. Adverbs:

1. Consider adverbs familiar which are formed by adding *ly* to a word on the list. In most cases *ly* will be indicated following the word.

soundly is familiar because *sound* is on the list.

2. Count as unfamiliar words which add more than *ly*, like *easily*.

F. Hyphenated Words:

Count hyphenated words as unfamiliar if either word in the compound does not appear on the word list. When both appear on the list, the word is familiar.

G. Miscellaneous Special Cases:

1. Words formed by adding *en* to a word on the list (unless the *en* is listed, in parentheses or the word itself appears on the list) are considered unfamiliar.

sharpen is considered unfamiliar although *sharp* is on the list.

golden is considered familiar because it appears on the list, *gold {en}*•

2. Count a word unfamiliar if two or more endings are added to a word on the list.

clippings is considered unfamiliar, although *clip* is on the list.

3. Words on the list to which *-tion*, *-ation*, *-ment*, and other suffixes not previously mentioned are added are considered unfamiliar, unless the word with the ending is included on the list.

treatment is unfamiliar although *treat* is on the list. *protection* is unfamiliar although *protect* is on the list. *preparation* is unfamiliar although *prepare* is on the list.

4. Numbers:

Numerals like *1947*, *18*, and so on, are considered familiar.

H. Record the total number of unfamiliar words, under No. 3 of the work sheet.

The number of words in the sample (No. 1 on the work sheet) have now been recorded, as well as the number of sentences in the sample (No. 2)

and the number of words not on the Dale list (No. 3). The next steps can be followed easily on the work sheet.

VI. Completing the Work Sheet:

1. The average sentence length (No. 4) is computed by dividing the number of words in the sample by the number of sentences in the sample.
2. The Dale score or percentage of words outside the Dale list is computed by dividing the number of words not on the Dale list by the number of words in the sample, and multiplying by 100.
3. Follow through Steps 6 and 7 on the work sheet.⁶
4. Add Nos. 6, 7, and 8 to get the formula raw score.
5. If you have more than one sample to analyze, get an average of the formula raw scores by adding all of these and dividing by the number of samples.
6. Convert the average formula raw score to a corrected grade-level according to the Correction Table given in Table I.

The corrected grade-level indicates the grade at which a book or article can be read with understanding. For example, a book with a corrected grade-level of 7-8 is one which should be within the reading ability of average children in Grades VII and VIII. For adults, the 7-8 grade-level can be compared to the last grade reached. If materials are being selected for persons who have had an average of eight grades of schooling, passages with a corrected grade-level of 7-8 should be within their ability. The corrected grade-levels corresponding to the raw scores obtained from the formula are given in Table I. These will serve to determine the grade-level of materials being appraised with the use of the Dale list.

The population reports of the Bureau of Census are a good source for determining the educational levels of large groups of adults. Statistics on the last grade reached are given in tables headed "Persons 25 Years Old by Years of School Completed," in the 1940 Population, Volume II, *Characteristics of the Population*. Part I contains the statistics for the states, cities, and counties. These are further broken down by sex, race, native and foreign born, urban and rural.

⁶ Copies of the table of multiplications may be obtained from Edgar Dale, Bureau of Educational Research, Ohio State University.

TABLE I

CORRECTION TABLE

Formula Raw Score	Corrected Grade-Levels
4.9 and below	4-th grade and below
5.0 to 5.9	5-6 th grade
6.0 to 6.9	7-8 th grade
7.0 to 7.9	9-10 th grade
8.0 to 8.9	11-12 th grade
9.0 to 9.9	13-15 th grade (college)
10.0 and above	16-(college graduate)

AN illustration of the mechanics of using the formula is given in this part of this article. The following three samples were chosen from a 15-page pamphlet, *Your Baby*, published by the National Tuberculosis Association. The words printed in italics were not found in the Dale list and are by definition unfamiliar words.

Sample I:

A happy, useful life—that's what you want for your baby, isn't it? And because a healthy mind and body are so *necessary* to happiness and long life, you must do all you can to get your baby off to a good start. There is much you can do while he is still a baby to lay the *foundation* for good health and good health habits.

Many things *affect* your baby's health. One was the state of your own health during *pregnancy*, and the *special* care your doctor gave you before the baby was born. Other things important to your child's health are food, clothes, baths, sleep, and habit training. A baby needs a clean, happy place to live, and he must be kept from having any sickness that can be *prevented*.

Sample 2:

Diphtheria used to kill many babies. Today no child need die of *diphtheria*. It is one of the *diseases* for which we have very good *treatment* and almost sure 'prevention. But your baby will not be safe from this *disease* unless he has been protected by *immunization*.

The way to protect your baby is simple. *Physicians usually give injections* of three *doses of toxoid*, three to four weeks apart, generally beginning when a baby is about six months old. Your doctor will tell you that your baby should have this *protection* before his first birthday.

Six months after the last *injection of toxoid*, the *physician* may test your baby to see if another *dose of toxoid* is *necessary*. Before the child enters school an extra shot of *toxoid* is often given.

Sample 3:

The *germs* that cause *tuberculosis* can enter the baby's body through his mouth or be breathed in through his nose. These *germs* come to him on *spray* or *moisture* which the person with *active tuberculosis* breathes or coughs out. *Germ-filled spray* from the mouth or nose may light on the baby's food, his dishes, his toys. The baby's hands may carry *germs* from soiled *objects* to his mouth. Kissing is one way of spreading *TB* as well as other *germs*.

Tuberculosis of the bones or *joints* or of certain organs of the body besides the *lungs* can come to the bottle-fed baby in milk which has not been •*pasteurized* or boiled.

The records for these three samples are given in the work sheet reproduced here as Table II. The average raw score for the three samples was 6.35. By referring to the grade equivalent given in Table I, the correction table, the grade-level of the readability of the pamphlet, 7-8, was determined.

TABLE II
A WORK SHEET FILLED IN FOR THE SAMPLES TAKEN FROM THE PAMPHLET "YOUR BABY"

Article:	Your Baby	Page No.	2	Page No.	7	Page No.	12
Author:				From "A happy...".		From "The germs...".	
Publisher:	Natl TB Assoc.	Date:	1945	To ... prevented."	To ... often given."	To ... or boiled."	
1. Number of words in the sample.....		132		131		111	
2. Number of sentences in the sample.....		7		9		6	
3. Number of words not on Dale List.....		6		20		17	
4. Average sentence length (divide 1 by 2).....		19		15		19	
5. Dale score (divide 3 by 1, multiply by 100).....		5		15		15	
6. Multiply average sentence length (4) by .0496.....		.9424		.7440		.9124	
7. Multiply Dale score (5) by .1579.....		.7895		2.3685		2.3685	
8. Constant.....		3.6365		3.6365		3.6365	
9. Formula raw score (add 6, 7, and 8).....		5.3684		6.7490		6.9174	
Average raw score of <u>2</u> samples....		<u>6.35</u>		Analyzed by	J. S. C.	Date	<u>1/28/48</u>
Average corrected grade-level.....		<u>7-8</u>		Checked by	C. D. C.	Date	<u>1/28/48</u>

THE Dale list of approximately three thousand familiar words represents words that are known in reading by at least 80 per cent of the children in Grade IV. It is presented primarily as a list which gives a significant correlation with reading difficulty. It is not intended as a list of the most important words for children or adults. It includes words that are relatively unimportant and excludes some important ones. To use the list for more than an over-all statistical device which gives a good prediction of readability would be out of harmony with the purpose for which it was constructed.

The technique used for constructing the list was crude. When 80 per cent of the fourth-graders questioned indicated that they knew a word, that word was included in the list. This arbitrary cutting off at the 80-per cent point and the lack of any measure of the importance of these words make exceedingly dubious the wisdom of using individual words in appraising the ease or difficulty of material. For purposes of computing a level of difficulty, however, the percentage of words outside this list is a very good index of the difficulty of reading materials. The terms *familiar* and *unfamiliar* describing words are therefore used here in a statistical sense.

There is, however, a real place for a list of important familiar words, graded in about four levels, for use in the preparation of materials for adults of limited reading ability. At the present time we are experimenting with such a list. It will include such words as *nation* and so on, which tested slightly below the 80-per cent criterion on children, but are important, and for all practical purposes are probably familiar to adults.

The three thousand words which comprise the Dale list are given in the pages which follow.

Dale List of 3,000 Words

a	ah	an	armful	awhile	barrel
able	ahead	and	army	ax	base
aboard	aid	angel	arose	baa	baseball
about	aim	anger	around	babe	basement
above	air	angry	arrange	baby (ies).	basket
absent	airfield	animal	arrive (d)	back	bat
accept	airport	another	arrow	background	batch
accident	airplane	answer	art	backward (s)	bath
account	airship	ant	artist	bacon	bathe
ache (ing)	airy	any	as	bad (ly)	bathing
acorn	alarm	anybody	ash (es)	badge	bathroom
acre	alike	anyhow	aside	bag	bathtub
across	alive	anyone	ask	bake (r)	battle
act	all	anything	asleep	baking	battleship
add	alley	anyway	at	bakery	bay
address	alligator	anywhere	ate	ball	be (ing)
admire	allow	apart	attack	balloon '	beach
adventure	almost	apartment	attend	banana	bead
afar	alone	ape	attention	band	beam
afraid	along	apiece	August	bandage	bean
after	aloud	appear	aunt	bang	bear
afternoon	already	apple	author	banjo	beard
afterward (s)	also	April	auto	bank (er)	beast
again	always	apron	automobile	bar	beaming
against	am	are	autumn	barber	beautiful
age	America	aren't	avenue	bare (ly)	beautify
aged	American	arise	awake (n)	barefoot	beauty
ago	among	arithmetic	away	bark	became
agree	amount	arm	awful (ly)	barn	because

1948—The Dale-Chall Readability Formula

become	bird	bookkeeper	bubble	calendar	cattle
becoming	birth	boom	bucket	calf	caught
bed	birthday	boot	buckle	call (er) (ing)	cause
bedbug	biscuit	born	bud	came	cave
bedroom	bit	borrow	buffalo	camel	ceiling
bedspread	bite	boss	bug	camp	cell
bedtime	biting	both	buggy	campfire	cellar
bee	bitter	bother	build	can	cent
beech	black	bottle	building	canal	center
beef	blackberry	bottom	built	canary	cereal
beefsteak	blackbird	bought	bulb	candle	certain (ly)
beehive	blackboard	bounce	bull	candlestick	chain
been	blackness	bow	bullet	candy	chair
beer	blacksmith	bowl	bum	cane	chalk
beet	blame	bow-wow	bumblebee	cannon	champion
before	blank	box (es)	bump	cannot	chance
beg	blanket	boxcar	bun	canoe	change
began	blast	boxer	bunch	can't	chap
beggar	blaze	boy	bundle	canyon	charge
begged	bleed	boyhood	bunny	cap	charm
begin	bless	bracelet	burn	cape	chart
begun	blew	brake	bury	captain	chatter
behave	blind (s)	bran	bus	car	cheap
behind	blindfold	branch	bush	card	cheat
believe	block	brass	bushel	cardboard	check
bell	blood	brave	business	care	checkers
belong	bloom	bread	busy	careful	cheek
below	blossom	break	but	careless	cheer
belt	blot	breakfast	butcher	carelessness	cheese
beneath	blow	breast	butt	carload	cherry
bench	blue	breath	butter	carpenter	chest
bend	blueberry	breathe	buttercup	carpet	chew
bent	bluebird	breeze	butterfly	carriage	chick
berry (ies)	blue jay	brick	buttermilk	carrot	chicken
besides (s)	blush	bride	butterscotch	carry	chief
best	board	bridge	button	cart	child
bet	boast	bright	buttonhole	carve	childhood
better	boat	brightness	buy	case	children
between	bob	bring	buzz	cash	chill (y)
bib	bobwhite	broad	by	cashier	chimney
bible	body (ies)	broadcast	bye	castle	chin
bicycle	boil (er)	broke(n)	cab	cat	china
bid	bold	brook	cabbage	catbird	chip
big (ger)	bone	broom	cabin	catch	chipmunk
bill	bonnet	brother	cabinet	catcher	chocolate
billboard	boo	brought	cackle	caterpillar	choice
bin	book	brown	cage	catfish	choose
bind	bookcase	brush	cake	catsup	chop

Unlocking Language

chorus	codfish	cow	dad	destroy	downstairs
chose	coffee.	coward (ly)	daddy	devil	downtown
christen	coffeepot	cowboy	daily	dew	dozen
Christmas	coin	cozy	dairy	diamond	drag
church	cold	crab	daisy	did	drain
churn	collar	crack	dam	didn't	drank
cigarette	college	cracker	damage	die (d) (s)	draw (er)
circle	color (ed)	cradle	dame	difference	draw (ing)
circus	colt	cramps	damp	different -	dream
citizen	column	cranberry	dance (r)	dig	-dress
city	comb	crank (y)	dancing	dim	dresser
clang	come	crash	dandy	dime	dressmaker
clap	comfort	crawl	danger (ous)	dine	drew
class	comic	crazy	dare	ding-dong	dried
classmate	coming	cream (y)	dark (ness)	dinner	drift
classroom	company	creek	darling	dip	drill
claw	compare	creep	darn	direct	drink
clay	conductor	crept	dart	direction	drip
clean(er)	cone	-cried	dash	dirt (y)	drive (n)
clear	connect	croak	date	discover	driver
clerk	coo	crook(ed)	daughter	dish	drop
clever	cook (ed)	crop	dawn	dislike	drove
click	cook (ing)	cross (ing)	day	dismiss	drown
cliff	cooky (ie) (s)	cross-eyed	daybreak	ditch	drowsy
climb	cool (er)	crow	daytime	dive	drug
clip	coop	crowd(ed)	dead	diver	drum
cloak	copper	crown	deaf	divide	drunk
clock	copy	cruel	deal	do	dry
close	cord	crumb	dear	dock	duck
closet	cork	crumble	death	doctor	due
cloth	corn	crush	December	does	dug
clothes	corner	crust.	decide	doesn't	dull
clothing	correct	cry (ies)	deck	dog	dumb
cloud (y)	cost	cub	deed	doll	dump
clover	cot	cuff	deep	dollar	during
clown	cottage	cup	deer	dolly	dust(y)
club	cotton	cupboard	defeat	done	duty
cluck	couch	cupful	defend	donkey	dwarf
clump	cough	cure	defense	don't	dwell
coach	could	curl(y)	delight	door	dwelt
coal	couldn't	curtain	den	doorbell	dying
coast	count	curve	dentist	doorknob	each
coat	counter	cushion	depend	doorstep	eager
cob	country	custard	deposit	dope	eagle
cobbler	county	customer	describe	dot	ear
cocoa	course	cut	desert	double	early
coconut	court	cute	deserve	dough	earn
cocoon	cousin	cutting	desire	dove	earth
cod	cover	dab	desk	down	eastern

1948—The Dale-Chall Readability Formula

easy	excited	fellow	flip-flop	French	gift
eat (en)	exciting	felt	float	fresh	gingerbread
edge	excuse	fence	flock	fret	girl
egg	exit	fever	flood	Friday	give (n);
eh	expect	few	floor	fried	giving
eight	explain	fib	flop	friend (ly)	glad (ly)
eighteen	extra	fiddle	flour	friendship	glance
eighth	eye	field	flow	frighten	glass (es)
eighty	eyebrow	fife	flower (y)	frog	glean
either	fable	fifteen	flutter	from	glide
elbow	face	fifth	fly	front	glory
elder	facing	fifty	foam	frost	glove
eldest	fact	fig	fog	frown	glow
electric	factory	fight	foggy	froze	glue
electricity	fail	figure	fold	fruit	go (ing)
elephant	faint	file	folks	fry	goes
eleven	fair	fill	follow (ing)	fudge	goal
elf	fairy	film	fond	fuel	goat
elm	faith	finally	food	full (y)	gobble
else	fake	find	fool	fun	God (g)
elsewhere	fall	fine	foolish	funny	godmother
empty	false	finger	foot	fur	gold (en)
end (ing)	family	finish	football	furniture	goldfish
enemy	fan	fire	footprint	further	golf
engine	fancy	firearm	for	fuzzy	gone
engineer	far	firecracker	forehead	gain	good (s)
English	faraway	fireplace	forest	gallon	good-by (bye)
enjoy	fare	fireworks	forget	gallop	good-looking
enough	farmer	firing	forgive	game	goodness
enter	farm (ing)	first	forgot (ten)	gang	goody
envelope	far-off	fish	fork	garage	goose
equal	farther	fisherman	form	garbage	gooseberry
erase (r)	fashion	fist	fort	garden	got
errand	fast	fit (s)	forth	gas	govern
escape	fasten	five"	fortune	gasoline	government
eve	fat	fix	forty	gate	gown
even	father	flag	forward	gather	grab
evening	fault	flake	fought	gave	gracious
ever	favor	flame	found	gay	grade
every	favorite	flap	fountain	gear	grain
everybody	fear	flash	four	geese	grand
everyday	feast	flashlight	fourteen	general	grandchild
everyone	feather	flat	fourth	gentle	grandchildren
everything	February	flea	fox	gentleman	granddaughter
everywhere	fed	flesh	frame	gentlemen	grandfather
evil	feed	flew	free	geography	grandma
exact	feel	flies	freedom	get	grandmother
except	feet	flight	freeze	getting	grandpa
exchange	fell	flip	freight	giant	grandson

Unlocking Language

grandstand	handle	held	homely	hush	January
grape (s)	handwriting	hell	homesick	hut	jar
grapefruit	hang	he'll	honest	hymn	jaw
grass	happen	hello	honey	I	jay
grasshopper	happily	helmet	honeybee	ice	jelly
grateful	happiness	help (er)	honeymoon	icy	jellyfish
grave	happy	helpful	honk	I'd	jerk
gravel	harbor	hem	honor	idea	jig
graveyard	hard	hen	hood	ideal	job
gravy	hardly	henhouse	hoof	if	jockey
gray	hardship	her (s)	hook	ill	join
graze	hardware	herd	hoop	I'll	joke
grease	hare	here	hop	I'm	joking
great	hark	here's	hope (ful)	important	jolly
green	harm	hero'	hopeless	impossible	journey
greet	harness	herself	horn	improve	joy (ful)
grew	harp	he's	horse	in	joyous
grind	harvest	hey	horseback	inch (es)	judge
groan	has	hickory	horseshoe	income	jug
grocery	hasn't	hid	hose	indeed	juice
ground	haste (n)	hidden	hospital	Indian	juicy
group	hasty	hide	host	indoors	July
grove	hat	high	hot	ink	jump
grow	hatch	highway	hotel	inn	June
guard	hatchet	hill	hound	insect	junior
guess	hate	hillside	hour	inside	junk
guest	haul	hilltop	house	instant	just
guide	have	hilly	housetop	instead	keen
gulf	haven't	him	housewife	insult	keep
gum	having	himself	housework	intend	kept
gun	hawk	hind	how	interested	kettle
gunpowder	hay	hint	however	interesting	key
guy	hayfield	hip	howl	into	kick
ha	haystack	hire	hug	invite	kid
habit	he	his	huge	iron	killed (ed)
had	head	hiss	hum	is	kind (ly)
hadn't	headache	history	humble	island	kindness
hail	heal	hit	hump	isn't	king
hair	health (y)	hitch	hundred	it	kingdom
haircut	heap	hive	hung	its	kiss
hairpin	hear (ing)	ho	hunger	it's	kitchen
half	heard	hoe	hungry	itself	kite
hall	heart	hog	hunk	I've	kitten
halt	heat (er)	hold (er)	hunt (er)	ivory	kitty
ham	heaven	hole	hurrah	ivy	knee
hammer	heavy	holiday	hurried	jacket	kneel
hand	he'd	hollow	hurry	jacks	knew
handful	heel	holy	hurt	jail	knife
handkerchief	height	home	husband	jam	knit

1948—The Dale-Chall Readability Formula

knives	lend	lonesome	market	minute	name
knob	length	long	marriage	mirror	nap
knock	less	look	married	mischief	napkin
knot	lesson	lookout	marry	miss (M)	narrow
know	let	loop	mask	misspell	nasty
known	let's	loose	mast	mistake	naughty
lace	letter	lord	master	misty	navy
lad	letting	lose (r)	mat	mitt	near
ladder	lettuce	loss	match	mitten	nearby
ladies	level	lost	matter	mix	nearly
lady	liberty	lot	mattress	moment	neat
laid	library	loud	may (M)	Monday	neck
lake	lice	love	maybe	money	necktie
lamb	lick	lovely	mayor	monkey	need
lame	lid	lover	maypole	month	needle
lamp	lie	low	me	moo	needn't
land	life	luck (y)	meadow	moon	Negro
lane	lift	lumber	meal	moonlight	neighbor
language	light (ness)	lump	mean (s)	moose	neighborhood
lantern	lightning	lunch	meant	mop	neither
lap	like	lying	measure	more	nerve
lard	likely	ma	meat	morning	nest
large	liking	machine	medicine	morrow	net
lash	lily	machinery	meet (ing)	moss	never
lass	limb	mad	melt	most (ly)	nevermore
last	lime	made	member	mother'	new
late	limp	magazine	men	motor	news
laugh	line	magic	mend	mount	newspaper
laundry	linen	maid	meow	mountain	next
law	lion	mail	merry	mouse	nibble
lawn	lip	mailbox	mess	mouth	nice
lawyer	list	mailman	message	move	nickel
lay	listen	major	met	movie	night
lazy	lit	make	metal	movies	nightgown
lead	little	making	mew	moving	nine
leader	live (s)	male	mice	mow	nineteen
leaf	lively	mama	middle	Mr., Mrs.	ninety
leak	liver	mamma	midnight	much	no
lean	living-	man	might (y)	mud	nobody
leap	lizard	manager	mile	muddy	nod
learn (ed)	load	mane	milk	mug	noise
least	loaf	manger	milkman	mule	noisy
leather	loan	many	mill	multiply	none
leave (ing)	loaves	map	miller	murder	noon
led	lock	maple	million	music	nor
left	locomotive	marble	mind	must	north (ern)
leg	log	march (M)	mine	my	nose
lemon	lone	mare	miner	myself	not
lemonade	lonely	mark	mint	nail	note

Unlocking Language

nothing	outfit	passenger	pin	popped	pussycat
notice	outlaw	past	pine	porch	put
November	outline	paste	pineapple	pork	putting
now	outside	pasture	pink	possible	puzzle
nowhere	outward	pat	pint	post	quack
number	oven	patch	pipe	postage	quart
nurse	over	path	pistol	postman	quarter
nut	overalls	patter	pit	pot	queen
oak	overcoat	pave	pitch	potato (es)	queer
oar	overeat	pavement	pitcher	pound	question
oatmeal	overhead	paw	pity	pour	quick (ly)
oats	overhear	pay	place	powder	quiet
obey	overnight	payment	plain	power (ful)	quilt
ocean	overturn	pea (s)	plan	praise	quit
o'clock	owe	peace (ful)	plane	pray	quite
October	owing	peach (es)	plant	prayer	rabbit
odd	owl	peak	plate	prepare	race
of	own (er)	peanut	platform	present	rack
off	ox	pear	platter	pretty	radio
offer	pa	pearl	play (er)	price	radish
office	pace	peck	playground	prick	tag
officer	pack	peek	playhouse	prince	rail
often	package	peel	playmate	princess	railroad
oh	pad	peep	plaything	print	railway
oil	page	peg	pleasant	prison	rain (y)
old	paid	pen	please	prize	rainbow
old-	pail	pencil	pleasure	promise	raise
fashioned	pain (ful)	penny	plenty	proper	raisin
on	paint (er)	people	plow	protect	rake
once	painting	pepper	plug	proud	ram
one	pair	peppermint	plum	prove	ran
onion	pal	perfume	pocket	prune	ranch
only	palace	perhaps	pocketbook	public	rang
onward	pale	person	poem	puddle	rap
open	pan	pet	point	puff	rapidly
or	pancake	phone	poison	pull	rat
orange	pane	piano	poke	pump	rate
orchard	pansy	pick	pole	pumpkin	rather
order	pants	pickle	police	punch	rattle
ore	papa	picnic	policeman	punish	raw
organ	paper	picture	polish	pup	ray
other	parade	pie	polite	pupil	reach
otherwise	pardon	piece	pond	puppy	read
ouch	parent	pig	ponies	pure	reader
ought	park	pigeon	pony	purple	reading
our (s)	part (ly)	piggy	pool	purse	ready
ourselves	partner	pile	poor	push	real
out	party	pill	pop	puss	really
outdoors	pass	pillow	popcorn	pussy	reap

1948—The Dale-Chall Readability Formula

rear	rock (y)	sand (y)	self	shirt	sixteen
reason	rocket	sandwich	selfish	shock	sixth
rebuild	rode	sang	sell	shoe	sixty
receive	roll	sank	send	shoemaker	size
recess	roller	sap	sense	shone	skate
record	roof	sash	sent	shook	skater
red	room	sat	sentence	shoot	ski
redbird	rooster	satin	separate	shop	skin
redbreast	root	satisfactory	September	shopping	skip
refuse	rope	Saturday	servant	shore	skirt
reindeer	rose	sausage	serve	short	sky
rejoice	rosebud	savage	service	shot	slam
remain	rot	save	set	should	slap
remember	rotten	savings	setting	shoulder	slate
remind	rough	saw	settle	shouldn't	slave
remove	round	say	settlement	shout	sled
rent	route	scab	seven	shovel	sleep (y)
repair	row	scales	seventeen	show	sleeve
repay	rowboat	scare	seventh	shower	sleigh
repeat	royal	scarf	seventy	shut	slept
report	rub	school	several	shy	slice
rest	rubbed	schoolboy	sew	sick (ness)	slid
return	rubber	schoolhouse	shade	side	slide
review	rubbish	schoolmaster	shadow	sidewalk	sling
reward	rug	schoolroom	shady	sideways	slip
rib	rule (r)	scorch	shake (r)	sigh	slipped
ribbon	rumble	score	shaking	sight	slipper
rice	run	scrap	shall	sign	slippery
rich	rung	scrape	shame	silence	slit
rid	runner	scratch	shan't	silent	slow (ly)
riddle	running	scream	shape	silk	sly
ride (r)	rush	screen	share	sill	smack
riding	rust (y)	screw	sharp	silly	small
right	rye	scrub	shave	silver	smart
rim	sack	sea	she	simple	smell
ring	sad	seal	she'd	sin	smile
rip	saddle	seam	she'll	since	smoke
ripe	sadness	search	she's	sing	smooth
rise	safe	season	shear (s)	singer	snail
rising	safety	seat	shed	single	snake
river	said	second	sheep	sink	snap
road	sail	secret	sheet	sip	snapping
roadside	sailboat	see (ing)	shelf	sir	sneeze
roar	sailor	seed	shell	sis	snow (y)
roast	saint	seek	shepherd	sissy	snowball
rob	salad	seem	shine	sister	snowflake
robber	sale	seen	shining	sit '	snuff
robe	salt	seesaw	shiny	sitting	snug
robin	same	select	ship	six	so

Unlocking Language

soak	splash	stocking	sunset	taught	tho
soap	spoil	stole	sunshine	tax	thorn
sob	spoke	stone	supper	tea	those
socks	spook	stood	suppose	teach (er)	though
sod	spoon	stool	sure (ly)	team	thought
soda	sport.	stoop	surface	tear	thousand
sofa	spot	stop	surprise	tease	thread
soft-	spread	stopped	swallow	teaspoon	three
soil	spring	stopping	swam	teeth	threw
sold	springtime	store	swamp	telephone	throat
soldier	sprinkle	stork	swan	tell	throne
sole	square	stories	swat	temper.	through
some	squash	storm (y)	swear	ten	throw (n)
somebody	squeak	story	sweat	tennis	thumb
somehow	squeeze	stove	sweater	tent	thunder
someone	squirrel	straight	sweep	term	Thursday
something	stable	strange (r)	sweet (ness)	terrible	thy
sometime (s)	stack	strap	sweetheart	test	tick
somewhere	stage	straw	swell	than	ticket
son	stair	strawberry	swept	thank (s)	tickle
song	stall	stream	swift	thankful	tie
soon	stamp	street	swim	Thanks	tiger
sore	stand	stretch	swimming	giving	tight
sorrow	star	string	swing	that	till
sorry	stare	strip	switch	that's	time
sort	start	stripes	sword	the	tin
soul	starve	strong	swore	theater	tinkle
sound	state	stuck	table	thee	tiny
soup	station	study	tablecloth	their	tip
sour	stay	stuff	tablespoon	them	tiptoe
south (ern)	steak	stump	tablet	then	tire
space	steal	stung	tack	there	tired
spade	steam	subject	tag	these	'tis
spank	steamboat	such	tail	they	title
sparrow	steamer	suck	tailor	they'd	to
speak (er)	steel	sudden	take (n)	they'll	toad
spear	steep	suffer	taking	they're	toadstool
speech	steeple	sugar	tale	they've	toast
speed	steer	suit	talk (er)	hick	tobacco
spell (ing)	stem	sum	tall	thief	today
spend	step	summer	tame	thimble	toe
spent	stepping	sun	tan	thin	together
spider	stick (y)	Sunday	tank	thing	toilet
spike	stiff	sunflower	tap	think	told
spill	still (ness)	sung	tape	third	tomato
spin	sting	sunk	tar	thirsty	tomorrow
spinach	stir	sunlight	tardy	thirteen	ton
spirit	stitch	sunny	task	thirty	tone
spit	stock	sunrise	taste	this	tongue

1948—The Dale-Chall Readability Formula

tonight	tulip	valentine	weaken	whom	workman
too	tumble	valley	wealth	who's	world
took	tune	valuable	weapon	whose	worm
tool	tunnel	value	wear	why	worn
toot	turkey	vase	weary	wicked	worry
tooth	turn	vegetable	weather	wide	worse
toothbrush	turtle	velvet	weave	wife	worst
toothpick	twelve	very	web	wiggle	worth
top	twenty	vessel	we'd	wild	would
tore	twice	victory	wedding	wildcat	wouldn't
torn	twig	view	Wednesday	will	wound
toss	twin	village	wee	willing	wove
touch	two	vine	weed	willow	wrap
tow	ugly	violet	week	win	wrapped
toward (s)	umbrella	visit	we'll	wind (y)	wreck
towel	uncle	visitor	weep	windmill	wren
tower	under	voice	weigh	window	wring
town	understand	vote	welcome	wine	write
toy	underwear	wag	well	wing	writing
trace	undress	wagon	went	wink	written
track	unfair	waist	were	winner	wrong
trade	unfinished	wait	we're	winter	wrote
tram	unfold	wake (n)	west (ern)	wipe	wrungr
tramp	unfriendly	walk	wet	wire	yard
trap	unhappy	wall	we've	wise	yarn
tray	unhurt	walnut	whale	wish	year-
treasure	uniform	want	what	wit	yell
treat	United	war	what's	witch	yellow
tree	States	warm	wheat	with	yes
trick	unkind	warn	wheel	without	yesterday
tricycle	unknown	was	when	woke	yet
tried	unless	wash (er)	whenever	wolf	yolk
trim	unpleasant	washtub	where	woman	yonder
trip	until	wasn't	which	women	you
trolley	unwilling	waste	while	won	you'd
trouble	up	watch	whip	wonder	you'll
truck	upon	watchman	whipped	wonderful	young
true	upper	water	whirl	won't	youngster
truly	upset	watermelon	whisky	wood (en)	your (s)
trunk	upside	waterproof	whisper	woodpecker	you're
trust	upstairs	wave	whistle	woods	yourself
truth	uptown	wax	white	wool	yourselves
try	upward	way	who	woolen	youth
tub	us	wayside	who'd	word	you've
Tuesday	use (d)	we	whole	wore	
tug	useful	weak (ness)	who'll	work (er)-	

1948—Rudolf Flesch and Reading Ease

Introduction

THE ONE most responsible for publicizing the need for readability was Rudolf Flesch, a colleague of Lorge at Columbia University. Besides working as a readability consultant, lecturer, and teacher of writing, he published a number of studies and nearly 20 popular books on writing, English usage, and readability.

His best-selling books included *The Art of Plain Talk* (1946), *The Art of Readable Writing* (1949), *The Art of Clear Thinking* (1951), *Why Johnny Can't Read —And What You Can Do About It* (1955), *The ABC of Style: A Guide to Plain English* (1964), *How to Write in Plain English: A Book for Lawyers and Consumers* (1979).

Flesch was born in Austria and got a degree in law from the University of Vienna in 1933. He practiced law until 1938, when he came to the U.S. as a refugee from the Nazis.

Since his law degree was not recognized, he worked several other jobs, one of them in the shipping department of a New York book manufacturer.

In 1939, he received a refugee's scholarship at Columbia University. In 1940, he received a bachelor's degree with honors in library science. That same year, he became an assistant to Lyman Bryson in the Teachers' College Readability Lab.

In 1942, Flesch received a master's degree in adult education. The next year, he received a Ph.D. in educational research for his dissertation, "Marks of a Readable Style" (1943). This paper set a course for his career and that of readability.

In his dissertation, Flesch published his first readability formula for measuring adult reading material. One of the variables it used was affixes and another was "personal references" such as personal pronouns and names. Publishers quickly discovered that Flesch's formula could increase readership by 40 to 60 percent. Investigators in many fields of communication began using it in their studies.

In the 1948 article reprinted here, "A New Readability Yardstick,"

Flesch published a second formula with two parts¹. The first part, the Reading Ease formula, dropped the use of affixes and used only two variables, the number of syllables and the number of sentences for each 100-word sample. It predicts reading ease on a scale from 1 to 100, with 30 being “very difficult” and 70 being “easy.” Flesch (p. 225) wrote that a score of 100 indicates reading matter understood by readers who have completed the fourth grade and are, in the language of the U.S. Census barely “functionally literate.”

The second part of Flesch’s formula predicts human interest by counting the number of personal words (such as pronouns and names) and personal sentences (such as quotes, exclamations, and incomplete sentences).

The formula for the updated Flesch Reading Ease score is:

$$\text{Score} = 206.835 - (1.015 \times \text{ASL}) - (84.6 \times \text{ASW})$$

Where:

Score = position on a scale of 0 (difficult) to 100 (easy), with 30 = very difficult and 70 = suitable for adult audiences.

ASL = average sentence length (the number of words divided by the number of sentences).

ASW = average number of syllables per word (the number of syllables divided by the number of words).

This formula correlated .70 with the 1925 McCall-Crabbs reading tests and .64 with the 1950 version of the same tests.

In *The Art of Readable Writing*, Flesch, described his Reading Ease scale in this way:

Reading Ease Score	Style Description	Estimated Reading Grade	Estimated Percent of U.S. Adults (1949)
0 to 30:	Very Difficult	College graduate	4.5
30 to 40:	Difficult	13 th to 16 th grade	33
50 to 60:	Fairly Difficult	10 th to 12 th grade	54
60 to 70:	Standard	8 th and 9 th grade	83
70 to 80:	Fairly Easy	7 th grade	88
80 to 90:	Easy	6 th grade	91
90 to 100:	Very Easy	5 th grade	93

¹ Flesch, R. 1948. “A New Readability Yardstick.” *Journal of Applied Psychology*, Vol. 32. No. 3, pp. 221-233.

Unlocking Language

Flesch's Reading Ease formula became the most widely used formula and one of the most tested and reliable (Chall 1958, Klare 1963).

In an attempt to further simplify the Flesch Reading Ease formula, Farr, Jenkins, and Paterson (1951) substituted the average number of one-syllable words per hundred words for Flesch's syllable count. The modified formula is:

$$\text{New Reading Ease score} = 1.599 \text{ nosw} - 1.015 \text{ sl} - 31.517$$

Where: nosw = number of one-syllable words per 100 words;

sl = average sentence length in words

This formula correlates better than .90 with the original Flesch Reading Ease formula and .70 with 75% comprehension of 100-word samplings of the McCall-Crabbs reading lessons. In 1976, a study commissioned by the U.S. Navy modified the Reading Ease formula to produce a grade-level score, This popular formula is known as the Flesch-Kincaid formula, the Flesch Grade-Scale formula or the Kincaid formula.

In 1949, Flesch published the results of a 10-year study of the editorial content of several magazines. He found that:

- About 45% of the population can read The Saturday Evening Post.
- Nearly 50% of the population can read McCall's, Ladies Home Journal, and Woman's Home Companion.
- Slightly over 50% can read American Magazine.
- 80% of the population can read Modern Screen, Photoplay, and three confession magazines.

Flesch, in *The Art of Plain Talk* (1949) compared the reading scores of popular magazines with other variables:

Style	Flesch Reading Ease Score	Average Sentence Length in Words	Average No. of Syll. Per 100 Words	Type of Magazine	Estimated School Grade Completed	Estimated Percent of U.S. Adults
Very Easy	90 to 100	8 or less	123 or less	Comics	4th grade	93
Easy	80 to 90	11	131	Pulp fiction	5th grade	91
Fairly Easy	70 to 80	14	139	Slick fiction	6th grade	88
Standard	60 to 70	17	147	Digests	7th or 8th grades	83
Fairly Difficult	50 to 60	21	155	Quality	Some high school	54
Difficult	30 to 50	25	167	Academic	High school or some college	33
Very Difficult	0 to 30	29 or more	192 or more	Scientific	College	4.5

Flesch's 1949 analysis of the readability of adult reading materials.

Flesch's work had an enormous impact on journalism. Like Robert Gunning, who worked with the United Press, Flesch was a consultant with the Associated Press. Their work helped to bring down the reading grade level of front-page stories from the 16th to the 11th grade level, where they remain today.

Flesch's most important and most popular work was the 1955 *Why Johnny Can't Read*, which put him on the front lines of the battle for teaching phonics. He followed up that work in 1981 with *Why Johnny Still Can't Read: A New Look at the Scandal of Our Schools*. He practiced what he preached by taking his argument to the public with his clear, vigorous, and straight-forward style. His writings changed educational policy, helped students read, and motivated millions of adults to think and write more clearly.

—WHD

Journal of Applied Psychology

Vol. 32

3 June, 1948

No. 3

A New Readability Yardstick*

Rudolf Flesch

Dobbs Ferry, N. Y.

IN 1943 the writer developed a statistical formula for the objective measurement of readability (comprehension difficulty) (5, 6). The formula was based on a count of three language elements: average sentence length in words, number of affixes, and number of references to people. Since its publication, the formula has been put to use in a wide variety of fields. For example, it has been applied to newspaper reports (9, 20), advertising copy (1), government publications (19), bulletins and leaflets for farmers (3), materials for adult education (4), and children's books (12). Its validity has been reaffirmed by five independent studies; the formula ratings of psychology textbooks substantially agreed with ratings by students and teachers (17); the formula scores rated specially edited radio news, newsmagazine, and Sunday news-summary copy "more readable" than comparable newspaper reports (18); advertisements, rated "more readable" by the formula, showed higher readership figures (7); and articles that were simplified with the aid of the formula brought increased readership in two successive split-run tests (13, 14). Since 1943, a number of academic institutions have incorporated the formula in the curriculum of courses in composition, creative writing, journalism, and advertising; it has also been used as the basis of several graduate research projects.

Because of this wide application, it seemed worthwhile to re-examine the formula and to analyze its shortcomings. One of these is to be traced

* Samples from the main body of this paper, when tested for readability by the method here proposed, had an average "reading ease" score of 30 and a "human interest" score of 0. Presumably, the paper is easier to read than most other articles appearing in scientific journals. The section, "The Formulas Restated/," which contains directions for users of the formulas, has a "reading ease" score of 79 and a "human interest" score of 42—which puts that portion of the article in the class of a good cookbook.

to the basic structure of the formula; others are the results of difficulties in its application.

The structural shortcoming of the formula is the fact that it does not always show the high readability of direct, conversational writing. For example, in the study of psychology texts mentioned above (17), the score for Koffka's *Principles of gestalt psychology* ("the students' choice for uniform readability") was 5.4 ("difficult"); yet William James' *Principles of Psychology*, a classic example of readability, rated 6.0 (bordering on "very difficult"). Similarly, the formula consistently rates the popular *Reader's Digest* more readable than the sophisticated *New Yorker* magazine, although many educated readers consider the *Reader's Digest* dull and the sprightly *New Yorker* ten times as readable.

Aside from that, the practical application of the formula led to several minor misinterpretations. Sentence length, for instance, is the element with the heaviest weight; it is also the easiest to measure. As a result, this feature of the formula is often overemphasized, sometimes to the exclusion of the others—as in the directives that have been issued to staff writers of the Associated Press and the *New York Times*, recommending the use of shorter sentences in "leads." On the other hand, the second element—number of affixes—seems often difficult to apply; users of the formula found this count particularly tedious and admitted to uncertainty in spotting affixes. The third element—references to people—raised no such questions; but it was sometimes felt to be arbitrary and the underlying principle was often misunderstood.

In addition, many people found it hard to get used to the scoring system, which generally ranges from 0 ("very easy") to 7 ("very difficult"). Also, the average time needed to test a 100-word sample is six minutes (4). This makes the application of the formula considerably faster than that of earlier formulas, which required reference to word lists (e.g. Gray-Leary (8) or Lorge (10)), but it is still too long for practical use.

The revision of the formula presented in this paper is an attempt to overcome these shortcomings and make the formula a more useful instrument.

Procedure

The criterion used in the original formula was McCall-Crabbs' *Standard test lessons in reading* (11). The formula was so constructed that it predicted the average grade level of a child who could answer correctly

three-quarters of the test questions asked about a given passage. Its multiple correlation coefficient was $R = .74$. It was partly based on statistical findings established in an earlier study by Lorge (10).

For many obvious reasons, the grade level of children answering test questions is not the best criterion for general readability. Data about the ease and interests with which adults will read selected passages would be far better. But such data were not available at the time the first formula was developed, and they are still unavailable today. So McCall-Crabbs' *Standard test lessons* are still the best and most extensive criterion that can be found; therefore they were used again for the revision. In reanalyzing the test passages, the following elements were used:

(1) *Average Sentence Length in Words.* The same element was used in the previous formula, but the correlation coefficient used was taken from Lorge's earlier findings. In the present study this coefficient was recomputed.

(2) *Average word length in syllables*, expressed as the number of syllables per 100 words. The hypothesis was that this measure would furnish results similar to the affix count in the earlier formula. Syllables are obviously easier to count than affixes since this work can be reduced to a mechanical routine.

(3) *Average Percentage of "Personal Words."* The same element was used in the earlier formula. However, the opportunity was used to test a clarified definition, which made no significant difference in correlation. The new definition was stated as follows: All nouns with natural gender; all pronouns except neuter pronouns; and the words *people* (used with the plural verb) and *folks*.

(4) *Average Percentage of "Personal Sentences."* This new element was designed to correct the structural shortcoming of the earlier formula, mentioned above. By hypothesis, it tests the conversational quality and the story interest of the passage analyzed. It was defined as the percentage of the following sentences: Spoken sentences, marked by quotation marks or otherwise; questions, commands, requests, and other sentences directly addressed to the reader; exclamations; and grammatically incomplete sentences whose meaning has to be inferred from the context.

To make the prediction more accurate, 13 of the 376 McCall-Crabbs' passages that contained poetry or problems in arithmetic were omitted in

the count of the first two elements, which are designed to test solely prose comprehension. However, these 13 passages were retained in the count of the last two elements, which are designed to test human interest.

Following the procedure in the earlier study, intercorrelations were then computed. However, multiple correlation of the four elements with the criterion showed no significant gain in prediction value over the earlier formula in spite of the significant prediction value of the additional fourth element by itself ($r = -.27$). Therefore, two multiple-correlation regression formulas were computed: one using the first two elements and one using the last two. This procedure had the advantage of giving independent predictions of the reading ease and the human interest of a given passage.

Finally, the resulting twin formulas were expressed in such a way that maximum readability (in both formulas) had a value of 100, and minimum readability a value of 0. This was done to make the scores more readily understandable for the practical user.

Table 1

Correlations, Means, Standard Deviations, and Regression Weights
of Word and Sentence Length

	<i>sl</i>	<i>C₅₀</i>	\bar{X}	s	β
<i>wl</i>	.4644	.6648	134.2208	13.6845	.5422
<i>sl</i>	—	.5157*	16.5213	5.5509	.2639

* After the preparation of this paper two articles appeared that pointed out a computational error affecting the writer's original formula (Dale, E. and Chall, Jeanne S. A formula for predicting readability. *Educ. Res. Bull.*, Ohio St. Univ., 1948, 27, 11-20, 28; Lorge, I. The Lorge and Flesch readability formulae: a correction. *Sch. & Soc.*, 1948, 67, 141-142). The error concerned the correlation coefficient between sentence length and the criterion, which had originally been reported by Lorge as .6174; the writer, acknowledging his debt to Lorge, used that figure without recomputation. The corrected correlation coefficient is now reported as .4681 by Dale and Chall, and as .467 by Lorge; this corresponds closely to the figure of .5157 reported in Table 1, considering the fact that the writer now used a slightly better criterion of 363 passages for sentence length. In other words, the formula presented in this paper incidentally and independently also corrects the error found by Dale and Chall and by Lorge.

Table 2

Correlations, Means, Standard Deviations, and Regression Weights
of Personal Words and Sentences

	<i>ps</i>	C_{50}	\bar{X}	s	β
<i>pw</i>	.2268	-.3881	7.34578	5.5175	-.3446
<i>ps</i>	—	-.2699	29.5745	35.58.22	-.1917

Findings

The intercorrelations, means, standard deviations, and regression weights found are shown in Tables 1, 2, and 3. The following symbols were used: *wl* for word length (syllables per 100 words), *sl* for sentence length in words, *pw* for percentage of "personal words," *ps* for percentage of "personal sentences," C_{50} for the average grade of children who could answer one-half of the test questions correctly, and C_{75} for the average grade of children who could answer three-quarters of the test questions correctly.

Table 3
Means and Standard Deviations of Two Criteria

	\bar{X}	s
C_{50}	5.4973	1.3877
C_{75}	7.3484	2.1345

The two regression formulas based on these correlations are:

Formula A (for predicting "reading ease"): $RE = 206.835 - .846 wl - 1.015 sl$.

The scores computed by this formula have a range from 0 to 100 for almost all samples taken from ordinary prose. A score of 100 corresponds to the prediction that a child who has completed fourth grade will be able to answer correctly three-quarters of the test questions to be asked about the passage that is being rated; in other words, a score of 100 indicates reading matter that is understandable for persons who have completed fourth grade and are, in the language of the U. S. Census, barely "functionally literate." The range of 100 points was arrived at by multiplying the grade level prediction by 10, so that a point on the formula scale cor-

responds to one-tenth of a grade. However, this relationship holds true only up to about seventh grade; beyond that, the formula under-rates grade level to an increasing degree. Finally, the formula—which predicted grade level and, therefore, difficulty—was "turned around" by reversing the signs to predict "reading ease." (Before this transformation, the formula read: $C_{75} = .0846 wl + .1015 sl - 5.6835$.) The multiple correlation coefficient of this formula is $R = .7047$.

Formula B (for predicting "human interest"): $HI = 3.635 pw + .314 ps$.

Scores computed by this formula, too, have a range from 0 to 100. A score of 100 has the same meaning as in Formula A. It indicates reading matter with enough human interest to suit the reading skills and habits of a barely "functionally literate" person. A score of 0, however, means here simply that the passage contains neither "personal words" nor "personal sentences"; in contrast to Formula A, the two elements counted here may be totally absent. Since the zero point could be fixed in this way, the scoring was arrived at by dividing the range between 0 (absence of both elements) and 100 (prediction of completed fourth grade) by 100. The formula therefore contains no statistical constant. The signs were reversed in the same fashion as in Formula A. (Before transformation, this formula read: $C_{75} = -.1333 pw - .0115 ps + 8.6673$.) The multiple correlation coefficient of this formula is $R = .4306$.

Since the correlations of three of the four elements with the criterion C_{75} , were higher than with the criterion C_{50} , the multiple correlation with the criterion C_{50} was computed first. As a second step, the values so found were used to predict criterion C_{75} , since it seemed obviously more desirable to predict 75% comprehension than 50% comprehension.

The correlation between the word length factor (syllable count) and the corresponding affix count in the earlier formula was found to be $r = .87$. For practical purposes the two measures may therefore be considered equivalent.

The number of affixes per 100 words (a) can be predicted from the syllable count (wl) by the formula: $a = .6832 wl - 66.6017$. Conversely, the number of syllables per 100 words (wl) can be predicted from the number of affixes (a) by the formula: $wl = 1.49 a + 94.56$.

Comment

It is hoped that the two new formulas will prove more useful than the earlier formula. Formula A alone, with a correlation coefficient of .70, has almost as high a prediction value as the combined earlier formula whose correlation coefficient was .74. Formula B has a much lower correlation coefficient of .43 and, accordingly, does not seem to contribute much to the measurement of readability. It should be remembered, however, that because of the criterion used, Formula B predicts only the effect of the two "human interest" elements on *comprehension*; in other words, the correlation coefficient shows only to what extent human interest in a given text will make the reader understand it better. The real value of this formula, however, lies in the fact that human interest will also increase the reader's attention and his motivation for continued reading.

In addition, the two new formulas will be more useful for the teaching of writing, since the added factor and the division into two parts will show specific faults in writing more clearly.

The significance of Formula A will be more easily understood when it is realized that the measurement of word length is indirectly a measurement of word complexity (as mentioned above, the correlation is $r = .87$) and that word complexity in turn is indirectly a measurement of abstraction: the correlation between the number of affixes and that of abstract words was found to be .78 (5). Similarly, the measurement of sentence length is indirectly a measurement of sentence complexity. In two independent studies the correlation between these two factors was found to be .775 (8) and .72 (15). Sentence complexity, in turn, may again be considered as a measure of abstraction. Formula A, therefore, is essentially a test of the level of abstraction.

It seems hardly necessary to prove the importance of human interest in reading, as tested by Formula B. That people are most interested in other people is an old truism. And the readability value of written dialogue, as tested by the added element, is well described in the following, oddly parallel quotations from a printer and a novelist: "Have you ever watched people at a library selecting books for home reading? Other things being equal, if they see enough pages that . . . promise interesting dialogue, they are much more apt to put the book under their arm and walk away with it, than if they see too many solid pages . . . which always suggest hard work" (16). "What is the use of a book without pictures or conversa-

tions?" thought Alice just before the White Rabbit ran by, in condemnation of the book her sister was reading, and this childish comment is supported by novel-readers of all degrees of intelligence. Long close paragraphs of print are in themselves apt to dismay the less serious readers and their instinct here is a sound one, for an excess of summary and an insufficiency of scene in a novel make the story seem remote, without bite, second-hand. . . . A great part of the vigor, the vivacity and the readability of Dickens derives from his innumerable interweavings of scene and summary; his general method is to keep summary to the barest essential minimum, a mere sentence or two here and there between the incredibly fertile burgeoning of his scenes" (2).

Table 4

Comparative Analysis of *The New Yorker* (October 26, 1946) and the
Reader's Digest (November, 1946)

	<i>New Yorker</i>	<i>Reader's Digest</i>
Old Formula:		
Average sentence length in words	20	16
Affixes per 100 words	36	34
Personal words per 100 words	10	8
Readability score	3.59	3.05
New Formula A:		
Average sentence length in words	20	16
Syllables per 100 words	148	145
"Reading ease" score	61	68
New Formula B:		
Personal words per 100 words	10	8
Personal sentences per 100 sentences	39	15
"Human interest" score	49	34

In preliminary tests of the formulas, the following results were found: When the newly isolated fourth element ("personal sentences") was applied to the psychology texts by Koffka and James mentioned above (17), it was found that the percentage of "personal sentences" in Koffka was negligible (4%), whereas in James's first volume it was 16% and in his second volume 10%. A striking example of this difference in style is the

following of James's "personal sentences": "Ask half the common drunkards you know why it is that they fall so often prey to temptation, and they will say that most of the time they cannot tell." This sentence shows well the aspect of readability that eluded the earlier formula.

When the old and the new formulas were applied to two random copies of the *New Yorker* (October 26, 1946) and the *Reader's Digest* (November 1946), the results were as shown in Table 4.

As can be seen, the old formula rated the *Reader's Digest* significantly more readable than the *New Yorker*; the new formula A also shows that the *Reader's Digest* is significantly easier to read. But the new formula B clearly shows a large difference in human interest in favor of the *New Yorker*.

The Formulas Restated

For practical application, the formulas may be restated this way: To measure the readability ("reading ease" and "human interest") of a piece of writing, go through the following steps:

Step 1. Unless you want to test a whole piece of writing, take samples. Take enough samples to make a fair test (say, three to five of an article and 25 to 30 of a book). Don't try to pick "good" or "typical" samples. Go by a strictly numerical scheme. For instance, take every third paragraph or every other page. Each sample should start at the beginning of a paragraph.

Step 2. Count the words in your piece of writing or, if you are using samples, take each sample and count each word in it up to 100. Count contractions and hyphenated words as one word. Count as words numbers or letters separated by space.

Step 3. Count the syllables in your 100-word samples or, if you are testing a whole piece of writing, compute the number of syllables per 100 words. If in doubt about syllabication rules, use any good dictionary. Count the number of syllables in symbols and figures according to the way they are normally read aloud, e.g. two for \$ ("dollars") and four for 1918 ("nineteen-eighteen"). If a passage contains several or lengthy figures, your estimate will be more accurate if you don't include these figures in your syllable count. In a 100-word sample, be sure to add instead a corresponding number of words in your syllable count. To save time, count all syllables

except the first in all words of more than one syllable and add the total to the number of words tested. It is also helpful to "read silently aloud" while counting.

Step 4. Figure the average sentence length in words for your piece of writing or, if you are using samples, for all your samples combined. In a 100-word sample, find the sentence that ends nearest to the 100-word mark—that might be at the 94th word or the 109th word. Count the sentences up to that point and divide the number of words in those sentences by the number of sentences. In counting sentences, follow the units of thought rather than the punctuation: usually sentences are marked off by periods; but sometimes they are marked off by colons or semicolons—like these. But don't break up sentences that are joined by conjunctions like *and* or *but*.

Step 5. Figure the number of "personal words" per 100 words in your piece of writing or, if you are using samples, in all your samples combined. "Personal words" are: (a) All first-, second-, and third-person pronouns except the neuter pronouns *it*, *its*, *itself*, and *they*, *them*, *their*, *theirs*, *themselves* if referring to things rather than people, (b) All words that have masculine or feminine natural gender, e.g. *Jones*, *Mary*, *father*, *sister*, *iceman*, *actress*. Do not count common-gender words like *teacher*, *doctor*, *employee*, *assistant*, *spouse*. Count singular and plural forms, (c) The group words *people* (with the plural verb) and *folks*.

Step 6. Figure the number of "personal sentences" per 100 sentences in your piece of writing or, if you use samples, in all your samples combined. "Personal sentences" are: (a) Spoken sentences, marked by quotation marks or otherwise, often including so-called speech tags like "he said" (e.g. "I doubt it."—We told him: "You can take it or leave it."—"That's all very well," he replied, showing clearly that he didn't believe a word of what we said). (b) Questions, commands, requests, and other sentences directly addressed to the reader. (c) Exclamations. (d) Grammatically incomplete sentences whose full meaning has to be inferred from the context (e.g. Doesn't know a word of English.—Handsome, though.—Well, he wasn't.—The minute you walked out). If a sentence fits two or more of these definitions, count it only once. Divide the number of these "personal sentences" by the total number of sentences you found in Step 4.

Step 7. Find your "reading ease" score by inserting the number of syllab-

bles per 100 words (word length, *wl*) and the average sentence length (*sl*) in the following formula:

$$\text{R.E. ("reading ease")} = 206.835 - .846 \text{ } wl - 1.015 \text{ } sl.$$

The "reading ease" score will put your piece of writing on a scale between 0 (practically unreadable) and 100 (easy for any literate person).

Step 8. Find your "human interest" score by inserting the percentage of "personal words" (*pw*) and the percentage of "personal sentences" (*ps*) in the following formula:

$$\text{H.L. ("human interest")} = 3.635 \text{ } pw + .314 \text{ } ps.$$

The "human interest" score will put your piece of writing on a scale between 0 (no human interest) and 100 (full of human interest). In applying the formulas, remember that Formula A measures *length* (the longer the words and sentences, the harder to read) and Formula B measures *percentages* (the more personal words and sentences, the more human interest).

Roughly, "reading ease" scores will tend to follow the pattern shown in Table 5.

"Human interest" scores will follow the general pattern shown in Table 6.

Table 5
Pattern of "Reading Ease" Scores

"Reading Ease" Score	Description of Style	Typical Magazine	Syllables per 100 words	Average Sentence Length in Words
0 to 20	Very Difficult	Scientific	192 or more	29 or more
30 to 50	Difficult	Academic	167	25
50 to 60	Fairly difficult	Quality	155	21
60 to 70	Standard	Digests	147	17
70 to 80	Fairly easy	Slick-fiction	139	14
80 to 90	Easy	Pulp-fiction	131	11
90 to 100	Very easy	Comics	123 or less	8 or less

Table 6
Pattern of "Reading Ease" Scores

"Human Interest" Score	Description of Style	Typical Magazine	Percentage of Personal Words	Percentage of Personal Sentences
0 to 10	Dull	Scientific	2 or less	0
10 to 20	Mildly Interesting	Trade	4	5
20 to 40	Interesting	Digests	7	15
40 to 60	Highly Interesting	<i>New Yorker</i>	11	22
60 to 100	Dramatic	Fiction	17 or more	58 or more

Sample Application

As an example of the application of the new formulas, two recent descriptions of the "nerve-block" method of anesthesia will be used.

By an odd coincidence, these two variations upon a theme appeared within the same week in *Life* (October 27, 1947) and *The New Yorker* (October 25, 1947). The *Life* story served as text accompanying a series of pictures; it is straight reporting, not particularly simple, and lacks human interest (which was supplied by the pictures). The *New Yorker* passage is part of a personality profile, vivid, dramatic, using all the tricks of the trade to get the reader interested and keep him in suspense.

From *Life*:

Except in the field of surgery, control of pain is still very much in the primitive stages. Countless thousands of patients suffer the tortures of cancer, angina pectoris and other distressing diseases while their physicians are helpless to relieve them. A big step toward help for these sufferers is now being made with a treatment known as nerve-blocking. This treatment, which consists of putting a "block" between the source of pain and the brain, is not a new therapy. But its potentialities are just now being realized., Using better drugs and a wider knowledge of the mechanics of pain gained during and since the war, Doctors E. A. Rovenstine and E. M. Papper of the New York University College of Medicine have been able to help two-thirds of the patients accepted for treatment in their "pain clinic" at Bellevue Hospital.

The nerve-block treatment is comparatively simple and does not have serious aftereffects. It merely involves the injection of an anesthetic drug along the path of the nerve carrying pain impulses from the diseased or injured tissue to the brain. Although its action is similar to that of spinal anesthesia used in surgery, nerve block generally lasts much longer and is only occasionally used for operations. The N. Y. U. doctors have found it effective in a wide range of diseases, including angina pec-

toris, sciatica, shingles, neuralgia and some forms of cancer. Relief is not always permanent, but usually the injection can be repeated. Some angina pectoris patients have had relief for periods ranging from six months to two years. While recognizing that nerve block is no panacea, the doctors feel that results obtained in cases like that of Mike Ostroich (*next page*) will mean a much wider application in the near future.

From *The New Yorker*:

... Recently, [Rovenstine] devoted a few minutes to relieving a free patient in Bellevue of a pain in an arm that had been cut off several years before. The victim of this phantom pain said that the tendons ached and that his fingers were clenched so hard he could feel his nails digging into his palm. Dr. Rovenstine's assistant, Dr. E. M. Papper, reminded Rovenstine that a hundred and fifty years ago the cure would have been to dig up the man's arm, if its burial place was known, and straighten out the hand. Rovenstine smiled. "I tell you," he said. "*We'll* use a two-percent solution of procaine, and if it works, in a couple of weeks we'll go on with an alcohol solution. Procaine, you know, lasts a couple of weeks, alcohol six months or longer. In most cases of this sort, I use the nerve block originated by Labat around 1910 and improved on in New Orleans about ten years back, plus one or two improvisations of my own." (Nerve blocking is a method of anesthetizing a nerve that is transmitting pain.) ...

The man with the pain in the nonexistent hand was an indigent, and Rovenstine was working before a large gallery of student anesthetists and visitors when he exorcised the ghosts that were paining him. Some of the spectators, though they felt awed, also felt inclined to giggle. Even trained anesthetists sometimes get into this state during nerve-block demonstrations because of the tenseness such feats of magic induce in them. The patient, thin, stark-naked, and an obvious product of poverty and cheap gin mills, was nervous and rather apologetic when he was brought into the operating theatre. He lay face down on the operating table. Rovenstine has an easy manner with patients, and as his thick, stubby hands roamed over the man's back, he gently asked, "How you doing?" "My hand, it is all closed together, Doc," the man answered, startled and evidently a little proud of the attention he was getting. "You'll be O.K. soon," Rovenstine said, and turned to the audience. "One of my greatest contributions to medical science has been the use of the eyebrow pencil," he said. He took one from the pocket of his white smock and made a series of marks on the patient's back, near the shoulder of the amputated arm, so that the spectators could see exactly where he was going to work. With a syringe and needle, he raised four small weals on the man's back and then shoved long needles into the weals. The man shuddered but said he felt no pain. Rovenstine then attached a syringe to the first needle, injected the procaine solution, unfastened the syringe, attached it to the next needle, injected more of the solution, and so on. The patient's face began to relax a little. "Lord, Doc," he said. "My hand is loosening up a bit already." "You'll be all right by tonight, I think," Rovenstine said. He was.

A comparative analysis of these two passages is shown in Table 7. The

two passages furnish a good illustration of the stylistic features measured and emphasized by the two new formulas.

Table 7

Comparative Analysis of Treatment of the Same Theme in *Life* and *The New Yorker*

	<i>Life</i> (290 words)	<i>New Yorker</i> (495 words)
Old Formula:		
Average sentence length in words	23	18
Affixes per 100 words	48	35
Personal words per 100 words	2	11
Readability score	5.16	3.20
New Formula A:		
Average sentence length in words	22	18
Syllables per 100 words	165	145
“Reading ease” score	46	66
New Formula B:		
Personal words per 100 words	2	11
Personal sentences per 100 sentences	0	41
“Human interest” score	7	53

Received November 3, 1947.

References

1. Alden, J. Lots of names—short sentences—simple words. *Printer's Ink*, June 29, 1945, 21-22.
2. Bentley, Phyllis. *Some observations on the art of the narrative*. New York: Macmillan, 1947.
3. Cowing, Amy G. They speak his language. *J. Home Econ.*, 1945, 37, 487-489.
4. Fihe, Pauline J., Wallace, Viola, and Schulz, Martha, compilers. *Books for adult beginners; grades I to VII*. Rev. ed. Chicago: American Library Association, 1946.
5. Flesch, R. *Marks of readable style; a study in adult education*. New York: Bur. of Publ., Teachers Coll., Columbia Univ., 1943. (Contr. to Educ. No. 897.)
6. Flesch, R. *The art of plain talk*. New York: Harper & Brothers, 1946.

7. Flesch, R. How to write copy that will be read. *Advertising & Selling*, March, 1947. 113ff.
8. Gray, W. S., and Leary, Bernice E. *What makes a book readable*. Chicago: Univ. of Chicago Press, 1935.
9. Gunning, R. Gunning finds papers too hard to read. *Editor & Publisher*, May 19, 1945, 12.
10. Lorge, I. Predicting reading difficulty of selections for children. *Elem. English Rev.*, 1939, 16, 229-233.
11. McCall, W. A., and Crabbs, Lelah M. *Standard test lessons in reading*. Books II, III, IV, and V. New York: Bur. of Publ., Teachers Coll., Columbia Univ., 1926.
12. Miller, L. R. Reading grade placement of the first 23 books awarded the John Newbery prize. *Elem. Sch. J.*, 1946, 394-399.
13. Murphy, D. R. Test proves short sentences and words get best readership. *Printer's Ink*, 1947, 218, 61-64.
14. Murphy, D. R. How plain talk increases readership 45% to 66%. *Printer's Ink*, 1947, 220, 35-37.
15. Sanford, F. H. *Individual differences in the mode of verbal expression*. Unpublished Ph.D. thesis. Harvard Univ., 1941.
16. Sherbow, B. *Making type work*. New York: Century, 1916.
17. Stevens, S. S., and Stone, Geraldine. Psychological writing, easy and hard. *Amer. Psychologist*, 1947, 2, 230-235. Discussion, 1947, 2, 523-525.
18. Foreign news written over heads of readers. *Editor & Publisher*, Dec. 28, 1946, 28.
19. *How does your writing read?* U. S. Civil Service Commission. Washington: U. S. Government Printing Office, 1946.
20. *Readability in news writing; report on an experiment by United Press*. New York: United Press Associations, 1945.