The Classic Readability Studies

William H. DuBay, Editor

Impact Information
Costa Mesa
Copyright


Introductions © 2006 William H. DuBay. All Rights Reserved.

Readers' Comments

Please send all comments and suggestions regarding this document to:

William DuBay
Impact Information
126 E. 18th Street, #C204
Costa Mesa, CA 92627
Phone: (949) 631-3309
Email: bdubay@impact-information.com
Website: http://www.impact-information.com
Contents

1893—L.A. Sherman: The Analytics of Literature ................................................................. 1
1921—E. L. Thorndike: The Teachers’ Book of Words ...................................................... 4
1923—The Lively and Pressey Measuring Method ............................................................... 6
    Introduction....................................................................................................................... 6
    A Method for Measuring the "Vocabulary Burden" of Textbooks .................................... 7
1928—Vogel and Washburne: The Winnetka Formula ....................................................... 16
    Introduction...................................................................................................................... 16
    An Objective Method of Determining Grade Placement of Children's Reading Material .... 18
The Early 1930s—New Directions for Readability ............................................................. 27
    Waples and Tyler: What Adults Want to Read About ...................................................... 27
    Ralph Ojemann: The Difficulty of Adult Materials ....................................................... 27
    Dale and Tyler: Adults of Limited Reading Ability ....................................................... 28
    Lyman Bryson: Books for the Average Reader .............................................................. 28
1931—Patty and Painter: The Vocabulary Burden ............................................................... 31
    Introduction...................................................................................................................... 31
    A Technique for Measuring the Vocabulary Burden of Textbooks ................................. 32
1944—The Lorge Readability Index ..................................................................................... 44
    Introduction...................................................................................................................... 44
    Predicting Readability .................................................................................................... 46
1948—The Dale-Chall Readability Formula ........................................................................ 61
    Introduction...................................................................................................................... 61
    A Formula for Predicting Readability ............................................................................ 63
    A Formula for Predicting Readability: Instructions ....................................................... 75
1948—The Flesch Formulas ............................................................................................... 96
    Introduction...................................................................................................................... 96
    A New Readability Yardstick .......................................................................................... 99
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.

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 re-
In Analytics of Literature, L.A. Sherman showed the importance of average sentence length and the relationship between spoken and written English.

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
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 *Teacher’s Word Book*, which came out in 1921, was the first extensive listing of words in English by frequency. 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. It was also the basis for the first readability formulas.

Its author, psychologist Edward L. Thorndike of 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. As we learn and grow, our vocabulary grows as does our ability to master longer and more complex sentences. How much that continues to grow depends on how much reading is done throughout life.

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. 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 writ-
ten material. The first 300 words make up about 65 percent of it.

Educators, publishers, and teachers still use word-frequency lists to evaluate reading materials for schools. After Thorndike, there was extensive research on vocabulary. The high mark came in 1949 with Human Behavior and The Principle of Least Effort by Harvard’s George Kingsley Zipf.

Zipf used a statistical analysis of language to show how the principle of least effort works in human speech. Zipf showed that, in many languages, there is a mathematical relationship between the hard and easy words, now called Zipf’s curve. This notion of saving energy is a central feature of language and is one of the principle bases of research on the frequency of words.

In 1968, psychologist George Klare wrote, “Not only do humans tend to used 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.”

—WHD
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.*

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.

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

---

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

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

---

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.
bus 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 samp-

<table>
<thead>
<tr>
<th>SUMMARY REGARDING VOCABULARY BURDEN—16 TYPES OF MATERIAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counts</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Second readers: Jones...............................</td>
</tr>
<tr>
<td>Horace Mann.................................</td>
</tr>
<tr>
<td>Aldine...........................................</td>
</tr>
<tr>
<td>Fourth readers: Jones............................</td>
</tr>
<tr>
<td>Aldine.........................................</td>
</tr>
<tr>
<td>Horace Mann.................................</td>
</tr>
<tr>
<td>Stevenson: Kidnapped.............................</td>
</tr>
<tr>
<td>Thackeray: Vanity Fair..........................</td>
</tr>
<tr>
<td>Columbus Dispatch................................</td>
</tr>
<tr>
<td>without local names............................</td>
</tr>
<tr>
<td>Muzzy: American History..........................</td>
</tr>
<tr>
<td>Clark: Introduction to Science...............</td>
</tr>
<tr>
<td>Clark: General Science..........................</td>
</tr>
<tr>
<td>Hunter: Elements of Biology....................</td>
</tr>
<tr>
<td>Elementary Chemistry...........................</td>
</tr>
<tr>
<td>Kimball: College Physics.......................</td>
</tr>
<tr>
<td>Howell: Physiology..............................</td>
</tr>
</tbody>
</table>

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.

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 apparently 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 trams 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.

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 important, 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. 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. The interpretation of median index numbers, and of other features of the total distributions of thousand word counts, would be much easier,

---

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.
and the significance of all results much enriched.\footnote{Or suppose Thorndike had used for index numbers a scale of ten, for the ten thousand words, using decimals to distinguish the position of a word within each thousand. It would then be possible to read directly, from the index number, the place of each word in the ten thousand, and interpretation would be greatly facilitated.} However, these are matters of relatively minor detail. The fundamental value of Thorndike’s contribution is obvious; the "Word Book" has opened up a whole new field for investigation.
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 overlapping 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 overlapping, 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.
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 corresponding 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 er-

ror 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% ($0.845^2$) of the variance of the text difficulty.

—WHD
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.  

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

---

The Classic Readability Studies

1928—Vogel and Washburne: The Winnetka Formula

graph-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 technique4)
   a) Number of different words occurring in a sampling of 1,000 words
   b) Median index number (based on Thorndike's indexed word list5) 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
   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 intran-

---

sitive, 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 makeup
   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 I shows the correlation between each of the elements selected 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.

---

6 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.

7 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.
### Table I

**CORRELATION OF VARIOUS ELEMENTS WITH MEDIAN READING SCORE**

<table>
<thead>
<tr>
<th>Element</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of different words occurring in a sampling of 1,000 words</td>
<td>.770</td>
</tr>
<tr>
<td>Median index number (based on Thorndike's list) of 1,000-word sampling</td>
<td>— .704</td>
</tr>
<tr>
<td>Number of words in 1,000-word sampling not occurring in Thorndike's list</td>
<td>.674</td>
</tr>
<tr>
<td>Number of words in book</td>
<td>— .704</td>
</tr>
<tr>
<td>Number of phrases in 1,000-word sampling</td>
<td>.592</td>
</tr>
<tr>
<td>Number of verbs in 1,000-word sampling</td>
<td>— .527</td>
</tr>
<tr>
<td>Number of words per paragraph</td>
<td>— .527</td>
</tr>
<tr>
<td>Number of prepositions in 1,000-word sampling</td>
<td>.518</td>
</tr>
<tr>
<td>Number of phrases of all kinds in 75 sample sentences</td>
<td>.518</td>
</tr>
<tr>
<td>Number of phrases and clauses of all kinds in 75 sample sentences</td>
<td>.474</td>
</tr>
<tr>
<td>Number of adverbial phrases and clauses in 75 sample sentences</td>
<td>.467</td>
</tr>
<tr>
<td>Number of adverbial phrases and clauses in 1,000-word sampling</td>
<td>.463</td>
</tr>
<tr>
<td>Number of adjective phrases and clauses in 75 sample sentences</td>
<td>.461</td>
</tr>
<tr>
<td>Number of adverbial phrases in 75 sample sentences</td>
<td>.458</td>
</tr>
<tr>
<td>Number of words in 75 sample sentences</td>
<td>.453</td>
</tr>
<tr>
<td>Number of simple sentences in 75 sample sentences</td>
<td>— .371</td>
</tr>
<tr>
<td>Number of conjunctions in 1,000-word sampling</td>
<td>.296</td>
</tr>
<tr>
<td>Number of adverbial clauses in 75 sample sentences</td>
<td>.291</td>
</tr>
<tr>
<td>Number of nouns in 1,000-word sampling</td>
<td>— .262</td>
</tr>
</tbody>
</table>

*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. 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 Classic Readability Studies

TABLE II
INTERCORRELATION OF VARIOUS ELEMENTS

<table>
<thead>
<tr>
<th></th>
<th>Different Words in 1,000</th>
<th>Prepositions in 1,000 Words</th>
<th>Verbs in 1,000 Words</th>
<th>Words per Paragraph</th>
<th>Words in 75 Sentences</th>
<th>Simple Sentences in 75</th>
<th>Uncommon Words</th>
<th>Adverbial Clause in 75 Sentences</th>
<th>Nouns in 1,000 Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median reading score</td>
<td>.770</td>
<td>.518</td>
<td>.527</td>
<td>.518</td>
<td>.453</td>
<td>-0.371</td>
<td>.674</td>
<td>.291</td>
<td>-0.262</td>
</tr>
<tr>
<td>Different words in 1,000</td>
<td>.546</td>
<td>-0.572</td>
<td>.516</td>
<td>.442</td>
<td>-0.306</td>
<td>.692</td>
<td>.308</td>
<td>-0.177</td>
<td></td>
</tr>
<tr>
<td>Prepositions in 1,000 words</td>
<td>-0.777</td>
<td>.462</td>
<td>.398</td>
<td>-0.134</td>
<td>.412</td>
<td>.131</td>
<td>.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbs in 1,000 words</td>
<td>-0.517</td>
<td>-0.543</td>
<td>.285</td>
<td>-0.431</td>
<td>-0.192</td>
<td>.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Words per paragraph</td>
<td>.706</td>
<td>-0.503</td>
<td>.322</td>
<td>.565</td>
<td>-0.356</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Words in 75 sentences</td>
<td>-0.741</td>
<td>.244</td>
<td>.818</td>
<td>-0.399</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple sentences in 75</td>
<td>.106</td>
<td>.674</td>
<td>.552</td>
<td>.074</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.069</td>
</tr>
<tr>
<td>Uncommon words in 1,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverbial clauses in 75 sentences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.424</td>
</tr>
</tbody>
</table>

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**

<table>
<thead>
<tr>
<th>Score</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>4–16</td>
<td>II</td>
</tr>
<tr>
<td>18–34</td>
<td>III</td>
</tr>
<tr>
<td>36–52</td>
<td>IV</td>
</tr>
<tr>
<td>54–62</td>
<td>V</td>
</tr>
<tr>
<td>64–70</td>
<td>VI</td>
</tr>
<tr>
<td>72–78</td>
<td>VII</td>
</tr>
<tr>
<td>80–86</td>
<td>VIII</td>
</tr>
<tr>
<td>88–94</td>
<td>IX</td>
</tr>
<tr>
<td>96–102</td>
<td>X</td>
</tr>
<tr>
<td>104–112</td>
<td>XI</td>
</tr>
</tbody>
</table>
Let us take *The Japanese Empire*⁸ by Harry A. Franck as an example of the application of the equation.

Number of different words in 1,000 = 445  
Number of prepositions in 1,000 words = 117  
Number of uncommon words in 1,000 = 22  
Number of simple sentences in 75 = 20

\[ X_1 = (0.085)(445) + (0.101)(117) + (0.604)(22) - (0.411)(20) + 17.43 \]

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 children 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.

---

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.

<table>
<thead>
<tr>
<th>Grades</th>
<th>III</th>
<th>IVs</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of different words in 1,000:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper quartile</td>
<td>338</td>
<td>401</td>
<td>417</td>
<td>435.5</td>
<td>457.5</td>
<td>460.8</td>
</tr>
<tr>
<td>Median</td>
<td>316</td>
<td>377</td>
<td>407.5</td>
<td>416.5</td>
<td>440.5</td>
<td>458</td>
</tr>
<tr>
<td>Lower quartile</td>
<td>258</td>
<td>329</td>
<td>386</td>
<td>397.5</td>
<td>411</td>
<td>447</td>
</tr>
<tr>
<td>Number of prepositions in 1,000 words:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper quartile</td>
<td>79</td>
<td>106</td>
<td>110</td>
<td>114</td>
<td>116</td>
<td>123</td>
</tr>
<tr>
<td>Median</td>
<td>71.3</td>
<td>96</td>
<td>100</td>
<td>107</td>
<td>99</td>
<td>115.5</td>
</tr>
<tr>
<td>Lower quartile</td>
<td>63</td>
<td>79</td>
<td>83</td>
<td>100</td>
<td>93</td>
<td>101</td>
</tr>
<tr>
<td>Number of uncommon words in 1,000:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper quartile</td>
<td>8</td>
<td>14</td>
<td>20</td>
<td>24.5</td>
<td>34.5</td>
<td>40</td>
</tr>
<tr>
<td>Median</td>
<td>6</td>
<td>11</td>
<td>14.5</td>
<td>19.5</td>
<td>26</td>
<td>32.5</td>
</tr>
<tr>
<td>Lower quartile</td>
<td>2</td>
<td>6</td>
<td>12</td>
<td>17</td>
<td>18.5</td>
<td>28.5</td>
</tr>
<tr>
<td>Number of simple sentences in 75:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper quartile</td>
<td>49</td>
<td>42.5</td>
<td>26</td>
<td>25</td>
<td>29</td>
<td>34</td>
</tr>
<tr>
<td>Median</td>
<td>39</td>
<td>30</td>
<td>21.5</td>
<td>19</td>
<td>22.5</td>
<td>26</td>
</tr>
<tr>
<td>Lower quartile</td>
<td>36</td>
<td>22</td>
<td>18</td>
<td>11</td>
<td>18</td>
<td>21.5</td>
</tr>
</tbody>
</table>
The Early 1930s—New Directions for Readability

Waples and Tyler: What Adults Want to Read About

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.

Ralph Ojemann: The Difficulty of Adult Materials

The year 1934 marked the beginning of more rigorous standards for the formulas. Ralph Ojemann did not invent a formula, but in 1934, he did invent a method of assessing the difficulty of materials for adult parent-education materials. 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 scaling (See “Text leveling” below). Although he was not able to express the qualitative variables in numeric terms, he succeeded in proving they could not be ignored.
Dale and Tyler: Adults of Limited Reading Ability

After working with Waples, Ralph Tyler became interested in adults of limited reading ability. He joined with Edgar Dale to publish in 1934 their own readability formula and the first study on adult readability formulas. 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.

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.

The Ojemann and Dale-Tyler studies mark the beginning of work on adult formulas that would continue unabated until the present time.

Lyman Bryson: Books for the Average Reader

During the depression of the 1930’s, the government in the U.S. put enormous resources into adult education. Bryson Lyman first became interested in non-fiction materials written for the average adult reader while serving as a leader in adult-education meetings in New York City. What he found was that what kept people from reading more was not lack of intelligence, but the lack of reading skills, a direct result of limited schooling.

He also found out there is a tendency to judge adults by the education their children receive and to assume the great bulk
of people have been through high school. At that time, 40 to 50 million people had a 7th to 9th grade education and reading ability.

Writers had assumed that readers had an equal education to their own or at least an equal reading ability. Highly educated people failed to realize just how much easier it is for them to read than it is for an average person. They found it difficult to recognize difficult writing because they read so well themselves.

Although college and business courses had long promoted ideas expressed in a direct and lucid style, Bryson found that simple and clear language was rare. He said such language results from “a discipline and artistry which few people who have ideas will take the trouble to achieve… If simple writing were easy, many of our problems would have been solved long ago” (Klare and Buck, p. 58).

Bryson helped set up the Readability Laboratory of the Columbia University Teachers College with Charles Beard and M. A. Cartwright. Bryson understood that people with enough motivation and time could read difficult material and improve their reading ability. Experience, however, showed him that most people do not do that.

Perhaps Bryson’s greatest contribution was the influence he had on his two students, Irving Lorge and Rudolf Flesch.

—WHD
1931—Patty and Painter: The Vocabulary Burden

Introduction

In 1931, W. W. Patty and W. I. Painter 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
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 a 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 Pressey 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* 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 of texts or of the vocabulary burden of units equal in length fails, however, to consider the extra burden imposed on the reader by additional length.

---

3 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.
Dolch\(^4\) 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.

Ward\(^5\) 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.

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 approxi-


mately 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**

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

**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, \( W.T. \) 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. ÷ 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 number 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:

<table>
<thead>
<tr>
<th>Text</th>
<th>T.W.S</th>
<th>R.</th>
<th>T.W.V.</th>
<th>A.W.W.V.</th>
<th>I.N.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book I</td>
<td>646</td>
<td>379</td>
<td>77.045</td>
<td>119.26</td>
<td>.315</td>
</tr>
<tr>
<td>Book II</td>
<td>1051</td>
<td>433</td>
<td>125.289</td>
<td>119.20</td>
<td>.275</td>
</tr>
</tbody>
</table>
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:

\[
A.W.W.V. = \frac{T.W.V.}{T.W.S.} \\
I.N. = \frac{(T.W.V. \div T.W.S.)}{R} \quad \text{or} \\
I.N. = \frac{(T.W.V. \div T.W.S.) \times (1 \div R)}{R} \quad \text{or} \\
I.N. = \frac{T.W.V.}{T.W.S. \times R}
\]

This same process should give satisfactory results in all cases.

**A SUB-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.

SAMPLE OF TEXTBOOK CHECK LIST

<table>
<thead>
<tr>
<th>Year of H.S.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our English</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Written and Spoken</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English Literature</td>
<td>Clippinger</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outlines of English Lit</td>
<td>Long</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outlines of English Lit. readings</td>
<td>Long</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Literature</td>
<td>Long</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outlines of American Lit</td>
<td>Long</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outline of American Lit.</td>
<td>Long</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early American History, Rev</td>
<td>Webster</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern European History, Rev</td>
<td>Webster</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern Times and Living Past</td>
<td>Elson</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of the U.S</td>
<td>Beard, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government in the U.S</td>
<td>Smith, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary Principles of Chem</td>
<td>Brownlee, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Book in Chem., 1928 ed</td>
<td>Bradbury</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemistry and Its Uses</td>
<td>McPherson, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beginners Chem. and Its Uses</td>
<td>Irwin, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elem. Prin. of Physics</td>
<td>Fuller, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elements of Physics</td>
<td>Millikan, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essentials of Mod. Physics</td>
<td>Dull</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physics in Everyday Life</td>
<td>Henderson</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commerce and Industry</td>
<td>Smith</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-School Geography</td>
<td>Whitbeck</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Course in Algebra</td>
<td>Nyberg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Course in Algebra</td>
<td>Nyberg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern Plane Geometry</td>
<td>Clark, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern Solid Geometry</td>
<td>Clark, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arithmetic of Business</td>
<td>Smith</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applied Arithmetic</td>
<td>Smith</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm Projects and Problems</td>
<td>Davis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal Husbandry</td>
<td>Harper</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soils and Crops</td>
<td>Mosier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies in Horticulture</td>
<td>Lloyd</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engineering on the Farm</td>
<td>Stewart</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Practical Botany (Agri)</td>
<td>Bergen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant Life and Plant Uses</td>
<td>Coulter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elem. Studies in Botany</td>
<td>Coulter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Course in Botany</td>
<td>Pool, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Textbook in Botany</td>
<td>Allen, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal Studies</td>
<td>Jordan, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Zoology</td>
<td>Linville, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heathful Living</td>
<td>Williams</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Practical Zoology</td>
<td>Hegner</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

In order that we might more objectively compare the diffi-
culty 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. 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:

\[ \text{B.F.} \times \text{I.N.} = \text{F.I.P.} \]

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

\[ \text{A.I.N.} = \frac{\text{T.F.I.P.}}{\text{N.}} \]

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.


Their criterion 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.

![Diagram showing the four basic elements of reading ease.](image)
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

---

**Table I**

<table>
<thead>
<tr>
<th>Major Category</th>
<th>All Persons</th>
<th>Librarians</th>
<th>Publishers</th>
<th>Others Interested in Adult Education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M.</td>
<td>σ</td>
<td>Range</td>
<td>M.</td>
</tr>
<tr>
<td>I. Format</td>
<td>20.26</td>
<td>7.68</td>
<td>45-5</td>
<td>24.13</td>
</tr>
<tr>
<td>II. General Features of Organization</td>
<td>15.58</td>
<td>7.04</td>
<td>50-3</td>
<td>15.71</td>
</tr>
<tr>
<td>III. Style of Expression and Presentation</td>
<td>30.71</td>
<td>9.17</td>
<td>59-0</td>
<td>12.74</td>
</tr>
<tr>
<td>IV. Content</td>
<td>33.64</td>
<td>13.14</td>
<td>75-7</td>
<td>27.42</td>
</tr>
</tbody>
</table>

**Fig. 1.** Opinion concerning the influence of classified factors on readability

The four major categories of readability (Gray and Leary, p. 31).
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 mor-
pheme 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.
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 Lymon Bryson.

In 1938, Irving Lorge 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,” 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, Lorge published corrections to his formula, which are given here in the footnote on p. 56.

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 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 formulae, which were introduced in 1948.

—WHD
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.¹

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 measure 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, Flesch,² 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

¹ Originally published as “Predicting Readability” in the Teachers College Record, Vol. 45, 404-419, March 1944—W.H.D.

² Studies referred to in this article, together with other pertinent references, are listed in the Bibliography, page 59.
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.).

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.³

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.⁴

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 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.⁵

---

³ 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.
⁴ The multiple correlations between the five predictors and the criterion used by Gray and Leary was .644.
⁵ 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 com-
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 56) 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 intercorrelated. 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.)

**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 separate 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 answer.

   (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 list\(^6\) to cross out in the sample every word on the Dale list, regardless of its meaning.\(^7\) The list is given on pages 56 to 59.

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.
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. Frankfurt and Frankfurth 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.

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 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.
1944—The Lorge Readability Index

The Classic Readability Studies

ILLUSTRATION OF LISTING OF HARD WORDS

<table>
<thead>
<tr>
<th></th>
<th>altogether</th>
<th>add</th>
<th>advanced</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>battle</td>
<td>battlefield</td>
<td>birth</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>continent</td>
<td>conceived /</td>
<td>created</td>
<td>civil</td>
</tr>
<tr>
<td>C</td>
<td>dedicated // ///</td>
<td>detract</td>
<td>devotion /</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>equal</td>
<td>engaged</td>
<td>endure</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>forth</td>
<td>final</td>
<td>fought</td>
<td>freedom</td>
</tr>
<tr>
<td>F</td>
<td>government</td>
<td>hallowed</td>
<td>honored</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>increased</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>Liberty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>nation ///</td>
<td>nobly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>proposition</td>
<td>portion</td>
<td>proper</td>
<td>power</td>
</tr>
<tr>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>remaining</td>
<td>resolve</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>score</td>
<td>sense</td>
<td>struggled</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>testing</td>
<td>thus</td>
<td>task</td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>unfinished</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>vain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WXYZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After the number of sentences has been counted, the work sheet may be completed as shown on page 56.

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.
FORMULA FOR ESTIMATING GRADE PLACEMENT OF READING MATERIAL

WORK SHEET

Title of article: Gettysburg Address
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

COMPUTATION

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.

THE DALE LIST OF 769 EASY WORDS

A
a already
about
above
across
act
afraid
after
afternoon
again
against
ago
air
all
almost
alone

along around bear best bone
as ask at am away American an B been bit box
before black boy
began bless branch
begin blind brave
behind blood bread
being blow break
belong boat bright
beside body bring

28 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
<table>
<thead>
<tr>
<th>broken</th>
<th>coal</th>
<th>drink</th>
<th>felt</th>
<th>golden</th>
<th>how</th>
</tr>
</thead>
<tbody>
<tr>
<td>brother</td>
<td>coat</td>
<td>drive</td>
<td>fence</td>
<td>gone</td>
<td>hundred</td>
</tr>
<tr>
<td>brought</td>
<td>cold</td>
<td>drop</td>
<td>few</td>
<td>good</td>
<td>hunt</td>
</tr>
<tr>
<td>brown</td>
<td>color</td>
<td>dry</td>
<td>field</td>
<td>got</td>
<td>hurry</td>
</tr>
<tr>
<td>building</td>
<td>coming</td>
<td>fill</td>
<td>grass</td>
<td></td>
<td></td>
</tr>
<tr>
<td>built</td>
<td>company</td>
<td>E</td>
<td>find</td>
<td>gray</td>
<td>I</td>
</tr>
<tr>
<td>burn</td>
<td>cook</td>
<td>each</td>
<td>fine</td>
<td>great</td>
<td>I</td>
</tr>
<tr>
<td>busy</td>
<td>cool</td>
<td>ear</td>
<td>finger</td>
<td>green</td>
<td>ice</td>
</tr>
<tr>
<td>but</td>
<td>corn</td>
<td>early</td>
<td>finish</td>
<td>grew</td>
<td>if</td>
</tr>
<tr>
<td>butter</td>
<td>corner</td>
<td>earth</td>
<td>fire</td>
<td>ground</td>
<td>in</td>
</tr>
<tr>
<td>buy</td>
<td>cost</td>
<td>east</td>
<td>first</td>
<td>grow</td>
<td>Indian</td>
</tr>
<tr>
<td>by</td>
<td>could</td>
<td>easy</td>
<td>fish</td>
<td>guess</td>
<td>instead</td>
</tr>
<tr>
<td>count</td>
<td>count</td>
<td>eat</td>
<td>fit</td>
<td>into</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>country</td>
<td>edge</td>
<td>five</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>iron</td>
</tr>
<tr>
<td>cake</td>
<td>course</td>
<td>egg</td>
<td>fix</td>
<td>had</td>
<td>is</td>
</tr>
<tr>
<td>call</td>
<td>cover</td>
<td>eight</td>
<td>floor</td>
<td>hair</td>
<td>it</td>
</tr>
<tr>
<td>came</td>
<td>cow</td>
<td>either</td>
<td>flower</td>
<td>half</td>
<td>its</td>
</tr>
<tr>
<td>can</td>
<td>cried</td>
<td>else</td>
<td>fly</td>
<td>hall</td>
<td></td>
</tr>
<tr>
<td>cap</td>
<td>cross</td>
<td>end</td>
<td>follow</td>
<td>hand</td>
<td>J</td>
</tr>
<tr>
<td>captain</td>
<td>crowd</td>
<td>England</td>
<td>food</td>
<td>hang</td>
<td>jump</td>
</tr>
<tr>
<td>car</td>
<td>crown</td>
<td>English</td>
<td>foot</td>
<td>happy</td>
<td>just</td>
</tr>
<tr>
<td>care</td>
<td>cry</td>
<td></td>
<td>for</td>
<td>hard</td>
<td>K</td>
</tr>
<tr>
<td>careful</td>
<td>cap</td>
<td>even</td>
<td>forget</td>
<td>has</td>
<td>keep</td>
</tr>
<tr>
<td>carry</td>
<td>cut</td>
<td>evening</td>
<td>fourth</td>
<td>hat</td>
<td>kept</td>
</tr>
<tr>
<td>case</td>
<td>ever</td>
<td>found</td>
<td>have</td>
<td>kill</td>
<td></td>
</tr>
<tr>
<td>catch</td>
<td>D</td>
<td>every</td>
<td>four</td>
<td>he</td>
<td>kind</td>
</tr>
<tr>
<td>cause</td>
<td>dance</td>
<td>everything</td>
<td>fresh</td>
<td>head</td>
<td>king</td>
</tr>
<tr>
<td>center</td>
<td>day</td>
<td>expect</td>
<td>from</td>
<td>heard</td>
<td>knee</td>
</tr>
<tr>
<td>chair</td>
<td>dead</td>
<td>eye</td>
<td>front</td>
<td>heart</td>
<td>knew</td>
</tr>
<tr>
<td>chance</td>
<td>dear</td>
<td>fruit</td>
<td>heavy</td>
<td>know</td>
<td></td>
</tr>
<tr>
<td>change</td>
<td>deep</td>
<td>F</td>
<td>full</td>
<td>help</td>
<td></td>
</tr>
<tr>
<td>chief</td>
<td>did</td>
<td>face</td>
<td>her</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>child</td>
<td>die</td>
<td>fair</td>
<td>G</td>
<td>here</td>
<td>lady</td>
</tr>
<tr>
<td>children</td>
<td>different</td>
<td>fall</td>
<td>game</td>
<td>herself</td>
<td>laid</td>
</tr>
<tr>
<td>choose</td>
<td>dinner</td>
<td>family</td>
<td>garden</td>
<td>hide</td>
<td>lake</td>
</tr>
<tr>
<td>Christmas</td>
<td>do</td>
<td>fancy</td>
<td>gate</td>
<td>high</td>
<td>land</td>
</tr>
<tr>
<td>church</td>
<td>doctor</td>
<td>far</td>
<td>gave</td>
<td>hill</td>
<td>large</td>
</tr>
<tr>
<td>circle</td>
<td>does</td>
<td>farm</td>
<td>get</td>
<td>him</td>
<td>last</td>
</tr>
<tr>
<td>city</td>
<td>dog</td>
<td>farmer</td>
<td>gift</td>
<td>himself</td>
<td>late</td>
</tr>
<tr>
<td>class</td>
<td>done</td>
<td>fast</td>
<td>girl</td>
<td>his</td>
<td>laugh</td>
</tr>
<tr>
<td>clean</td>
<td>don’t</td>
<td>fat</td>
<td>give</td>
<td>hold</td>
<td>lay</td>
</tr>
<tr>
<td>clear</td>
<td>door</td>
<td>father</td>
<td>glad</td>
<td>hole</td>
<td>lead</td>
</tr>
<tr>
<td>clock</td>
<td>double</td>
<td>feed</td>
<td>glass</td>
<td>home</td>
<td>learn</td>
</tr>
<tr>
<td>close</td>
<td>down</td>
<td>feel</td>
<td>go</td>
<td>hope</td>
<td>leave</td>
</tr>
<tr>
<td>cloth</td>
<td>draw</td>
<td>feet</td>
<td>God</td>
<td>horse</td>
<td>left</td>
</tr>
<tr>
<td>clothes</td>
<td>dream</td>
<td>fell</td>
<td>going</td>
<td>hot</td>
<td>leg</td>
</tr>
<tr>
<td>cloud</td>
<td>dress</td>
<td>fellow</td>
<td>gold</td>
<td>house</td>
<td>lesson</td>
</tr>
<tr>
<td>let</td>
<td>month</td>
<td>off</td>
<td>queen</td>
<td>seed</td>
<td>something</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>------</td>
<td>-----------</td>
</tr>
<tr>
<td>letter</td>
<td>moon</td>
<td>office</td>
<td>quick</td>
<td>seem</td>
<td>sometime</td>
</tr>
<tr>
<td>lie</td>
<td>more</td>
<td>often</td>
<td>quiet</td>
<td>seen</td>
<td>song</td>
</tr>
<tr>
<td>lift</td>
<td>morning</td>
<td>old</td>
<td>quite</td>
<td>self</td>
<td>soon</td>
</tr>
<tr>
<td>light</td>
<td>most</td>
<td>on</td>
<td>sell</td>
<td>sound</td>
<td></td>
</tr>
<tr>
<td>like</td>
<td>mother</td>
<td>once</td>
<td>R</td>
<td>send</td>
<td>south</td>
</tr>
<tr>
<td>line</td>
<td>mountain</td>
<td>one</td>
<td>race</td>
<td>sent</td>
<td>space</td>
</tr>
<tr>
<td>lion</td>
<td>mouth</td>
<td>only</td>
<td>rain</td>
<td>serve</td>
<td>speak</td>
</tr>
<tr>
<td>lips</td>
<td>move</td>
<td>open</td>
<td>ran</td>
<td>set</td>
<td>spot</td>
</tr>
<tr>
<td>listen</td>
<td>Mr.</td>
<td>or</td>
<td>seven</td>
<td>spread</td>
<td></td>
</tr>
<tr>
<td>little</td>
<td>Mrs.</td>
<td>other</td>
<td>reach</td>
<td>several</td>
<td>spring</td>
</tr>
<tr>
<td>live</td>
<td>much</td>
<td>out</td>
<td>read</td>
<td>shake</td>
<td>square</td>
</tr>
<tr>
<td>lead</td>
<td>music</td>
<td>out</td>
<td>ready</td>
<td>shall</td>
<td>stand</td>
</tr>
<tr>
<td>long</td>
<td>must</td>
<td>outside</td>
<td>real</td>
<td>shape</td>
<td>star</td>
</tr>
<tr>
<td>look</td>
<td>my</td>
<td>over</td>
<td>reason</td>
<td>she</td>
<td>start</td>
</tr>
<tr>
<td>lost</td>
<td>myself</td>
<td>own</td>
<td>red</td>
<td>sheep</td>
<td>station</td>
</tr>
<tr>
<td>lot</td>
<td></td>
<td>remember</td>
<td></td>
<td>shine</td>
<td>stay</td>
</tr>
<tr>
<td>loud</td>
<td>N</td>
<td>P</td>
<td>rest</td>
<td>ship</td>
<td>step</td>
</tr>
<tr>
<td>love</td>
<td>name</td>
<td>page</td>
<td>rich</td>
<td>shoe</td>
<td>stick</td>
</tr>
<tr>
<td>low</td>
<td>near</td>
<td>paint</td>
<td>ride</td>
<td>shop</td>
<td>still</td>
</tr>
<tr>
<td>neck</td>
<td>pair</td>
<td>right</td>
<td>short</td>
<td>stone</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>need</td>
<td>paper</td>
<td>ring</td>
<td>should</td>
<td>stood</td>
</tr>
<tr>
<td>made</td>
<td>neighbor</td>
<td>part</td>
<td>river</td>
<td>shoulder</td>
<td>stop</td>
</tr>
<tr>
<td>mail</td>
<td>neither</td>
<td>party</td>
<td>road</td>
<td>show</td>
<td>store</td>
</tr>
<tr>
<td>make</td>
<td>nest</td>
<td>pass</td>
<td>rock</td>
<td>shut</td>
<td>storm</td>
</tr>
<tr>
<td>man</td>
<td>never</td>
<td>path</td>
<td>roll</td>
<td>sick</td>
<td>story</td>
</tr>
<tr>
<td>many</td>
<td>new</td>
<td>pay</td>
<td>roof</td>
<td>side</td>
<td>straight</td>
</tr>
<tr>
<td>march</td>
<td>New York</td>
<td>pen</td>
<td>room</td>
<td>sign</td>
<td>street</td>
</tr>
<tr>
<td>mark</td>
<td>next</td>
<td>people</td>
<td>rose</td>
<td>silk</td>
<td>strike</td>
</tr>
<tr>
<td>market</td>
<td>nice</td>
<td>pick</td>
<td>round</td>
<td>silver</td>
<td>strong</td>
</tr>
<tr>
<td>matter</td>
<td>night</td>
<td>picture</td>
<td>row</td>
<td>sing</td>
<td>such</td>
</tr>
<tr>
<td>may</td>
<td>nine</td>
<td>piece</td>
<td>run</td>
<td>sir</td>
<td>sugar</td>
</tr>
<tr>
<td>me</td>
<td>no</td>
<td>place</td>
<td>sister</td>
<td>suit</td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>noise</td>
<td>plain</td>
<td>S</td>
<td>sit</td>
<td>summer</td>
</tr>
<tr>
<td>measure</td>
<td>none</td>
<td>plan</td>
<td>said</td>
<td>six</td>
<td>sun</td>
</tr>
<tr>
<td>meat</td>
<td>noon</td>
<td>play</td>
<td>sail</td>
<td>size</td>
<td>suppose</td>
</tr>
<tr>
<td>meet</td>
<td>nor</td>
<td>please</td>
<td>salt</td>
<td>skin</td>
<td>sure</td>
</tr>
<tr>
<td>men</td>
<td>north</td>
<td>point</td>
<td>same</td>
<td>sky</td>
<td>surprise</td>
</tr>
<tr>
<td>met</td>
<td>nose</td>
<td>poor</td>
<td>sand</td>
<td>sleep</td>
<td>sweet</td>
</tr>
<tr>
<td>middle</td>
<td>not</td>
<td>post</td>
<td>sat</td>
<td>slow</td>
<td></td>
</tr>
<tr>
<td>might</td>
<td>note</td>
<td>pound</td>
<td>save</td>
<td>small</td>
<td>T</td>
</tr>
<tr>
<td>mile</td>
<td>nothing</td>
<td>present</td>
<td>saw</td>
<td>smile</td>
<td>table</td>
</tr>
<tr>
<td>milk</td>
<td>now</td>
<td>press</td>
<td>say</td>
<td>smoke</td>
<td>tail</td>
</tr>
<tr>
<td>mill</td>
<td>number</td>
<td>pretty</td>
<td>school</td>
<td>snow</td>
<td>take</td>
</tr>
<tr>
<td>mind</td>
<td>pul</td>
<td>sea</td>
<td>so</td>
<td>talk</td>
<td></td>
</tr>
<tr>
<td>mine</td>
<td>O</td>
<td>put</td>
<td>season</td>
<td>soft</td>
<td>tall</td>
</tr>
<tr>
<td>minute</td>
<td>oak</td>
<td></td>
<td>seat</td>
<td>sold</td>
<td>taste</td>
</tr>
<tr>
<td>miss</td>
<td>ocean</td>
<td>Q</td>
<td>second</td>
<td>soldier</td>
<td>teach</td>
</tr>
<tr>
<td>money</td>
<td>of</td>
<td>quarter</td>
<td>see</td>
<td>some</td>
<td>teacher</td>
</tr>
</tbody>
</table>
The Classic Readability Studies 1944—The Lorge Readability Index

tear three try walk wheel without
tell through turn wall when women
ten throw twelve want where wonder
than tie twenty war whether would
thank till two warn which word
that time was while work
the tire (d) U wash white world
their to uncle waste who would
them today under watch whole write
then together until water whom wrong
there told up wave whose
these tomorrow upon way why
they tongue us we wide Y
thick too use wear wild yard
thin took weather will year
thing top V week win yellow
think touch valley well wind yes
this town very went window yesterday
those trade visit were wing yet
though train west winter you
thought tree W what wish young
thousand true wait wheat with your

BIBLIOGRAPHY


LEWRENCE, ALFRED S. “Measurement of the Difficulty of the Difficulty of Reading Mate-
1944—The Lorge Readability Index


The Classic Readability Studies


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 online 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. 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 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
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." ¹

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. 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. 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 dictionary to be certain that all affixes were included and that no non-affixes were included, we found that the work was too time-

---

Mr. Flesch's reasons for using affixes as a count of difficulty are very well stated, with statistical evidence, in both his books. 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." 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." 

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.

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?

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

---

5 Flesch, Marks of Readable Style, p. 32.
6 Ibid., p. 24.
7 Lorge, loc. cit., 406.
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 Geraldine 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.8

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.\(^9\)

This passage has 7 personal references per hundred words. According to Flesch’s Quick Reference Chart,\(^10\) 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,\(^11\) the same passages used by Mr. Lorge and Mr. 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.\(^12\) These data sheets also included the Flesch counts of affixed morphemes and personal references.

---


\(^10\) Flesch, The Art of Plain Talk, p. 205.


\(^12\) The authors wish to thank Mr. Lorge for making the data sheets available and for permission to publish the intercorrelations of his factors.
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,\textsuperscript{13} Buckingham and Dolch,\textsuperscript{14} 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.

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.\textsuperscript{15}

We punched this information on Hollerith cards, along with the information made available by Mr. Lorge. The intercorrelations appear in Table I.\textsuperscript{16}

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,

\textsuperscript{13} Thorndike, Edward L. \textit{A Teacher’s Word Book of Twenty Thousand Words}. New York: Teachers College, Columbia University, 1931.
\textsuperscript{14} Buckingham, B. R., and Dolch, E. W. \textit{A Combined Word List}. Boston: Ginn and Company, 1936
\textsuperscript{15} The specific instructions for counting will be included in the instructions which will appear in next month’s issue.
\textsuperscript{16} 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.
.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.
TABLE I

Intercorrelations between Four Style elements and grade Score of a Pupil Who Answered One-Half of the Questions on McCall and Crabbs

<table>
<thead>
<tr>
<th></th>
<th>Dale Score (3,000 List)</th>
<th>Flesch A-fixed Morphemes</th>
<th>Flesch Personal References</th>
<th>Lorge Hard Words (Dale 769)</th>
<th>Criterion C_{50}</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average sentence length</td>
<td>0.5108</td>
<td>0.4428</td>
<td>-0.2201</td>
<td>0.4913</td>
<td>0.4681</td>
<td>16.803</td>
<td>5.381</td>
</tr>
<tr>
<td>Dale score (words outside 3,000 list)</td>
<td>...</td>
<td>0.7932</td>
<td>-0.4033</td>
<td>0.7988</td>
<td>0.6833</td>
<td>8.1011</td>
<td>6.305</td>
</tr>
<tr>
<td>Flesch a-fixed morphemes†</td>
<td>...</td>
<td>...</td>
<td>-0.3254</td>
<td>0.7441</td>
<td>0.6017</td>
<td>25.281</td>
<td>11.06</td>
</tr>
<tr>
<td>Flesch personal references†</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>-0.3422</td>
<td>7.8245</td>
<td>5.543</td>
</tr>
<tr>
<td>Lorge hard words (outside Dale 769 list)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>17.416</td>
<td>7.165</td>
</tr>
<tr>
<td>Criterion (C_{50})</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>5.7492</td>
<td>1.656</td>
</tr>
</tbody>
</table>

* 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 0.6174. 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.

The next highest measure of difficulty is average sentence length—which correlates 0.4681 with the criterion.

After making several combinations of factors, we found that the following two, plus a constant, gave the most efficient empirical formula:

\[ X_{C_{50}} = 0.1579X_1 + 0.0496X_2 + 3.6365 \]

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 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.

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:

<table>
<thead>
<tr>
<th>Formula Score</th>
<th>Corrected Grade Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.9 and below</td>
<td>Grade 4 and below</td>
</tr>
<tr>
<td>5.0 to 5.9</td>
<td>Grades 5-6</td>
</tr>
<tr>
<td>6.0 to 6.9</td>
<td>Grades 7-8</td>
</tr>
<tr>
<td>7.0 to 7.9</td>
<td>Grades 9-10</td>
</tr>
<tr>
<td>8.0 to 8.9</td>
<td>Grades 11-12</td>
</tr>
<tr>
<td>9.0 to 9.9</td>
<td>Grades 13-15 (college)</td>
</tr>
<tr>
<td>10 and above</td>
<td>Grades 16 and above (college graduate)</td>
</tr>
</tbody>
</table>
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]
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. 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 sheet 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. 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. 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 sen-

---

3 See page 43 of this issue. Mimeographed copies of the work sheet may be obtained from Edgar Dale, Bureau of Educational Research, Ohio State University.
4 See pages 42 and 44.
5 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, XXXVII (February, 1944), pp. 441-50.
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 list are considered unfamiliar. Underline all unfamiliar words, even if they appear more than once.

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).

6 See the Dale list on pages 25-34
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.

4. **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 of 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.⁷

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

---

⁷ Copies of the table of multiplications may be obtained from Edgar Dale, Bureau of Educational Research, Ohio State University.
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.

<table>
<thead>
<tr>
<th>Formula Raw Score</th>
<th>Corrected Grade-Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.9 and below</td>
<td>4-th grade and below</td>
</tr>
<tr>
<td>5.0 to 5.9</td>
<td>5-6th grade</td>
</tr>
<tr>
<td>6.0 to 6.9</td>
<td>7-8th grade</td>
</tr>
<tr>
<td>7.0 to 7.9</td>
<td>9-10th grade</td>
</tr>
<tr>
<td>8.0 to 8.9</td>
<td>11-12th grade</td>
</tr>
<tr>
<td>9.0 to 9.9</td>
<td>13-15th grade (college)</td>
</tr>
<tr>
<td>10.0 and above</td>
<td>16-(college graduate)</td>
</tr>
</tbody>
</table>

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 worksheet 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 2
A Work Sheet Filled in for the Samples Taken from the Pamphlet “Your Baby”

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Your Baby</td>
<td>2</td>
<td>7</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author:</th>
<th>From:</th>
<th>From:</th>
<th>From:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“A happy...”</td>
<td>“Diphtheria...”</td>
<td>“The germs...”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Publisher:</th>
<th>Date:</th>
<th>To:</th>
<th>To:</th>
<th>To:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nat’l TB Assoc.</td>
<td>1945</td>
<td>“...prevented.”</td>
<td>“...often given.”</td>
<td>“...or boiled.”</td>
</tr>
</tbody>
</table>

| 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 the Dale list | 6 | 20 | 17 |
| 4. Average sentence length (divide 1 by 2) | 19 | 15 | 19 |
| 5. Dale scored (divide 3 by 1, multiply by 100) | 7 | 9 | 6 |
| 6. Multiply average sentence length by .0496 | .9424 | .7440 | .9424 |
| 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 | 6.9474 |

Average raw score of 3 samples...... 6.35

Average corrected grade level...... 7-8

Analyzed by J.S.C Date 1/28/48

Checked by C.D.C Date 1/28/48
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

<table>
<thead>
<tr>
<th>a</th>
<th>ah</th>
<th>an</th>
<th>armful</th>
<th>awhile</th>
<th>barrel</th>
</tr>
</thead>
<tbody>
<tr>
<td>able</td>
<td>ahead</td>
<td>and</td>
<td>army</td>
<td>ax</td>
<td>base</td>
</tr>
<tr>
<td>aboard</td>
<td>aid</td>
<td>angel</td>
<td>arose</td>
<td>baa</td>
<td>baseball</td>
</tr>
<tr>
<td>about</td>
<td>aim</td>
<td>anger</td>
<td>around</td>
<td>babe</td>
<td>basement</td>
</tr>
<tr>
<td>above</td>
<td>air</td>
<td>angry</td>
<td>arrange</td>
<td>baby (ies)</td>
<td>basket</td>
</tr>
<tr>
<td>absent</td>
<td>airfield</td>
<td>animal</td>
<td>arrive (d)</td>
<td>back</td>
<td>bat</td>
</tr>
<tr>
<td>accept</td>
<td>airport</td>
<td>another</td>
<td>arrow</td>
<td>background</td>
<td>batch</td>
</tr>
<tr>
<td>accident</td>
<td>airplane</td>
<td>answer</td>
<td>art</td>
<td>backward (s)</td>
<td>bath</td>
</tr>
<tr>
<td>account</td>
<td>airship</td>
<td>ant</td>
<td>artist</td>
<td>bacon</td>
<td>bathe</td>
</tr>
<tr>
<td>ache (ing)</td>
<td>airy</td>
<td>any</td>
<td>as</td>
<td>bad (ly)</td>
<td>bathing</td>
</tr>
<tr>
<td>acorn</td>
<td>alarm</td>
<td>anybody</td>
<td>ash (es)</td>
<td>badge</td>
<td>bathroom</td>
</tr>
<tr>
<td>acre</td>
<td>alike</td>
<td>anyhow</td>
<td>aside</td>
<td>bag</td>
<td>bathtub</td>
</tr>
<tr>
<td>across</td>
<td>alive</td>
<td>anyone</td>
<td>ask</td>
<td>bake (r)</td>
<td>battle</td>
</tr>
<tr>
<td>act</td>
<td>all</td>
<td>anything</td>
<td>asleep</td>
<td>baking</td>
<td>battleship</td>
</tr>
<tr>
<td>add</td>
<td>alley</td>
<td>anyway</td>
<td>at</td>
<td>bakery</td>
<td>bay</td>
</tr>
<tr>
<td>address</td>
<td>alligator</td>
<td>anywhere</td>
<td>ate</td>
<td>ball</td>
<td>be (ing)</td>
</tr>
<tr>
<td>admire</td>
<td>allow</td>
<td>apart</td>
<td>attack</td>
<td>balloon '</td>
<td>beach</td>
</tr>
<tr>
<td>adventure</td>
<td>almost</td>
<td>apartment</td>
<td>attend</td>
<td>banana</td>
<td>bead</td>
</tr>
<tr>
<td>afar</td>
<td>alone</td>
<td>ape</td>
<td>attention</td>
<td>band</td>
<td>beam</td>
</tr>
<tr>
<td>afraid</td>
<td>along</td>
<td>apiece</td>
<td>August</td>
<td>bandage</td>
<td>bean</td>
</tr>
<tr>
<td>after</td>
<td>aloud</td>
<td>appear</td>
<td>aunt</td>
<td>bang</td>
<td>bear</td>
</tr>
<tr>
<td>afternoon</td>
<td>already</td>
<td>apple</td>
<td>author</td>
<td>banjo</td>
<td>beard</td>
</tr>
<tr>
<td>afterward (s)</td>
<td>also</td>
<td>April</td>
<td>auto</td>
<td>bank (er)</td>
<td>beast</td>
</tr>
<tr>
<td>again</td>
<td>always</td>
<td>apron</td>
<td>automobile</td>
<td>bar</td>
<td>beaming</td>
</tr>
<tr>
<td>against</td>
<td>am</td>
<td>are</td>
<td>autumn</td>
<td>barber</td>
<td>beautiful</td>
</tr>
<tr>
<td>age</td>
<td>America</td>
<td>aren't</td>
<td>avenue</td>
<td>bare (ly)</td>
<td>beautify</td>
</tr>
<tr>
<td>aged</td>
<td>American</td>
<td>arise</td>
<td>awake (n)</td>
<td>barefoot</td>
<td>beauty</td>
</tr>
<tr>
<td>ago</td>
<td>among</td>
<td>arithmetic</td>
<td>away</td>
<td>bark</td>
<td>became</td>
</tr>
<tr>
<td>agree</td>
<td>amount</td>
<td>arm</td>
<td>awful (ly)</td>
<td>barn</td>
<td>because</td>
</tr>
<tr>
<td>become</td>
<td>bird</td>
<td>bookkeeper</td>
<td>bubble</td>
<td>calendar</td>
<td>cattle</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>------------</td>
<td>--------</td>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>becoming</td>
<td>birth</td>
<td>boom</td>
<td>bucket</td>
<td>calf</td>
<td>caught</td>
</tr>
<tr>
<td>bed</td>
<td>birthday</td>
<td>boot</td>
<td>buckle</td>
<td>call (er) (ing)</td>
<td>cause</td>
</tr>
<tr>
<td>bedbug</td>
<td>biscuit</td>
<td>born</td>
<td>bud</td>
<td>came</td>
<td>cave</td>
</tr>
<tr>
<td>bedroom</td>
<td>bit</td>
<td>borrow</td>
<td>buffalo</td>
<td>camel</td>
<td>ceiling</td>
</tr>
<tr>
<td>beds</td>
<td>bite</td>
<td>boss</td>
<td>bug</td>
<td>camp</td>
<td>cell</td>
</tr>
<tr>
<td>bedtime</td>
<td>biting</td>
<td>both</td>
<td>buggy</td>
<td>campfire</td>
<td>cellar</td>
</tr>
<tr>
<td>bee</td>
<td>bitter</td>
<td>bother</td>
<td>build</td>
<td>can</td>
<td>cent</td>
</tr>
<tr>
<td>beech</td>
<td>black</td>
<td>bottle</td>
<td>building</td>
<td>canal</td>
<td>center</td>
</tr>
<tr>
<td>beef</td>
<td>blackberry</td>
<td>bottom</td>
<td>built</td>
<td>canary</td>
<td>cereal</td>
</tr>
<tr>
<td>beefsteak</td>
<td>blackbird</td>
<td>bought</td>
<td>bulb</td>
<td>candle</td>
<td>certain (ly)</td>
</tr>
<tr>
<td>bee hive</td>
<td>blackboard</td>
<td>bounce</td>
<td>bull</td>
<td>candlestick</td>
<td>chair</td>
</tr>
<tr>
<td>been</td>
<td>blackness</td>
<td>bow</td>
<td>bullet</td>
<td>candy</td>
<td>chair</td>
</tr>
<tr>
<td>beer</td>
<td>blacksmith</td>
<td>bowl</td>
<td>bum</td>
<td>cane</td>
<td>chalk</td>
</tr>
<tr>
<td>beet</td>
<td>blame</td>
<td>bow-wow</td>
<td>bumblebee</td>
<td>cannon</td>
<td>champion</td>
</tr>
<tr>
<td>before</td>
<td>blank</td>
<td>box (es)</td>
<td>bump</td>
<td>cannot</td>
<td>chance</td>
</tr>
<tr>
<td>beg</td>
<td>blanket</td>
<td>boxcar</td>
<td>bun</td>
<td>canoe</td>
<td>change</td>
</tr>
<tr>
<td>began</td>
<td>blast</td>
<td>boxer</td>
<td>bunch</td>
<td>can't</td>
<td>chap</td>
</tr>
<tr>
<td>beggar</td>
<td>blaze</td>
<td>boy</td>
<td>bundle</td>
<td>canyon</td>
<td>charge</td>
</tr>
<tr>
<td>begged</td>
<td>bleed</td>
<td>boyhood</td>
<td>bunny</td>
<td>cap</td>
<td>charm</td>
</tr>
<tr>
<td>begin</td>
<td>bless</td>
<td>bracelet</td>
<td>burn</td>
<td>cape</td>
<td>chart</td>
</tr>
<tr>
<td>begun</td>
<td>blew</td>
<td>brake</td>
<td>bury</td>
<td>captain</td>
<td>chatter</td>
</tr>
<tr>
<td>behave</td>
<td>blind (s)</td>
<td>bran</td>
<td>bus</td>
<td>car</td>
<td>cheap</td>
</tr>
<tr>
<td>behind</td>
<td>blindfold</td>
<td>branch</td>
<td>bush</td>
<td>card</td>
<td>cheat</td>
</tr>
<tr>
<td>believe</td>
<td>block</td>
<td>brass</td>
<td>bushel</td>
<td>cardboard</td>
<td>check</td>
</tr>
<tr>
<td>bell</td>
<td>blood</td>
<td>brave</td>
<td>business</td>
<td>care</td>
<td>checkers</td>
</tr>
<tr>
<td>belong</td>
<td>bloom</td>
<td>bread</td>
<td>busy</td>
<td>careful</td>
<td>cheek</td>
</tr>
<tr>
<td>below</td>
<td>blossom</td>
<td>break</td>
<td>but</td>
<td>careless</td>
<td>cheer</td>
</tr>
<tr>
<td>belt</td>
<td>blot</td>
<td>breakfast</td>
<td>butcher</td>
<td>carelessness</td>
<td>cheese</td>
</tr>
<tr>
<td>beneath</td>
<td>blow</td>
<td>breast</td>
<td>but</td>
<td>carload</td>
<td>cherry</td>
</tr>
<tr>
<td>bench</td>
<td>blue</td>
<td>breath</td>
<td>butter</td>
<td>carpenter</td>
<td>chest</td>
</tr>
<tr>
<td>bend</td>
<td>blueberry</td>
<td>breathe</td>
<td>buttercup</td>
<td>carpet</td>
<td>chew</td>
</tr>
<tr>
<td>bent</td>
<td>bluebird</td>
<td>breeze</td>
<td>butterfly</td>
<td>carriage</td>
<td>chick</td>
</tr>
<tr>
<td>berry (ies)</td>
<td>blue jay</td>
<td>brick</td>
<td>buttermilk</td>
<td>carrot</td>
<td>chicken</td>
</tr>
<tr>
<td>besides (s)</td>
<td>blush</td>
<td>bride</td>
<td>butterscotch</td>
<td>carry</td>
<td>chief</td>
</tr>
<tr>
<td>best</td>
<td>board</td>
<td>bridge</td>
<td>button</td>
<td>cart</td>
<td>child</td>
</tr>
<tr>
<td>bet</td>
<td>boast</td>
<td>bright</td>
<td>buttonhole</td>
<td>carve</td>
<td>childhood</td>
</tr>
<tr>
<td>better</td>
<td>boat</td>
<td>brightness</td>
<td>buy</td>
<td>case</td>
<td>children</td>
</tr>
<tr>
<td>between</td>
<td>bob</td>
<td>bring</td>
<td>buzz</td>
<td>cash</td>
<td>chill (y)</td>
</tr>
<tr>
<td>bib</td>
<td>bobwhite</td>
<td>broad</td>
<td>by</td>
<td>cashier</td>
<td>chimney</td>
</tr>
<tr>
<td>bible</td>
<td>body (ies)</td>
<td>broadcast</td>
<td>bye</td>
<td>castle</td>
<td>chin</td>
</tr>
<tr>
<td>bicycle</td>
<td>boil (er)</td>
<td>broke(n)</td>
<td>cab</td>
<td>cat</td>
<td>china</td>
</tr>
<tr>
<td>bid</td>
<td>bold</td>
<td>brook</td>
<td>cabbage</td>
<td>catbird</td>
<td>chip</td>
</tr>
<tr>
<td>big (ger)</td>
<td>bone</td>
<td>broom</td>
<td>cabin</td>
<td>catch</td>
<td>chipmunk</td>
</tr>
<tr>
<td>bill</td>
<td>bonnet</td>
<td>brother</td>
<td>cabinet</td>
<td>catcher</td>
<td>chocolate</td>
</tr>
<tr>
<td>billboard</td>
<td>boo</td>
<td>brought</td>
<td>cackle</td>
<td>caterpillar</td>
<td>choice</td>
</tr>
<tr>
<td>bin</td>
<td>book</td>
<td>brown</td>
<td>cage</td>
<td>catfish</td>
<td>choose</td>
</tr>
<tr>
<td>bind</td>
<td>bookcase</td>
<td>brush</td>
<td>cake</td>
<td>catsup</td>
<td>chop</td>
</tr>
<tr>
<td>chorus</td>
<td>codfish</td>
<td>cow</td>
<td>dad</td>
<td>destroy</td>
<td>downstairs</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>chose</td>
<td>coffee</td>
<td>coward</td>
<td>daddy</td>
<td>devil</td>
<td>downtown</td>
</tr>
<tr>
<td>christen</td>
<td>coffeepot</td>
<td>cowboy</td>
<td>daily</td>
<td>dew</td>
<td>dozen</td>
</tr>
<tr>
<td>Christmas</td>
<td>coin</td>
<td>cozy</td>
<td>dairy</td>
<td>diamond</td>
<td>drag</td>
</tr>
<tr>
<td>church</td>
<td>cold</td>
<td>crab</td>
<td>daisy</td>
<td>didn't</td>
<td>drank</td>
</tr>
<tr>
<td>churn</td>
<td>collar</td>
<td>crack</td>
<td>dam</td>
<td>didn't</td>
<td>drank</td>
</tr>
<tr>
<td>cigarette</td>
<td>college</td>
<td>cracker</td>
<td>damage</td>
<td>die (d)</td>
<td>draw (er)</td>
</tr>
<tr>
<td>circle</td>
<td>color</td>
<td>cradle</td>
<td>dame</td>
<td>difference</td>
<td>draw (ing)</td>
</tr>
<tr>
<td>circus</td>
<td>colt</td>
<td>cramps</td>
<td>damp</td>
<td>different</td>
<td>dream</td>
</tr>
<tr>
<td>citizen</td>
<td>column</td>
<td>cranberry</td>
<td>dance (r)</td>
<td>dig</td>
<td>-dress</td>
</tr>
<tr>
<td>city</td>
<td>comb</td>
<td>crank (y)</td>
<td>dancing</td>
<td>dim</td>
<td>dresser</td>
</tr>
<tr>
<td>clang</td>
<td>come</td>
<td>crash</td>
<td>dandy</td>
<td>dime</td>
<td>dressmaker</td>
</tr>
<tr>
<td>clap</td>
<td>comfort</td>
<td>crawl</td>
<td>danger (ous)</td>
<td>dine</td>
<td>drew</td>
</tr>
<tr>
<td>class</td>
<td>comic</td>
<td>crazy</td>
<td>dare</td>
<td>ding-dong</td>
<td>dried</td>
</tr>
<tr>
<td>classmate</td>
<td>coming</td>
<td>cream (y)</td>
<td>dark (ness)</td>
<td>dinner</td>
<td>drift</td>
</tr>
<tr>
<td>classroom</td>
<td>company</td>
<td>creek</td>
<td>darling</td>
<td>dip</td>
<td>drill</td>
</tr>
<tr>
<td>claw</td>
<td>compare</td>
<td>creep</td>
<td>darn</td>
<td>direct</td>
<td>drink</td>
</tr>
<tr>
<td>clay</td>
<td>conductor</td>
<td>crept</td>
<td>dart</td>
<td>direction</td>
<td>drip</td>
</tr>
<tr>
<td>clean(er)</td>
<td>cone</td>
<td>-cried</td>
<td>dash</td>
<td>dirt (y)</td>
<td>drive (n)</td>
</tr>
<tr>
<td>clear</td>
<td>connect</td>
<td>croak</td>
<td>date</td>
<td>discover</td>
<td>driver</td>
</tr>
<tr>
<td>clerk</td>
<td>coo</td>
<td>crook(ed)</td>
<td>daughter</td>
<td>dish</td>
<td>drop</td>
</tr>
<tr>
<td>clever</td>
<td>cook (ed)</td>
<td>crop</td>
<td>dawn</td>
<td>dislike</td>
<td>drove</td>
</tr>
<tr>
<td>click</td>
<td>cook (ing)</td>
<td>cross (ing)</td>
<td>day</td>
<td>dismiss</td>
<td>drown</td>
</tr>
<tr>
<td>cliff</td>
<td>cooky (ie)(s)</td>
<td>cross-eyed(d)</td>
<td>daybreak</td>
<td>ditch</td>
<td>drowsy</td>
</tr>
<tr>
<td>climb</td>
<td>cool (er)</td>
<td>crow</td>
<td>daytime</td>
<td>dive</td>
<td>drug</td>
</tr>
<tr>
<td>clip</td>
<td>coop</td>
<td>crowd(ed)</td>
<td>dead</td>
<td>diver</td>
<td>drum</td>
</tr>
<tr>
<td>cloak</td>
<td>copper</td>
<td>crown</td>
<td>deaf</td>
<td>divide</td>
<td>drunk</td>
</tr>
<tr>
<td>clock</td>
<td>copy</td>
<td>cruel</td>
<td>deal</td>
<td>do</td>
<td>dry</td>
</tr>
<tr>
<td>close</td>
<td>cord</td>
<td>crumb</td>
<td>dear</td>
<td>dock</td>
<td>duck</td>
</tr>
<tr>
<td>closet</td>
<td>cork</td>
<td>crumble</td>
<td>death</td>
<td>doctor</td>
<td>due</td>
</tr>
<tr>
<td>cloth</td>
<td>corn</td>
<td>crush</td>
<td>December</td>
<td>does</td>
<td>dug</td>
</tr>
<tr>
<td>clothes</td>
<td>corner</td>
<td>crust.</td>
<td>decide</td>
<td>doesn't</td>
<td>dull</td>
</tr>
<tr>
<td>clothing</td>
<td>correct</td>
<td>cry (ies)</td>
<td>deck</td>
<td>dog</td>
<td>dumb</td>
</tr>
<tr>
<td>cloud (y)</td>
<td>cost</td>
<td>cub</td>
<td>deed</td>
<td>doll</td>
<td>dump</td>
</tr>
<tr>
<td>clever</td>
<td>cot</td>
<td>cuff</td>
<td>deep</td>
<td>dollar</td>
<td>during</td>
</tr>
<tr>
<td>clown</td>
<td>cottage</td>
<td>cup</td>
<td>deer</td>
<td>dolly</td>
<td>dust(y)</td>
</tr>
<tr>
<td>club</td>
<td>cotton</td>
<td>cupboard</td>
<td>defeat</td>
<td>done</td>
<td>duty</td>
</tr>
<tr>
<td>cluck</td>
<td>couch</td>
<td>cupful</td>
<td>defend</td>
<td>donkey</td>
<td>dwarf</td>
</tr>
<tr>
<td>clump</td>
<td>cough</td>
<td>cure</td>
<td>defense</td>
<td>don't</td>
<td>dwell</td>
</tr>
<tr>
<td>coach</td>
<td>could</td>
<td>curl(y)</td>
<td>delight</td>
<td>door</td>
<td>dwelt</td>
</tr>
<tr>
<td>coal</td>
<td>couldn't</td>
<td>curtain</td>
<td>den</td>
<td>doorknell</td>
<td>each</td>
</tr>
<tr>
<td>coast</td>
<td>count</td>
<td>curve</td>
<td>dentist</td>
<td>doorknell</td>
<td>each</td>
</tr>
<tr>
<td>coat</td>
<td>counter</td>
<td>cushion</td>
<td>depend</td>
<td>doorstep</td>
<td>eager</td>
</tr>
<tr>
<td>cob</td>
<td>country</td>
<td>custard</td>
<td>deposit</td>
<td>dope</td>
<td>eagle</td>
</tr>
<tr>
<td>cobbler</td>
<td>county</td>
<td>customer</td>
<td>describe</td>
<td>dot</td>
<td>ear</td>
</tr>
<tr>
<td>cocoa</td>
<td>course</td>
<td>cut</td>
<td>desert</td>
<td>double</td>
<td>early</td>
</tr>
<tr>
<td>coconut</td>
<td>court</td>
<td>cute</td>
<td>deserve</td>
<td>dough</td>
<td>earn</td>
</tr>
<tr>
<td>cocoons</td>
<td>cousin</td>
<td>cutting</td>
<td>desire</td>
<td>dove</td>
<td>earth</td>
</tr>
<tr>
<td>cod</td>
<td>cover</td>
<td>dab</td>
<td>desk</td>
<td>down</td>
<td>eastern</td>
</tr>
<tr>
<td>easy</td>
<td>excited</td>
<td>fellow</td>
<td>flip-flop</td>
<td>French</td>
<td>gift</td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>--------</td>
<td>-----------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>eat (en)</td>
<td>exciting</td>
<td>felt</td>
<td>float</td>
<td>fresh</td>
<td>gingerbread</td>
</tr>
<tr>
<td>edge</td>
<td>excuse</td>
<td>fence</td>
<td>flock</td>
<td>fret</td>
<td>girl</td>
</tr>
<tr>
<td>egg</td>
<td>exit</td>
<td>fever</td>
<td>flood</td>
<td>Friday</td>
<td>give (n);</td>
</tr>
<tr>
<td>eh</td>
<td>expect</td>
<td>few</td>
<td>floor</td>
<td>fried</td>
<td>giving</td>
</tr>
<tr>
<td>eight</td>
<td>explain</td>
<td>fib</td>
<td>flop</td>
<td>friend (ly)</td>
<td>glad (ly)</td>
</tr>
<tr>
<td>eighteen</td>
<td>extra</td>
<td>fiddle</td>
<td>flour</td>
<td>friendship</td>
<td>glance</td>
</tr>
<tr>
<td>eighth</td>
<td>eye</td>
<td>field</td>
<td>flow</td>
<td>frighten</td>
<td>glass (es)</td>
</tr>
<tr>
<td>eighty</td>
<td>eyebrow</td>
<td>fife</td>
<td>flower (y)</td>
<td>frog</td>
<td>glean</td>
</tr>
<tr>
<td>either</td>
<td>fable</td>
<td>fifteen</td>
<td>flutter</td>
<td>from</td>
<td>glide</td>
</tr>
<tr>
<td>elbow</td>
<td>face</td>
<td>fifth</td>
<td>fly</td>
<td>front</td>
<td>glory</td>
</tr>
<tr>
<td>elder</td>
<td>facing</td>
<td>fifty</td>
<td>foam</td>
<td>frost</td>
<td>glove</td>
</tr>
<tr>
<td>eldest</td>
<td>fact</td>
<td>fig</td>
<td>fog</td>
<td>brown</td>
<td>glow</td>
</tr>
<tr>
<td>electric</td>
<td>factory</td>
<td>fight</td>
<td>foggy</td>
<td>froze</td>
<td>glue</td>
</tr>
<tr>
<td>electricity</td>
<td>fail</td>
<td>figure</td>
<td>fold</td>
<td>fruit</td>
<td>go (ing)</td>
</tr>
<tr>
<td>elephant</td>
<td>faint</td>
<td>file</td>
<td>folks</td>
<td>fry</td>
<td>goes</td>
</tr>
<tr>
<td>eleven</td>
<td>fair</td>
<td>fill</td>
<td>follow (ing)</td>
<td>fudge</td>
<td>goal</td>
</tr>
<tr>
<td>elf</td>
<td>fairy</td>
<td>film</td>
<td>fond</td>
<td>fuel</td>
<td>goat</td>
</tr>
<tr>
<td>elm</td>
<td>faith</td>
<td>finally</td>
<td>food</td>
<td>full (y)</td>
<td>gobble</td>
</tr>
<tr>
<td>else</td>
<td>fake</td>
<td>find</td>
<td>fool</td>
<td>fun</td>
<td>God (g)</td>
</tr>
<tr>
<td>elsewhere</td>
<td>fall</td>
<td>fine</td>
<td>foolish</td>
<td>funny</td>
<td>godmother</td>
</tr>
<tr>
<td>empty</td>
<td>false</td>
<td>finger</td>
<td>foot</td>
<td>fur</td>
<td>gold (en)</td>
</tr>
<tr>
<td>end (ing)</td>
<td>family</td>
<td>finish</td>
<td>football</td>
<td>furniture</td>
<td>goldfish</td>
</tr>
<tr>
<td>enemy</td>
<td>fan</td>
<td>fire</td>
<td>footprint</td>
<td>further</td>
<td>golf</td>
</tr>
<tr>
<td>engine</td>
<td>fancy</td>
<td>firearm</td>
<td>for</td>
<td>fuzzy</td>
<td>gone</td>
</tr>
<tr>
<td>engineer</td>
<td>far</td>
<td>firecracker</td>
<td>forehead</td>
<td>gain</td>
<td>good (s)</td>
</tr>
<tr>
<td>English</td>
<td>faraway</td>
<td>fireplace</td>
<td>forest</td>
<td>gallon</td>
<td>good-by (bye)</td>
</tr>
<tr>
<td>enjoy</td>
<td>fare</td>
<td>fireworks</td>
<td>forget</td>
<td>gallop</td>
<td>good-looking</td>
</tr>
<tr>
<td>enough</td>
<td>farmer</td>
<td>firing</td>
<td>forgive</td>
<td>game</td>
<td>goodness</td>
</tr>
<tr>
<td>enter</td>
<td>farm (ing)</td>
<td>first</td>
<td>forgot (ten)</td>
<td>gang</td>
<td>goody</td>
</tr>
<tr>
<td>envelope</td>
<td>far-off</td>
<td>fish</td>
<td>fork</td>
<td>garage</td>
<td>goose</td>
</tr>
<tr>
<td>equal</td>
<td>farther</td>
<td>fisherman</td>
<td>form</td>
<td>garbage</td>
<td>gooseberry</td>
</tr>
<tr>
<td>erase (r)</td>
<td>fashion</td>
<td>fist</td>
<td>fort</td>
<td>garden</td>
<td>got</td>
</tr>
<tr>
<td>errand</td>
<td>fast</td>
<td>fit (s)</td>
<td>forth</td>
<td>gas</td>
<td>govern</td>
</tr>
<tr>
<td>escape</td>
<td>fasten</td>
<td>five&quot;</td>
<td>fortune</td>
<td>gasoline</td>
<td>government</td>
</tr>
<tr>
<td>eve</td>
<td>fat</td>
<td>fix</td>
<td>forty</td>
<td>gate</td>
<td>gown</td>
</tr>
<tr>
<td>even</td>
<td>father</td>
<td>flag</td>
<td>forward</td>
<td>gather</td>
<td>grab</td>
</tr>
<tr>
<td>evening</td>
<td>fault</td>
<td>flame</td>
<td>fought</td>
<td>gave</td>
<td>gracious</td>
</tr>
<tr>
<td>ever</td>
<td>favor</td>
<td>flame</td>
<td>found</td>
<td>gay</td>
<td>grade</td>
</tr>
<tr>
<td>every</td>
<td>favorite</td>
<td>flap</td>
<td>fountain</td>
<td>gear</td>
<td>grain</td>
</tr>
<tr>
<td>everybody</td>
<td>fear</td>
<td>flash</td>
<td>four</td>
<td>geese</td>
<td>grand</td>
</tr>
<tr>
<td>everyday</td>
<td>feast</td>
<td>flashlight</td>
<td>fourteen</td>
<td>general</td>
<td>grandchild</td>
</tr>
<tr>
<td>everyone</td>
<td>feather</td>
<td>flat</td>
<td>fourth</td>
<td>gentle</td>
<td>grandchildren</td>
</tr>
<tr>
<td>everything</td>
<td>February</td>
<td>flea</td>
<td>fox</td>
<td>gentleman</td>
<td>granddaughters</td>
</tr>
<tr>
<td>everywhere</td>
<td>fed</td>
<td>flesh</td>
<td>frame</td>
<td>gentlemen</td>
<td>grandfather</td>
</tr>
<tr>
<td>evil</td>
<td>feed</td>
<td>flew</td>
<td>free</td>
<td>geography</td>
<td>grandma</td>
</tr>
<tr>
<td>exact</td>
<td>feel</td>
<td>flies</td>
<td>freedom</td>
<td>get</td>
<td>grandmother</td>
</tr>
<tr>
<td>except</td>
<td>feet</td>
<td>flight</td>
<td>freeze</td>
<td>getting</td>
<td>grandpa</td>
</tr>
<tr>
<td>exchange</td>
<td>fell</td>
<td>flip</td>
<td>freight</td>
<td>giant</td>
<td>grandson</td>
</tr>
<tr>
<td>grandstand</td>
<td>handle</td>
<td>held</td>
<td>homely</td>
<td>hush</td>
<td>January</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>grape (s)</td>
<td>handwriting</td>
<td>hell</td>
<td>homesick</td>
<td>hut</td>
<td>jar</td>
</tr>
<tr>
<td>grapefruit</td>
<td>hang</td>
<td>he'll</td>
<td>honest</td>
<td>hymn</td>
<td>jaw</td>
</tr>
<tr>
<td>grass</td>
<td>happen</td>
<td>hello</td>
<td>honey</td>
<td>I</td>
<td>jay</td>
</tr>
<tr>
<td>grasshopper</td>
<td>happily</td>
<td>helmet</td>
<td>honeybee</td>
<td>ice</td>
<td>jelly</td>
</tr>
<tr>
<td>grateful</td>
<td>happiness</td>
<td>help (er)</td>
<td>honeymoon</td>
<td>icy</td>
<td>jellyfish</td>
</tr>
<tr>
<td>grave</td>
<td>happy</td>
<td>helpful</td>
<td>honk</td>
<td>I'd</td>
<td>jerk</td>
</tr>
<tr>
<td>gravel</td>
<td>harbor</td>
<td>hem</td>
<td>honor</td>
<td>idea</td>
<td>jig</td>
</tr>
<tr>
<td>graveyard</td>
<td>hard</td>
<td>hen</td>
<td>hood</td>
<td>ideal</td>
<td>job</td>
</tr>
<tr>
<td>gravy</td>
<td>hardly</td>
<td>henhouse</td>
<td>hoof</td>
<td>if</td>
<td>jockey</td>
</tr>
<tr>
<td>gray</td>
<td>hardship</td>
<td>her (s)</td>
<td>hook</td>
<td>ill</td>
<td>join</td>
</tr>
<tr>
<td>graze</td>
<td>hardware</td>
<td>herd</td>
<td>hoop</td>
<td>I'll</td>
<td>joke</td>
</tr>
<tr>
<td>grease</td>
<td>hare</td>
<td>here</td>
<td>hop</td>
<td>I'm</td>
<td>joking</td>
</tr>
<tr>
<td>great</td>
<td>hark</td>
<td>here's</td>
<td>hope (ful)</td>
<td>important</td>
<td>jolly</td>
</tr>
<tr>
<td>green</td>
<td>harm</td>
<td>hero'</td>
<td>hopeless</td>
<td>impossible</td>
<td>journey</td>
</tr>
<tr>
<td>greet</td>
<td>harness</td>
<td>herself</td>
<td>horn</td>
<td>improve</td>
<td>joy (ful)</td>
</tr>
<tr>
<td>grew</td>
<td>harp</td>
<td>he's</td>
<td>horse</td>
<td>in</td>
<td>joyous</td>
</tr>
<tr>
<td>grind</td>
<td>harvest</td>
<td>hey</td>
<td>horseback</td>
<td>inch (es)</td>
<td>judge</td>
</tr>
<tr>
<td>groan</td>
<td>has</td>
<td>hickory</td>
<td>horseshoe</td>
<td>income</td>
<td>jug</td>
</tr>
<tr>
<td>grocery</td>
<td>hasn't</td>
<td>hid</td>
<td>hose</td>
<td>indeed</td>
<td>juice</td>
</tr>
<tr>
<td>ground</td>
<td>haste (n)</td>
<td>hidden</td>
<td>hospital</td>
<td>Indian</td>
<td>juicy</td>
</tr>
<tr>
<td>group</td>
<td>hasty</td>
<td>hide</td>
<td>host</td>
<td>indoors</td>
<td>July</td>
</tr>
<tr>
<td>grove</td>
<td>hat</td>
<td>high</td>
<td>hot</td>
<td>ink</td>
<td>jump</td>
</tr>
<tr>
<td>grow</td>
<td>hatch</td>
<td>highway</td>
<td>hotel</td>
<td>inn</td>
<td>June</td>
</tr>
<tr>
<td>guard</td>
<td>hatchet</td>
<td>hill</td>
<td>hound</td>
<td>insect</td>
<td>junior</td>
</tr>
<tr>
<td>guess</td>
<td>hate</td>
<td>hillside</td>
<td>hour</td>
<td>inside</td>
<td>junk</td>
</tr>
<tr>
<td>guest</td>
<td>haul</td>
<td>hilltop</td>
<td>house</td>
<td>instant</td>
<td>just</td>
</tr>
<tr>
<td>guide</td>
<td>have</td>
<td>hilly</td>
<td>household</td>
<td>instead</td>
<td>keen</td>
</tr>
<tr>
<td>gulf</td>
<td>haven't</td>
<td>him</td>
<td>housewife</td>
<td>insult</td>
<td>keep</td>
</tr>
<tr>
<td>gum</td>
<td>having</td>
<td>himself</td>
<td>housework</td>
<td>intend</td>
<td>kept</td>
</tr>
<tr>
<td>gun</td>
<td>hawk</td>
<td>hind</td>
<td>how</td>
<td>interested</td>
<td>kettle</td>
</tr>
<tr>
<td>gunpowder</td>
<td>hay</td>
<td>hint</td>
<td>however</td>
<td>interesting</td>
<td>key</td>
</tr>
<tr>
<td>guy</td>
<td>hayfield</td>
<td>hip</td>
<td>howl</td>
<td>into</td>
<td>kick</td>
</tr>
<tr>
<td>ha</td>
<td>haystack</td>
<td>hire</td>
<td>hug</td>
<td>invite</td>
<td>kid</td>
</tr>
<tr>
<td>habit</td>
<td>he</td>
<td>his</td>
<td>huge</td>
<td>iron</td>
<td>killed (ed)</td>
</tr>
<tr>
<td>had</td>
<td>head</td>
<td>hiss</td>
<td>hum</td>
<td>is</td>
<td>kind (ly)</td>
</tr>
<tr>
<td>hadn't</td>
<td>headache</td>
<td>history</td>
<td>humble</td>
<td>island</td>
<td>kindness</td>
</tr>
<tr>
<td>hail</td>
<td>heal</td>
<td>hit</td>
<td>hump</td>
<td>isn't</td>
<td>king</td>
</tr>
<tr>
<td>hair</td>
<td>health (y)</td>
<td>hitch</td>
<td>hundred</td>
<td>it</td>
<td>kingdom</td>
</tr>
<tr>
<td>haircut</td>
<td>heap</td>
<td>hive</td>
<td>hung</td>
<td>its</td>
<td>kiss</td>
</tr>
<tr>
<td>hairpin</td>
<td>hear (ing)</td>
<td>ho</td>
<td>hunger</td>
<td>it's</td>
<td>kitchen</td>
</tr>
<tr>
<td>half</td>
<td>heard</td>
<td>hoe</td>
<td>hungry</td>
<td>itself</td>
<td>kite</td>
</tr>
<tr>
<td>hall</td>
<td>heart</td>
<td>hog</td>
<td>hunk</td>
<td>I've</td>
<td>kitten</td>
</tr>
<tr>
<td>halt</td>
<td>heat (er)</td>
<td>hold (er)</td>
<td>hunt (er)</td>
<td>ivory</td>
<td>kitty</td>
</tr>
<tr>
<td>ham</td>
<td>heaven</td>
<td>hole</td>
<td>hurrah</td>
<td>ivy</td>
<td>knee</td>
</tr>
<tr>
<td>hammer</td>
<td>heavy</td>
<td>holiday</td>
<td>hurried</td>
<td>jacket</td>
<td>kneel</td>
</tr>
<tr>
<td>hand</td>
<td>he'd</td>
<td>hollow</td>
<td>hurry</td>
<td>jacks</td>
<td>knew</td>
</tr>
<tr>
<td>handful</td>
<td>heel</td>
<td>holy</td>
<td>hurt</td>
<td>jail</td>
<td>knife</td>
</tr>
<tr>
<td>handkerchief</td>
<td>height</td>
<td>home</td>
<td>husband</td>
<td>jam</td>
<td>knit</td>
</tr>
</tbody>
</table>
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 milksman 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
<table>
<thead>
<tr>
<th>1948—The Dale-Chall Readability Formula</th>
<th>The Classic Readability Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>nothing</td>
<td>outfit</td>
</tr>
<tr>
<td>notice</td>
<td>outlaw</td>
</tr>
<tr>
<td>November</td>
<td>outline</td>
</tr>
<tr>
<td>now</td>
<td>outside</td>
</tr>
<tr>
<td>nowhere</td>
<td>outward</td>
</tr>
<tr>
<td>number</td>
<td>oven</td>
</tr>
<tr>
<td>nurse</td>
<td>over</td>
</tr>
<tr>
<td>nut</td>
<td>overalls</td>
</tr>
<tr>
<td>oak</td>
<td>overcoat</td>
</tr>
<tr>
<td>oar</td>
<td>overeat</td>
</tr>
<tr>
<td>oatmeal</td>
<td>overhead</td>
</tr>
<tr>
<td>oats</td>
<td>overhear</td>
</tr>
<tr>
<td>obey</td>
<td>overnight</td>
</tr>
<tr>
<td>ocean</td>
<td>overturn</td>
</tr>
<tr>
<td>o'clock</td>
<td>owe</td>
</tr>
<tr>
<td>October</td>
<td>owing</td>
</tr>
<tr>
<td>odd</td>
<td>owl</td>
</tr>
<tr>
<td>of</td>
<td>own (er)</td>
</tr>
<tr>
<td>off</td>
<td>ox</td>
</tr>
<tr>
<td>offer</td>
<td>pa</td>
</tr>
<tr>
<td>office</td>
<td>pace</td>
</tr>
<tr>
<td>officer</td>
<td>pack</td>
</tr>
<tr>
<td>often</td>
<td>package</td>
</tr>
<tr>
<td>oh</td>
<td>pad</td>
</tr>
<tr>
<td>oil</td>
<td>page</td>
</tr>
<tr>
<td>old</td>
<td>paid</td>
</tr>
<tr>
<td>old-fashioned</td>
<td>pain (ful)</td>
</tr>
<tr>
<td>on</td>
<td>paint (er)</td>
</tr>
<tr>
<td>once</td>
<td>painting</td>
</tr>
<tr>
<td>one</td>
<td>pair</td>
</tr>
<tr>
<td>onion</td>
<td>pal</td>
</tr>
<tr>
<td>only</td>
<td>palace</td>
</tr>
<tr>
<td>onward</td>
<td>pale</td>
</tr>
<tr>
<td>open</td>
<td>pan</td>
</tr>
<tr>
<td>or</td>
<td>pancake</td>
</tr>
<tr>
<td>orange</td>
<td>pane</td>
</tr>
<tr>
<td>orchard</td>
<td>pansy</td>
</tr>
<tr>
<td>order</td>
<td>pants</td>
</tr>
<tr>
<td>ore</td>
<td>papa</td>
</tr>
<tr>
<td>organ</td>
<td>paper</td>
</tr>
<tr>
<td>other</td>
<td>parade</td>
</tr>
<tr>
<td>otherwise</td>
<td>pardon</td>
</tr>
<tr>
<td>ouch</td>
<td>parent</td>
</tr>
<tr>
<td>ought</td>
<td>park</td>
</tr>
<tr>
<td>our (s)</td>
<td>part (ly)</td>
</tr>
<tr>
<td>ourselves</td>
<td>partner</td>
</tr>
<tr>
<td>out</td>
<td>party</td>
</tr>
<tr>
<td>outdoors</td>
<td>pass</td>
</tr>
</tbody>
</table>
The Classic Readability Studies
1948—The Dale-Chall Readability Formula

<table>
<thead>
<tr>
<th>rear</th>
<th>rock (y)</th>
<th>sand (y)</th>
<th>self</th>
<th>shirt</th>
<th>sixteen</th>
</tr>
</thead>
<tbody>
<tr>
<td>reason</td>
<td>rocket</td>
<td>sandwich</td>
<td>selfish</td>
<td>shock</td>
<td>sixth</td>
</tr>
<tr>
<td>rebuild</td>
<td>rode</td>
<td>sang</td>
<td>sell</td>
<td>shoe</td>
<td>sixty</td>
</tr>
<tr>
<td>receive</td>
<td>roll</td>
<td>sank</td>
<td>send</td>
<td>shoemaker</td>
<td>size</td>
</tr>
<tr>
<td>recess</td>
<td>roller</td>
<td>sap</td>
<td>sense</td>
<td>shone</td>
<td>skate</td>
</tr>
<tr>
<td>record</td>
<td>roof</td>
<td>sash</td>
<td>sent</td>
<td>shoo</td>
<td>skater</td>
</tr>
<tr>
<td>red</td>
<td>room</td>
<td>sat</td>
<td>sentence</td>
<td>shoot</td>
<td>ski</td>
</tr>
<tr>
<td>redbird</td>
<td>rooster</td>
<td>satin</td>
<td>separate</td>
<td>shop</td>
<td>skin</td>
</tr>
<tr>
<td>redbreast</td>
<td>root</td>
<td>satisfactory</td>
<td>September</td>
<td>shopping</td>
<td>skip</td>
</tr>
<tr>
<td>refuse</td>
<td>rope</td>
<td>Saturday</td>
<td>servant</td>
<td>shore</td>
<td>skirt</td>
</tr>
<tr>
<td>reindeer</td>
<td>rose</td>
<td>sausage</td>
<td>serve</td>
<td>short</td>
<td>sky</td>
</tr>
<tr>
<td>rejoice</td>
<td>rosebud</td>
<td>savage</td>
<td>service</td>
<td>shot</td>
<td>slam</td>
</tr>
<tr>
<td>remain</td>
<td>rot</td>
<td>save</td>
<td>set</td>
<td>should</td>
<td>slap</td>
</tr>
<tr>
<td>remember</td>
<td>rotten</td>
<td>savings</td>
<td>setting</td>
<td>shoulder</td>
<td>slate</td>
</tr>
<tr>
<td>remind</td>
<td>rough</td>
<td>saw</td>
<td>settle</td>
<td>shouldn't</td>
<td>slave</td>
</tr>
<tr>
<td>remove</td>
<td>round</td>
<td>say</td>
<td>settlement</td>
<td>shout</td>
<td>sled</td>
</tr>
<tr>
<td>rent</td>
<td>route</td>
<td>scab</td>
<td>seven</td>
<td>shovel</td>
<td>sleep (y)</td>
</tr>
<tr>
<td>repair</td>
<td>row</td>
<td>scales</td>
<td>seventeen</td>
<td>show</td>
<td>sleeve</td>
</tr>
<tr>
<td>repay</td>
<td>rowboat</td>
<td>scare</td>
<td>seventh</td>
<td>shower</td>
<td>sleigh</td>
</tr>
<tr>
<td>repeat</td>
<td>royal</td>
<td>scarf</td>
<td>seventy</td>
<td>shut</td>
<td>slept</td>
</tr>
<tr>
<td>report</td>
<td>rub</td>
<td>school</td>
<td>several</td>
<td>shy</td>
<td>slice</td>
</tr>
<tr>
<td>rest</td>
<td>rubbed</td>
<td>schoolboy</td>
<td>sew</td>
<td>sick (ness)</td>
<td>slid</td>
</tr>
<tr>
<td>return</td>
<td>rubber</td>
<td>schoolhouse</td>
<td>shade</td>
<td>side</td>
<td>slide</td>
</tr>
<tr>
<td>review</td>
<td>rubbish</td>
<td>schoolmaster</td>
<td>shadow</td>
<td>sidewalk</td>
<td>sling</td>
</tr>
<tr>
<td>reward</td>
<td>rug</td>
<td>schoolroom</td>
<td>shady</td>
<td>sideways</td>
<td>slip</td>
</tr>
<tr>
<td>rib</td>
<td>rule (r)</td>
<td>scorched</td>
<td>shake (r)</td>
<td>sigh</td>
<td>slipped</td>
</tr>
<tr>
<td>ribbon</td>
<td>rumble</td>
<td>score</td>
<td>shaking</td>
<td>sight</td>
<td>slipper</td>
</tr>
<tr>
<td>rice</td>
<td>run</td>
<td>scrap</td>
<td>shall</td>
<td>sign</td>
<td>slippery</td>
</tr>
<tr>
<td>rich</td>
<td>rung</td>
<td>scrape</td>
<td>shame</td>
<td>silence</td>
<td>slit</td>
</tr>
<tr>
<td>rid</td>
<td>runner</td>
<td>scratch</td>
<td>shan't</td>
<td>silent</td>
<td>slow (ly)</td>
</tr>
<tr>
<td>riddle</td>
<td>running</td>
<td>scream</td>
<td>shape</td>
<td>silk</td>
<td>sly</td>
</tr>
<tr>
<td>ride (r)</td>
<td>rush</td>
<td>screen</td>
<td>share</td>
<td>sill</td>
<td>smack</td>
</tr>
<tr>
<td>riding</td>
<td>rust (y)</td>
<td>screw</td>
<td>sharp</td>
<td>silly</td>
<td>small</td>
</tr>
<tr>
<td>right</td>
<td>rye</td>
<td>scrub</td>
<td>shave</td>
<td>silver</td>
<td>smart</td>
</tr>
<tr>
<td>rim</td>
<td>sack</td>
<td>sea</td>
<td>she</td>
<td>simple</td>
<td>smell</td>
</tr>
<tr>
<td>ring</td>
<td>sad</td>
<td>seal</td>
<td>she'd</td>
<td>sin</td>
<td>smile</td>
</tr>
<tr>
<td>rip</td>
<td>saddle</td>
<td>seam</td>
<td>she'll</td>
<td>since</td>
<td>smoke</td>
</tr>
<tr>
<td>ripe</td>
<td>sadness</td>
<td>search</td>
<td>she's</td>
<td>sing</td>
<td>smooth</td>
</tr>
<tr>
<td>rise</td>
<td>safe</td>
<td>season</td>
<td>shear (s)</td>
<td>singer</td>
<td>snail</td>
</tr>
<tr>
<td>rising</td>
<td>safety</td>
<td>seat</td>
<td>shed</td>
<td>single</td>
<td>snake</td>
</tr>
<tr>
<td>river</td>
<td>said</td>
<td>second</td>
<td>sheep</td>
<td>sink</td>
<td>snap</td>
</tr>
<tr>
<td>road</td>
<td>sail</td>
<td>secret</td>
<td>sheet</td>
<td>sip</td>
<td>snapping</td>
</tr>
<tr>
<td>roadside</td>
<td>sailboat</td>
<td>see (ing)</td>
<td>shelf</td>
<td>sir</td>
<td>sneeze</td>
</tr>
<tr>
<td>roar</td>
<td>sailor</td>
<td>seed</td>
<td>shell</td>
<td>sis</td>
<td>snow (y)</td>
</tr>
<tr>
<td>roast</td>
<td>saint</td>
<td>seek</td>
<td>shepherd</td>
<td>sissey</td>
<td>snowball</td>
</tr>
<tr>
<td>rob</td>
<td>salad</td>
<td>seem</td>
<td>shine</td>
<td>sister</td>
<td>snowflake</td>
</tr>
<tr>
<td>robber</td>
<td>sale</td>
<td>seen</td>
<td>shining</td>
<td>sit ’</td>
<td>snuff</td>
</tr>
<tr>
<td>robe</td>
<td>salt</td>
<td>seesaw</td>
<td>shiny</td>
<td>sitting</td>
<td>snug</td>
</tr>
<tr>
<td>robin</td>
<td>same</td>
<td>select</td>
<td>ship</td>
<td>six</td>
<td>so</td>
</tr>
</tbody>
</table>
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  swallowed  teaspoon  three
soil   spring   stopping  swam    teeth   threw
sold   springtime  store   swamped  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    tiptoe
sour    stay    stuff    tablespoon  them    tiptoe
south (ern)  steak  stump    tablet    then    tire
space   steal    stump    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  steepel    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
| 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 | wrung   |
| 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—The Flesch Formulas

Introduction


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 a 1948 article printed here, “A New Readability Yardstick,” Flesch published a second formula with two parts. 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:

\[
\text{Score} = \text{position on a scale of 0 (difficult) to 100 (easy), with 30 = very difficult and 70 = suitable for adult audiences.}
\]

\[
\text{ASL} = \text{average sentence length (the number of words divided by the number of sentences).}
\]

\[
\text{ASW} = \text{average number of syllables per word (the number of syllables divided by the number of words).}
\]

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:

<table>
<thead>
<tr>
<th>Reading Ease Score</th>
<th>Style Description</th>
<th>Estimated Reading Grade</th>
<th>Estimated Percent of U.S. Adults (1949)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 30:</td>
<td>Very Difficult</td>
<td>College graduate</td>
<td>4.5</td>
</tr>
<tr>
<td>30 to 40:</td>
<td>Difficult</td>
<td>13\textsuperscript{th} to 16\textsuperscript{th} grade</td>
<td>33</td>
</tr>
<tr>
<td>50 to 60:</td>
<td>Fairly Difficult</td>
<td>10\textsuperscript{th} to 12\textsuperscript{th} grade</td>
<td>54</td>
</tr>
<tr>
<td>60 to 70:</td>
<td>Standard</td>
<td>8\textsuperscript{th} and 9\textsuperscript{th} grade</td>
<td>83</td>
</tr>
<tr>
<td>70 to 80:</td>
<td>Fairly Easy</td>
<td>7\textsuperscript{th} grade</td>
<td>88</td>
</tr>
<tr>
<td>80 to 90:</td>
<td>Easy</td>
<td>6\textsuperscript{th} grade</td>
<td>91</td>
</tr>
<tr>
<td>90 to 100:</td>
<td>Very Easy</td>
<td>5\textsuperscript{th} grade</td>
<td>93</td>
</tr>
</tbody>
</table>

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 Flesh’s syllable count. The modified formula is:

\[
\text{New Reading Ease score} = 1.599 \times \text{nosw} - 1.015 \times \text{sl} - 31.517
\]

Where:  
\[
\text{nosw} = \text{number of one-syllable words per 100 words;}
\]
\[
\text{sl} = \text{average sentence length in words}
\]

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 edito-
rial 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:

<table>
<thead>
<tr>
<th>Style</th>
<th>Flesch Reading Ease Score</th>
<th>Average Sentence Length in Words</th>
<th>Average No. of Syll. Per 100 Words</th>
<th>Type of Magazine</th>
<th>Estimated School Grade Completed</th>
<th>Estimated Percent of U.S. Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Easy</td>
<td>90 to 100</td>
<td>8 or less</td>
<td>123 or less</td>
<td>Comics</td>
<td>4th grade</td>
<td>93</td>
</tr>
<tr>
<td>Easy</td>
<td>80 to 90</td>
<td>11</td>
<td>131</td>
<td>Pulp fiction</td>
<td>5th grade</td>
<td>91</td>
</tr>
<tr>
<td>Fairly Easy</td>
<td>70 to 80</td>
<td>14</td>
<td>139</td>
<td>Slick fiction</td>
<td>6th grade</td>
<td>88</td>
</tr>
<tr>
<td>Standard</td>
<td>60 to 70</td>
<td>17</td>
<td>147</td>
<td>Digests</td>
<td>7th or 8th grades</td>
<td>83</td>
</tr>
<tr>
<td>Fairly Difficult</td>
<td>50 to 60</td>
<td>21</td>
<td>155</td>
<td>Quality</td>
<td>Some high school</td>
<td>54</td>
</tr>
<tr>
<td>Difficult</td>
<td>30 to 50</td>
<td>25</td>
<td>167</td>
<td>Academic</td>
<td>High school or some college</td>
<td>33</td>
</tr>
<tr>
<td>Very Difficult</td>
<td>0 to 30</td>
<td>29 or more</td>
<td>192 or more</td>
<td>Scientific</td>
<td>College</td>
<td>4.5</td>
</tr>
</tbody>
</table>

*Table 1. 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. Together, they helped to bring down the reading grade level of front-page stories from the 16th to the 11th grade, where they remain today.

—WHD
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 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 un-i readability") was 5.4 ("difficult"); yet William James' Principles of Psychology, a classic example of readability, rated 6.0

* 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.
(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 of ten-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

<table>
<thead>
<tr>
<th></th>
<th>sl</th>
<th>C_{50}</th>
<th>\bar{X}</th>
<th>s</th>
<th>\beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>wl</td>
<td>.4644</td>
<td>.6648</td>
<td>134.2208</td>
<td>13.6845</td>
<td>.5422</td>
</tr>
<tr>
<td>sl</td>
<td></td>
<td>—</td>
<td>16.5213</td>
<td>5.5509</td>
<td>.2639</td>
</tr>
</tbody>
</table>

* 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

<table>
<thead>
<tr>
<th></th>
<th>ps</th>
<th>C_{50}</th>
<th>\bar{X}</th>
<th>s</th>
<th>\beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>pw</td>
<td>.2268</td>
<td>-.3881</td>
<td>7.34578</td>
<td>5.5175</td>
<td>-.3446</td>
</tr>
<tr>
<td>ps</td>
<td></td>
<td>-.2699</td>
<td>29.5745</td>
<td>35.5822</td>
<td>-.1917</td>
</tr>
</tbody>
</table>

Findings

The intercorrelations, means, standard deviations, and regression weights found are shown in Tables 1, 2, and 3. The following symbols were used: \( w_l \) for word length (syllables per 100 words), \( s_l \) for sentence length in words, \( p_w \) for percentage of "personal words," \( p_s \) for percentage of "personal sentences," \( C_{75} \) 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.
The Classic Readability Studies 1948—The Flesch Formulas

Table 3
Means and Standard Deviations of Two Criteria

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>C50</td>
<td>5.4973</td>
<td>1.3877</td>
</tr>
<tr>
<td>C75</td>
<td>7.3484</td>
<td>2.1345</td>
</tr>
</tbody>
</table>

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 corresponds 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
The Flesch Formulas, 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: $a = .6832 \times wl - 66.6017$. Conversely, the number of syllables per 100 words ($wl$) can be predicted from the number of affixes ($a$) by the formula: $wl = 1.49 \times a + 94.56$.

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 dia-
logue, as tested by the added element, is well described in the follow-
ing, 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 conversations?" thought Alice just before the
White Rabbit ran by, in condemnation of the book her sister was read-
ing, and this childish comment is supported by novel-readers of all
degrees of intelligence. Long close paragraphs of print are in them-
selves 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 de-
\v{r}ives 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).

\begin{table}
\centering
\caption{Comparative Analysis of The New Yorker (October 26, 1946) and the Reader's Digest (November, 1946)}
\begin{tabular}{lll}
\hline
 & New Yorker & Reader's Digest \\
\hline
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 \\
\hline
New Formula A: & & \\
Average sentence length in words & 20 & 16 \\
Syllables per 100 words & 148 & 145 \\
"Reading ease" score & 61 & 68 \\
\hline
New Formula B: & & \\
Personal words per 100 words & 10 & 8 \\
Personal sentences per 100 sentences & 39 & 15 \\
"Human interest" score & 49 & 34 \\
\hline
\end{tabular}
\end{table}

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 dif-
ference 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 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,
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, 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 syllables per 100 words (word length, \(wl\)) and the average sentence length (\(sl\)) in the following formula:

\[
R.E. \text{("reading ease")} = 206.835 - .846 \, \text{wl} - 1.015 \, \text{sl}.
\]

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:

\[
H.L \text{("human interest")} = 3.635 \, pw + 314 \, ps.
\]

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

<table>
<thead>
<tr>
<th>&quot;Reading Ease&quot; Score</th>
<th>Description of Style</th>
<th>Typical Magazine</th>
<th>Syllables per 100 words</th>
<th>Average Sentence Length in Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 20</td>
<td>Very Difficult</td>
<td>Scientific</td>
<td>192 or more</td>
<td>29 or more</td>
</tr>
<tr>
<td>30 to 50</td>
<td>Difficult</td>
<td>Academic</td>
<td>167</td>
<td>25</td>
</tr>
<tr>
<td>50 to 60</td>
<td>Fairly difficult</td>
<td>Quality</td>
<td>155</td>
<td>21</td>
</tr>
<tr>
<td>60 to 70</td>
<td>Standard</td>
<td>Digests</td>
<td>147</td>
<td>17</td>
</tr>
<tr>
<td>70 to 80</td>
<td>Fairly easy</td>
<td>Slick-fiction</td>
<td>139</td>
<td>14</td>
</tr>
<tr>
<td>80 to 90</td>
<td>Easy</td>
<td>Pulp-fiction</td>
<td>131</td>
<td>11</td>
</tr>
<tr>
<td>90 to 100</td>
<td>Very easy</td>
<td>Comics</td>
<td>123 or less</td>
<td>8 or less</td>
</tr>
</tbody>
</table>

### Table 6

Pattern of “Reading Ease” Scores

<table>
<thead>
<tr>
<th>&quot;Human Interest&quot; Score</th>
<th>Description of Style</th>
<th>Typical Magazine</th>
<th>Percentage of Personal Words</th>
<th>Percentage of Personal Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 10</td>
<td>Dull</td>
<td>Scientific</td>
<td>2 or less</td>
<td>0</td>
</tr>
<tr>
<td>10 to 20</td>
<td>Mildly Interesting</td>
<td>Trade</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>20 to 40</td>
<td>Interesting</td>
<td>Digests</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>40 to 60</td>
<td>Highly Interesting</td>
<td>New Yorker</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>60 to 100</td>
<td>Dramatic</td>
<td>Fiction</td>
<td>17 or more</td>
<td>58 or more</td>
</tr>
</tbody>
</table>

### 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
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 pectoris, 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

<table>
<thead>
<tr>
<th></th>
<th>Life (290 words)</th>
<th>New Yorker (495 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Old Formula:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average sentence length in words</td>
<td>23</td>
<td>18</td>
</tr>
<tr>
<td>Affixes per 100 words</td>
<td>48</td>
<td>35</td>
</tr>
<tr>
<td>Personal words per 100 words</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Readability score</td>
<td>5.16</td>
<td>3.20</td>
</tr>
<tr>
<td>New Formula A:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average sentence length in words</td>
<td>22</td>
<td>18</td>
</tr>
<tr>
<td>Syllables per 100 words</td>
<td>165</td>
<td>145</td>
</tr>
<tr>
<td>“Reading ease” score</td>
<td>46</td>
<td>66</td>
</tr>
<tr>
<td>New Formula B:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal words per 100 words</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Personal sentences per 100 sentences</td>
<td>0</td>
<td>41</td>
</tr>
<tr>
<td>“Human interest” score</td>
<td>7</td>
<td>53</td>
</tr>
</tbody>
</table>

Received November 3, 1947.

References

7. Flesch, R. How to write copy that will be read. *Advertising & Selling*, March, 1947. 113ff.

9. Gunning, R. Gunning finds papers too hard to read. *Editor & Publisher,* May 19, 1945, 12.


