The Principles of Readability

William H. DuBay

Impact Information
Costa Mesa, California

Copyright

Abstract

The Principles of Readability gives a brief introduction to the literacy studies in the U.S. and the research on readability and the readability formulas.

Readers' Comments

Please send all comments and suggestions regarding this document to:

William DuBay
Impact Information
Plain-Language Services
Phone: (360) 261-8955
Website: http://www.impact-information.com

Contents

RESEARCH

Summary

Over 80 years of research and testing have contributed to the worldwide use in many languages of the readability formulas.

They help us improve the text on the level of words and sentences, the first causes of reading difficulty.

The principles of readability are in every style manual. Readability formulas are in every word processor. What is missing is the research and theory on which they stand.

The Principles of Readability

William H. DuBay

Introduction

In 1998, traffic accidents caused 46 percent of all accidental deaths of infants and children aged 1 to 14 (National Center for Health Statistics, 2000). One study (Johnston et al. 1994) showed that the single strongest risk factor for injury in a traffic accident is the improper use of child-safety seats. Another study (Kahane 1986) showed that, when correctly used, child safety seats reduce the risk of fatal injury by 71 percent and hospitalization by 67 percent.

To be effective, however, the seats must be installed correctly. Other studies, showed that 79 to 94 percent of car seats are used improperly (National Highway Traffic Safety Administration 1996, Decina and Knoebel 1997, Lane et al. 2000).

Public-health specialists Dr. Mark Wegner and Deborah Girasek (2003) suspected that poor comprehension of the installation instructions might contribute to this problem. They looked into the readability of the instructions and published their findings in the medical journal *Pediatrics*. The story was covered widely in the media.

The authors referred to the National Adult Literacy Study (National Center for Educational Statistics, 1993), which states the average adult in the U.S. reads at the 7th grade level. They also cited experts in health literacy who recommend that materials for the public be written at the fifth or sixth-grade reading level (Doak et al., 1996; Weiss and Coyne, 1997).

Their study found that the average reading level of the 107 instructions they examined was the 10th grade, too difficult for 80 percent adult readers in the U.S. When texts exceed the reading ability of readers, they usually stop reading. The authors did not address the design, completeness, or the organization of the instructions. They did not say that the instructions were badly written. Armed with the SMOG readability formula, they found the instructions were written at the wrong grade level. You can be sure the manufacturers of the car safety seats are scrambling to re-write their instructions.

Guidelines for Readability

In works about technical communication, we are often told how to avoid such problems. For example, JoAnn Hackos and Dawn Stephens in *Standards for Online Communication* (1997) ask us to "conform to accepted style standards." They explain:

> Many experts, through much research, have compiled golden rules of documentation writing. These rules apply regardless of medium:
>
> - Use short, simple, familiar words
> - Avoid jargon.
> - Use culture-and-gender-neutral language.
> - Use correct grammar, punctuation, and spelling.
> - Use simple sentences, active voice, and present tense.
> - Begin instructions in the imperative mode by starting sentences with an action verb.
> - Use simple graphic elements such as bulleted lists and numbered steps to make information visually accessible.
>
> For more suggestions, we recommend referring to one of many excellent books on writing style, especially technical style.

We all know of technical publications that do not follow these guidelines and are read only by a small fraction of the potential readership. One reason may be that the writers are not familiar with the background and research of these guidelines.

This paper looks most carefully at two of the most important elements of communication, the reading skills of the audience and the readability of the text.

The readability formulas

In the 1920s, educators discovered a way to use vocabulary difficulty and sentence length to predict the difficulty level of a text. They embedded this method in readability formulas, which have proven their worth in over 80 years of application.

Progress and research on the formulas was something of a secret until the 1950s. Writers like Rudolf Flesch, George Klare, Edgar Dale, and Jeanne Chall brought the formulas and the research supporting them to the marketplace. The formulas were widely used in journalism, research, health care, law, insurance, and industry. The U.S. military developed its own set of formulas for technical-training materials.

By the 1980s, there were 200 formulas and over a thousand studies published on the readability formulas attesting to their strong theoretical and statistical validity.

Are the readability formulas a problem?

In spite of the success of the readability formulas, they were always the center of controversy. When the "plain language" movement in the 1960s resulted in legislation requiring plain language in public and commercial documents a number of articles attacked the use of readability formulas. They had titles like, "Readability: A Postscript" (Manzo 1970), "Readability: Have we gone too far?" (Maxwell 1978), "Readability is a Four-letter Word" (Selzer 1981), "Why Readability Formulas Fail" (Bruce et al. 1981), "Readability Formulas: Second Looks, Second Thoughts" (Lange 1982), "Readability Formulas: What's the

Use?" (Duffy 1985) and "Last Rites for Readability Formulas in Technical Communication" (Connaster 1999).

Many of the critics were honestly concerned about the limitations of the formulas and some of them offered alternatives such as usability testing. Although the alternatives are useful and even necessary, they fail to do what the formulas do: provide an objective prediction of text difficulty.

Although the concerns of the formula critics have been amply addressed elsewhere (Chall 1984, Benson 1984-1985, Fry 1989b, Dale and Chall 1995, Klare 2000), we will examine them again in some detail, with a special regard for the needs of technical communication.

The purpose of this article is to very briefly review the landmark studies on readability and the controversy regarding the formulas. I will be happy if you learn something of the background of the formulas, what they are good for, and what they are not. That knowledge will give you greater confidence and method in tailoring your text for a specific audience.

What is readability?

Readability is what makes some texts easier to read than others. It is often confused with **legibility**, which concerns typeface and layout.

George Klare (1963) defines readability as "the ease of understanding or comprehension due to the style of writing." This definition focuses on writing style as separate from issues such as content, coherence, and organization. In a similar manner, Gretchen Hargis and her colleagues at IBM (1998) state that readability, the "ease of reading words and sentences," is an attribute of clarity.

The creator of the SMOG readability formula G. Harry McLaughlin (1969) defines readability as: "the degree to which a given class of people find certain reading matter compelling and comprehensible." This definition stresses the interaction between the text and a class of readers of known characteristics such as reading skill, prior knowledge, and motivation.

Edgar Dale and Jeanne Chall's (1949) definition may be the most comprehensive: "The sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimal speed, and find it interesting."

Content

Beginning early in the last century in the U.S., studies of the reading ability of adults and the readability of texts developed in tandem. Our subject matter falls under these headings:

The Adult Literacy Studies These studies discovered great differences in the reading skills of adults in the U.S. and their implications for society.

The Classic Readability Studies This section looks at the early readability studies, which started in the late 19th century and concluded in the 1940s, with the publication of the popular Flesch and Dale-Chall formulas. During this period, publishers, educators, and teachers were concerned with finding practical methods to match texts to the skills of readers, both students and adults.

The New Readability Studies Beginning in the 1950s, new developments transformed the study of readability, including a new test of reading comprehension and the contributions of linguistics and cognitive psychology. Researchers explored how the reader's interest,

motivation, and prior knowledge affect readability. These studies in turn stimulated the creation of new and more accurate formulas.

The Adult Literacy Studies

Grading the reading skills of students

Before the mid-19th century, schools in the U.S. did not group students according to grade. Students learned from books that their families had, often Bibles and hornbooks. American educator Horace Mann, who had studied the supervision, graded classes, and well-articulated classes of Prussian schools, struggled to bring those reforms to America.

It was not until 1847 that the first graded school opened in Boston with a series of books prepared for each grade. Educators found that students learn reading in steps, and they learn best with materials written for their current reading level. Since then, grouping by grades has functioned as an instructional process that continues from the first year of school through high school and beyond.

Although reading standards were set for each grade, we know that not all students in the same class read at the same level. A 7th-grade teacher, for example can typically face a classroom of students with reading ability from the 2nd to the 12th grade. Good teaching practice has long separated students in the same class by reading ability for separate instruction (Betts 1946, Barr and Dreeben 1984).

Educators promoted the target reading levels for each class with the use of **standardized reading tests**. William A. McCall and Lelah Crabbs (1926) of the Teachers College of Columbia University published *Standard Test Lessons in Reading*. Revised in 1950, 1961, and 1979, these tests became an important measure of the reading ability of students in the U.S. These and later reading tests typically measure comprehension by having students first read a passage and then answer multiple-choice questions.

The Mc Call-Crabbs reading tests also became important in the development and validation of the readability formulas. Later reading tests also used for creating and testing formulas for adults and children include the Gates-MacGinitie Reading Tests, the Stanford Diagnostic Reading Test, the California Reading Achievement Test, the Nelson-Denny Reading Test, the Diagnostic Assessment of Reading with Trial Teaching Strategies and the National Assessment of Educational Progress (NAEP).

Grading adult readers

For a long time, no one thought of grading adults, who were considered either literate or illiterate. This began to change with the first systematic testing of adults in the U.S. military in 1917. The testing of civilians began in Chicago in 1937.

During that first period, investigators discovered that general readers in the U. S. were adults of limited reading ability. The average adult was able to read with pleasure nothing but the simplest adult materials, usually cheap fiction or graphically presented news of the day.

Educators, corporations, and government agencies responded by providing more materials at different reading levels for adults.

U.S. military literacy surveys—reading on the job

General George Washington first addressed concerns about the reading skills of fighters during the Revolutionary War. He directed chaplains at Valley Forge to teach basic skills of reading, writing, and arithmetic to soldiers. Since then, the

U.S. armed services has invested more in studying workplace literacy than any other organization.

Since the 50s in the U.S., you have to pass a literacy test to join the Armed Services. From such a test and others, the military learns a lot about your aptitudes, cognitive skills, and ability to perform on the job.

It took a while for the military to develop these tests. Over the years, it changed the content of the tests and what they measure. Testing literacy advanced in these general stages:

1. During World War I, they focused on testing **native intelligence**.
2. The military decided that what they were testing was not so much raw intelligence as reading skills. By World War II, they were focusing on classifying **general learning ability** for job placement.
3. In the 1950s, Congress mandated a literacy requirement for all the armed services. The resulting **Armed Forces Qualification Test** (**AFQT**) prevented people of the lowest 10% of reading ability from entering military service. The military then combined AFQT subtest with other tests, which differed for each service and sorted recruits into different jobs.
4. In 1976, with the arrival of the All-Volunteer Force, the military introduced the **Armed Services Vocational Aptitude Battery** (**ASVAB**). All military services used this test battery for both screening qualified candidates and assessing trainability for classified jobs.
5. In 1978, an error resulted in the recruitment of more than 200,000 candidates in the lowest 10% category. The military, with the aid of Congress, decided to keep them. The four military services each created workplace literacy programs, with contract and student costs over $70 million. This was a greater enrollment in adult basic education than in all such programs of 25 states combined. The results of the workplace literacy programs were considered highly successful, with performance and promotions “almost normal.”
6. In 1980, the military further launched the largest study ever in job literacy, the **Job Performance Measurement/Enlistment Standards Project**. They invested $36 million in developing measures of job performance. Over ten years, the project involved more than 15,000 troops from all four military services. Dozens of professionals in psychological measurement took part in this study.
7. In 1991, based on these findings, the military raised its standards and combined the ASVAB with the AFQT and special aptitude tests from all the services into one battery of 10 tests. Both the Army and Navy continue to provide workplace-literacy programs for entering recruits and for upgrading the literacy skills of experienced personnel (Sticht 1995, pp 37-38).

The major findings of the military research were:

1. Measures of literacy correlate closely with measures of intelligence and aptitude.
2. Measures of literacy correlate closely with the breadth of one’s knowledge.
3. Measures of literacy correlate closely to job performance. Hundreds of military studies found no gap between literacy and job performance.
4. Workplace literacy programs are highly effective in producing, in a brief period, significant improvements in job-related reading.

5. Advanced readers have vast bodies of knowledge and perform well across a large set of domains of knowledge. Poor readers perform poorly across these domains of knowledge. This means that, if programs of adult literacy are to move students to high levels of literacy, they must help them explore and learn across a wide range of knowledge (Sticht and Armstrong 1994, pp. 37-38).

The military studies indicated that achieving high levels of literacy requires continued opportunities for life-long learning. Investments in adult literacy provide a unique and cost-effective strategy for improving the economy, the home, the community, and the schools.

U.S. civilian literacy surveys

University of Chicago Study Guy Buswell (1937) of the University of Chicago surveyed 1,000 adults in Chicago with different levels of education. He measured skills in reading materials such as food ads, telephone directories, and movie ads. He also used more traditional tests of comprehension of paragraphs and vocabulary.

Buswell found that reading skills and practices increase as years of education increase**.** He suggested that an important role of education is to guide readers to read more, and that reading more leads to greater reading skill. In turn, this may lead one to continue more education, thus leading to greater reading skill.

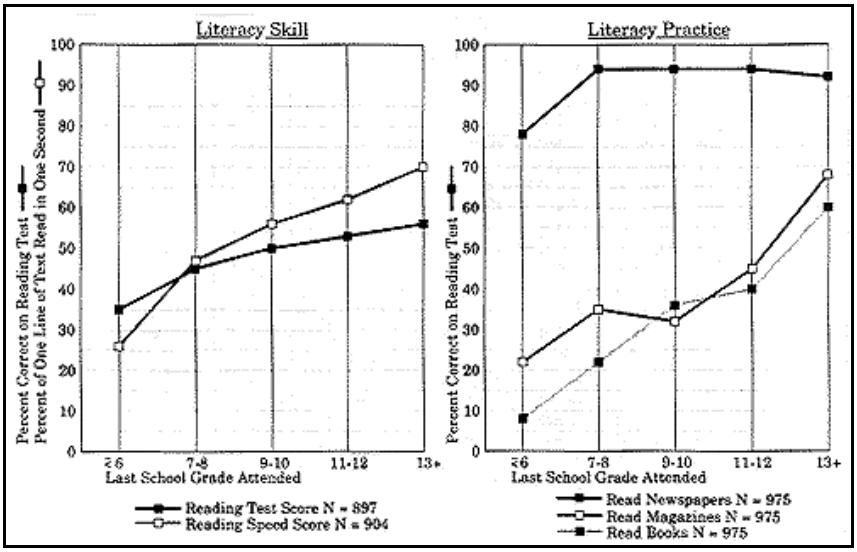


Fig. 1. Adult literacy in 1937. This study confirmed the relationship between reading skill and years of education completed. Sources: Buswell, G. 1937 pp. 27, 57, 71).

The National Assessment of Educational Progress (NAEP) of 1970-1971 This study tested how students 9, 13, and 17 years old as well adults 26 to 35 years old perform on 21 different tasks. The results showed for the first time how age affects performance on the same items. This survey showed as children grow up, attend school, and become adults, they grow progressively more literate (Sticht and Armstrong, pp. 51-58).

Louis Harris survey of 1970 The Louis Harris polling organization surveyed adults representing a cross section of the U.S. population. The subjects filled out five common application forms, including an application for a driver's license and a Medicaid application.

The poll was the first of many to show that many U.S. citizens have difficulty with filling out forms. The Medicaid form was difficult, with only 54 percent of those with an 8th grade education or less getting 90-100 percent correct. Even

many college-educated adults had trouble completing the Medicaid form (Sticht and Armstrong, pp. 59-62).

Adult Functional Reading Study of 1973 This study used household interviews to find out the literacy practices of adults. It used a second household sample to assess literacy skills.

Over all 170 items used in the study, over 70 percent of the respondents scored 70 percent correct or better. As a trend, adults with more education performed better on the test than those with less.

As with Buswell's study, both literacy skills and literacy practices correlated closely with education. Book and magazine reading correlated more closely with years of education than did newspaper reading. Altogether, the adults reported that they spent about 90 minutes a day in reading materials such as forms, labels, signs, bills, and mail. (Sticht and Armstrong, pp. 63-66).

Adult Performance Level Study of 1971 This study began as a project funded by the U. S. Office of Education. It introduced "competency-based" education, directing adult education to focus on achieving measurable outcomes. By 1977, two-thirds of the states had set up some form of "competency-based" adult basic education.

The test included over 40 common and practical tasks, such as filling out a check, reading the want ads, addressing an envelope, comparing advertised products, filling out items on a 1040 tax form, reading a tax table, and filling out a Social Security application. Results showed the high correlation between performance on all tasks and literacy (Sticht and Armstrong, pp. 67-98).

What a Reading Grade Level Means

The reading grade level assigned to a text depends on the use of the text. If the text is used for independent, unassisted, or recreational use, the reading grade level will be higher than a text destined for classroom use and optimum learning gain. In other words, the same text will be easier for those with more advanced reading skills (with a higher grade level) and harder for those with less (and with a lower grade level). See the “Problem of Optimal Difficulty” below.

The grade of completed education is no indication of one’s reading level. Average high-school graduates read at the 9th-grade level, which means a large number reads below that level. Those who pursue special domains of knowledge may develop higher levels of reading skill in those specialties than they have for general reading. Thus, college graduates, who prefer to read general materials at the 10th-grade level, may prefer more difficult texts within their own specialty. Students who are poor readers of general classroom material are often able to master difficult treatments of subjects that appeal to them.

Young Adult Literacy Survey of 1985 This study of young adults (17-25) and the adult study that followed in 1992 both measured the literacy the same way in three areas:

- Prose literacy—meaning of selected texts
- Document literacy—finding information on a form such as a bus schedule.
- Quantitive literacy—mathematical and spatial tasks.

Both studies used a literacy scoring range of 1 to 500 and the five levels of skill defined by the National Assessment of Educational Progress (1985). John Carroll (1987) estimated the corresponding reading-grade levels as shown below in Table 1.

NAEP Level	Literacy Score	Grade Level
I Rudimentary	150	1.5
II Basic	200	3.6
III Intermediate	250	7.2
IV Adept	300	12
V Advanced	350	16+

Table 1. NAEP proficiency levels and the reading-grade-level equivalents.

The young adult survey by the NAEP (1985) found that only 40 percent of young adults 17 to 25 no longer in high school, and 17 years old and in high school, read at a 12th-grade level. Large numbers leave high school still reading at the 8th-grade level or lower. The 1990 census showed that 24.8 percent of adults did not graduate from high school.

The National Adult Literacy Survey (NALS) of 1992 This U.S. Government study sampled 26,000 adults, representing 191 million adults. In 1993, it published the first of a number of reports on this survey entitled, "Adult Literacy in America" (National Center for Education Statistics 1993, 1999, 2001).

This study used the same tests as the Young Adult Literacy Survey and reported data with the same five levels of skill.

Literacy Skill	Level 1	Level 2	Level 3	Level 4	Level 5
Prose	21%	27%	32%	17%	3%
Document	23%	28%	31%	15%	3%
Quantitative	22%	25%	31%	17%	4%

Table 2. Percentages of adults in the U.S. in each of the five NAEP skill levels for each literacy skill (Sticht and Armstrong 1995, p. 113).

The data in this table suggest 40 to 44 million adults in the U.S. are in Level 1, defined as "functionally illiterate, not having enough reading skills for daily life." Some 50 million are in Level 2. This means the percentage of adults who struggle at Levels 1 and 2 (below the 5th-grade level) in the U.S. reaches 48 percent.

The report confirmed that numeracy (quantitative) skills increase with reading skills. Adults of different reading skills not only have different worldviews but also different life experiences. Forty-three percent of adults with low-literacy skills live in poverty, 17% receive food stamps, and 70% have no job or part-time job. Over 60 % of frontline workers producing goods have difficulty applying information from a text to a task. More than 20% of adults read below the sixth-grade level, far below the level needed to earn a living wage.

Adults at Level 1 earned a median income of $240 a week, while those at Level 5 earned $681. Seventy percent of prisoners are in the lowest two levels.

In support of these figures, the number of companies reporting shortages of skilled workers doubled between 1995 and 1998. Ninety percent of Fortune 1000 executives reported that low literacy is hurting productivity and profitability. In one survey, more than half of the responding company representatives said that high school graduates applying for jobs are not literate enough to hire.

Low levels of literacy have caused costly and dangerous mistakes in the workplace. There are other costs in billions of dollars in the workplace resulting

from low productivity, poor quality of products and services, mistakes, absenteeism, and lost management time.

The Adult Literacy Survey also confirmed the effects of literacy on health care. Since 1974, when health officials became aware of the effects of low literacy on health, literacy problems have grown. A more complex health-care system requires better reading skills to negotiate the system and take more responsibility for self-care.

Using a nationally representative sample of the U.S. adult population age 16 and older, the National Academy (2002) on an Aging Society examined the impact of literacy on the use of health care services. The study found that people with low health-literacy skills use more health care services.

Among adults who stayed overnight in a hospital in 1994, those with low health literacy skills averaged 6 percent more hospital visits, and stayed in the hospital nearly 2 days longer than adults with higher health literacy skills. The added health-care costs of low literacy are estimated at $73 billion in 1998 dollars. This includes $30 billion for the Level 2 population plus $43 billion for the Level 1 population. The total is about what Medicare pays for doctor services, dental services, home health care, prescription drugs, and nursing-home care combined.

Low literacy is not chiefly the problem of immigrants, the elderly, high school dropouts, or people whose first language is not English. Low literacy is a problem that knows no age, education, income levels, or national origins. Most people with low literacy skills were born in this country and have English as their first language.

One solution to the problem of low literacy of adults is more government and corporate support for adult literacy programs. Workplace literacy programs have cost-effective and lasting results. Another solution is to produce more texts that are written for people of diverse reading skills.

Challenges for technical communicators

The lessons of the literacy studies for technical communicators are obvious:

- Low and intermediate literacy skills are a big problem for large numbers of users of technical documents. Providing technical documents at their levels will advance both their technical and reading skills.
- The larger the audience, the more it will include the average reading habits and skills of the public as determined by the literacy surveys.
- The more critical the information is for safety and health, the greater is the need for increased readability.

The finding that the great majority of adult readers are mid-range, intermediate readers brings to us in technical communication new opportunities and challenges.

Intermediate readers represent a large audience that technical documents have been missing. Go into any library or bookstore, and you will find few technical or scientific publications in the "Young Adult" section, or elsewhere written at the 7th to 9th-grade level. On the Internet, there is the same scarcity of intermediate technical materials.

For example, a small sampling of the author's shows that the support sections of the Apple and Microsoft Web sites are written at advanced level of 10th grade and up. The technical books for *Dummies* and *Idiots*, while written in a casual style, are often at the 10th-grade level and up. Like the car-safety seat instructions, these technical documents are too difficult for 80 percent of adult

readers in the U.S. Ironically, the user manual that comes with the CorelDraw program is written at the 7th-grade level, making it fit for a much larger audience than its *Dummies* counterpart.

Considering the keen interest that intermediate readers of all ages can have in technical matters, this literacy gap is troubling. While some highly motivated readers are able to master difficult technical materials, we cannot assume that everyone will do so. To the contrary, the difficulty of technical materials has taught many if not most readers of intermediate skill not to look for technical help in written texts. Helpful text means not only providing readers accurate information but also information written at the reading levels they need.

The Classic Readability Studies

The first aim of the classic readability studies was to develop practical methods to match reading materials with the abilities of students and adults. These efforts centered on making easily applied readability formulas which teachers and librarians could use.

The first adult literacy surveys in the U.S. in the 1930s brought new concerns about providing graded texts for adults. For the rest of the century, publishers, librarians, teachers, and investigators addressed that need with new methods of determining the reading level of texts.

The classic readability studies include these landmark issues:

- L. A. Sherman and the statistical analysis of literature.
- The vocabulary-frequency lists
- The classic readability formulas

L. A. Sherman and the statistical analysis of literature

Down through the centuries, many had written about the differences between an "ornate" and "plain" style in English.

In 1880, a professor of English Literature at the University of Nebraska, Lucius Adelno Sherman, began to teach literature from a historical and statistical point of view.

He compared the older prose writers with more popular modern writers such as Macaulay (*The History of England*) and Ralph Waldo Emerson. He 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 (1893), *Analytics of Literature*, *A Manual for the Objective Study of English Prose and Poetry*, 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.

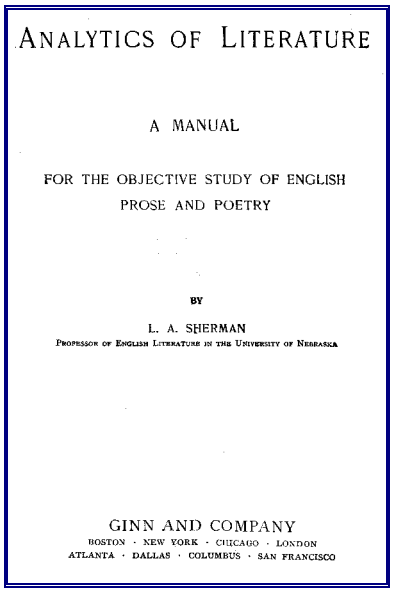
Sherman was the first to use statistical analysis for the task of analyzing readability, introducing a new and objective method of literary criticism. Another of Sherman's discoveries was that over time sentences not only became shorter but also simpler and less abstract. He believed this process was due to the influence of the spoken language on written English. He wrote (p. 312):

> 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.

Linguistic research later confirmed Sherman's view of the relationship between spoken and written language. Rudolf Flesch (1946) wrote that English is following written Chinese in making language simpler by substituting standard word order (subject-verb-object) for more complex grammar.

According to Flesch, Chinese is "the most grown-up talk of mankind. It is the way people speak who started to simplify their language thousands of years ago and have kept at it ever since…. (p. 12).

"Among the world's great languages, the runner-up to Chinese is English. It's simpler, more flexible, more practical than any other Western language because it has gone furthest in losing inflections and straightening out irregularities" (p. 20)


ANALYTICS OF LITERATURE

A MANUAL

FOR THE OBJECTIVE STUDY OF ENGLISH PROSE AND POETRY

BY

L. A. SHERMAN

PROFESSOR OF ENGLISH LITERATURE IN THE UNIVERSITY OF NEBRASKA

GINN AND COMPANY

BOSTON · NEW YORK · CHICAGO · LONDON

ATLANTA · DALLAS · COLUMBUS · SAN FRANCISCO

Fig.2. In Analytics of Literature*, L.A. Sherman looked at literature statistically. He 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).

Vocabulary-frequency lists

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* (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.

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.

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 (1995, p. 84) write, "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* (Johnson, 1946). The first 100 most frequent words make up almost half of all written material. The first 300 words make up about 65 percent of it (Fry et al, 1993).

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*.

Until computers came along, educators, publishers, and teachers commonly used word-frequency lists to evaluate reading materials for their classes. Thorndike's work also was the basis for the first readability formulas for children's books.

After Thorndike, there was extensive research on vocabulary. The high mark came in *Human Behavior and The Principle of Least Effort* by Harvard's George Kingsley Zipf (1949).

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.

Klare (1968), reviewing the research on word frequency, concludes: "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."

Dale and O'Rourke: the words Americans know In 1981, publishers of the *World Book Encyclopedia* published *The Living Word Vocabulary: A National Vocabulary Inventory* by Edgar Dale and Joseph O'Rourke. The authors based this work on the earlier work of Thorndike and others as well as on a 25-year study of their own. It contained the grade-level scores of the familiarity of 44,000 words. For the first time, it gave scores for each of the meanings a word

can have and the percentage of readers in the specified grade who are familiar with the word.

The authors obtained the familiarity scores by giving a three-choice test to students from the 4th to the 16th grade in schools and colleges throughout the U.S. The editors of the encyclopedia also used the scores to test the readability of the articles they published. Field tests of the encyclopedia later confirmed the validity of the word scores. This work is exceptional in every respect and is considered by many to be the best aid in writing for a targeted grade level.

Grade	Score	Word — Word Meaning
16	78%	**abruption** — a sudden breaking off
08	71%	**abscess** — wound with pus
12	31%	**abscind** — to cut apart
16	72%	**abscissa** — horizontal coordinate
16	84%	**abscond** — run away and hide
04	67%	**absence** — being away
06	91%	**absence** — not having something
04	84%	**absent** — not here

Fig. 3. Sample entries from The Living Word Vocabulary. *This work featured not only grade level and a short definition, but also the percentage of readers in that grade who know the word. The editors of* World Book Encyclopedia *used this information as one of the reading-level tests for their entries (Dale and O'Rourke 1981).*

In the preface, the Editorial Director of the encyclopedia W. H. Nault wrote (p. v) that this work marked "the beginning of a revolutionary approach to the preparation and presentation of materials that fit not only the reading abilities, but the experience and background of the reader as well."

Although this work is out of print, you can find it at libraries and used bookshops along with other graded vocabularies and word-frequency lists such as *The American Heritage Word Frequency Book.*

The classic readability formulas

Harry D. Kitson—Different readers, different styles Psychologist Harry D. Kitson (1921) published *The Mind of the Buyer*, in which he showed how and why readers of different magazines and newspapers differed from one another. Although he was not aware of Sherman's work, he found that sentence length and word length measured in syllables are important measures of readability. Rudolph Flesch would incorporate both these variables in his Reading Ease formula 30 years later.

Although Kitson did not create a readability formula, he showed how his principles worked in analyzing two newspapers, the *Chicago Evening Post* and the *Chicago American* and two magazines, the *Century* and the *American*. He analyzed 5000 consecutive words and 8000 consecutive sentences in the four publications. His study showed that the average word and sentence length were shorter in the *Chicago American* newspaper than in the *Post*, and the *American* magazine's style simpler than the *Century*'s, accounting for the differences in their readership.

The first readability formula Bertha A. Lively and Sidney L. Pressey (1923) were concerned with the practical problem of selecting science textbooks for

junior high school. The books were so overlaid with technical words that teachers spent all class time teaching vocabulary. They argued that it would be helpful to have a way to measure and reduce the "vocabulary burden" of textbooks.

Their article featured the first children's readability formula. In each count of a thousand words, it measured the number of different words, the number of words not on the Thorndike list of 10,000 words, and the median index number of the words found in the Thorndike list of 10,000 words.

They tested their formula on 11 textbooks of different difficulties, along with one newspaper. At the low end, there were a second and a fourth-grade reader and Stevenson's *Kidnapped*. At the high end, there was 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 had a great influence on the readability studies that would shortly follow.

Other early school formulas Mabel Vogel and Carleton Washburne (1928) of Winnetka, Illinois carried out one of the most important studies of readability. 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.

Vogel and Washburne's work stimulated the interest of Alfred S. Lewerenz (1929, 1929a, 1935, 1939), who produced several new readability formulas for the Los Angeles School District.

W. W. Patty and W. I. Painter (1931) discovered the year of highest burden in high school is the sophomore year. They also developed a formula to measure the relative difficulty of textbooks based on a combination of frequency as determined by the Thorndike list and vocabulary diversity (the number of different words in a text).

With the rise of the plain-language movement in the 1960s, several critics of the formulas claimed that the formulas do not test comprehensibility (Kern 1979, Duffy and Kabance 1981, Duffy 1985). The history of the formulas, however, shows that from the beginning their scores correlate well with comprehension difficulty as measured by reading tests. The formulas rate very well when compared with other widely used psychometric measurements such as reading tests (Chall and Dale 1995). Their validity correlations make them useful for predicting the comprehension difficulty of texts (Bormuth 1966).

Correlations with difficulty 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 9^{th}-grade level of difficulty on a 7^{th}-grade text, and, if at all grade levels, the error is in the same direction and by a corresponding amount, the correlation could be +1.00 or at least quite high. If, on the other hand, a formula predicts a 9^{th}-grade level for a 6^{th}-grade text, an 8^{th} grade level for a 10^{th}-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 Lively and Pressey formula above accounts for 64% ($.80^2$) of the variance of the text difficulty.

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

Douglas Waples and Ralph W. Tyler (1931) 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 (1934) did not invent a formula, but 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 (1934) their own readability formula and the first

study on adult readability formulas. 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 7^{th} to 9^{th} 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.

Gray and Leary: what makes a book readable William S. Gray and Bernice Leary (1935) published a landmark work in reading research, *What Makes a Book Readable*. Like Dale and Tyler's work, it attempted to discover what makes a book readable for adults of limited reading ability.

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 Figure 4).

TABLE I

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

Major Category	All Persons			Librarians			Publishers			Others Interested in Adult Education		
	M.	σ	Range	M.	σ	Range	M.	σ	Range	M.	σ	Range
I. Format	20.26	7.68	45–5	24.13	7.64	45–7	17.08	5.20	25–10	17.92	6.27	30–5
II. General Features of Organization	15.38	7.04	40–3	15.71	6.59	26–3	15.42	6.27	30–5	15.20	7.24	40–5
III. Style of Expression and Presentation	30.71	9.17	50–0	32.74	8.39	50–20	32.92	8.27	50–25	27.50	9.24	40–5
IV. Content	33.64	13.11	75–7	27.42	9.95	50–7	34.58	12.83	50–10	39.37	12.54	75–20

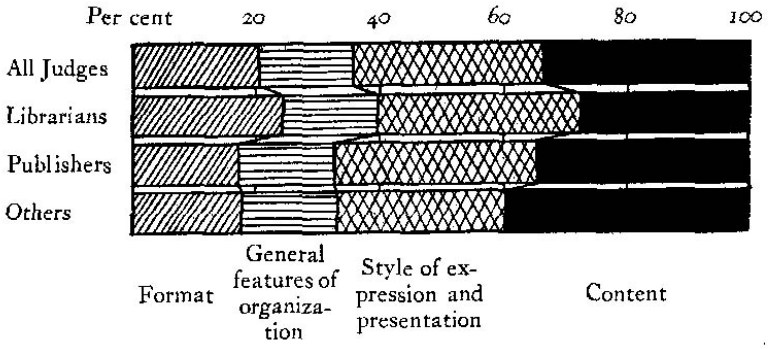


FIG. 1.—Opinion concerning the influence of classified factors on readability

Fig 4. The four major categories of readability (Gray and Leary, p. 31).

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.

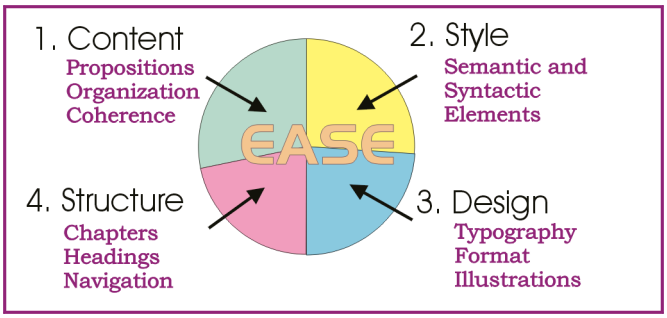


Fig. 5. The four basic elements of reading ease.

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 relationship.

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

1. Average sentence length in words: -.52 (a negative correlation, that is, the longer the sentence the more difficult it is).
2. Percentage of easy words: .52 (the larger the number of easy words the easier the material).
3. Number of words not known to 90% of sixth-grade students: -.51
4. Number of "easy" words: .51
5. Number of different "hard" words: -.50
6. Minimum syllabic sentence length: -.49
7. Number of explicit sentences: .48
8. Number of first, second, and third-person pronouns: .48
9. Maximum syllabic sentence length, -.47
10. Average sentence length in syllables, -.47
11. Percentage of monosyllables: .43
12. Number of sentences per paragraph: .43
13. Percentage of different words not known to 90% of sixth-grade students: -.40
14. Number of simple sentences: .39
15. Percentage of different words: -.38
16. Percentage of polysyllables: -.38
17. Number of prepositional phrases: -35

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

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

measure and apply. Later formulas that use fewer variables may have higher correlations.

Gray and Leary's work stimulated an enormous effort to find the perfect formula, using different combinations of the style variables. In 1954, Klare and Buck listed 25 formulas for children and another 14 for adult readers. By 1981, Klare noted there were over 200 published formulas.

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

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

Formula limitations Readability researchers have long taken pains to recommend that, because of their limitations, formulas are best used in conjunction with other methods of grading and writing texts. Ojemann (1934) warned that the formulas are not to be applied mechanically, a caution expressed throughout readability literature. Other investigators concerned with the difficulty and density of concepts were Morriss and Holversen (1938) and Dolch (1939). E. Horn (1937) warned against the mechanical use of the word lists in the re-writing of books for social studies.

George Klare and colleagues (1969) stated, "For these reasons, formula scores are better thought of as rough guides than as highly accurate values. Used as rough guides, however, scores derived from readability formulas provide quick, easy help in the analysis and placement of educational material."

Readability researchers such as Flesch (1949, 1964, 1979), Klare and Buck (1954), Klare (1980), Gunning (1952), Dale (1967), Gilliland (1972), and Fry (1988) wrote extensively on the other rhetorical factors that require attention such as organization, content, coherence, and design. Using the formulas creatively along with techniques of good writing results in greater comprehension by an audience of a specified reading ability (Klare 1976, Chall and Conard 1991).

Irving Lorge: Consolidating the research Irving Lorge (1938) published *The Semantic Count of the 570 Commonest English Words*, a frequency count of the meaning of words rather than the words themselves. He was co-author with E. L. Thorndike's of Thorndike's last book, *The Teacher's Word Book of 30,000 Words* (1944).

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.

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." 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.

In 1948, Lorge published corrections to his 1939 article and the formulas that were based on those findings.

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

Rudolf Flesch and the art of plain writing The one perhaps most responsible for publicizing the need for readability was Rudolf Flesch, a colleague of Lorge at Columbia University. Besides working as a readability consultant, lecturer, and teacher of writing, he published a number of studies and nearly 20 popular books on English usage and readability. His best-selling books included *The Art of Plain Talk* (1946), *The Art of Readable Writing* (1949), *The Art of Clear Thinking* (1951), *Why Johnny Can't Read —And What You Can Do About It* (1955), *The ABC of Style: A Guide to Plain English* (1964), *How to Write in Plain English: A Book for Lawyers and Consumers* (1979).

Flesch was born in Austria and got a degree in law from the University of Vienna in 1933. He practiced law until 1938, when he came to the U.S. as a refugee from the Nazis.

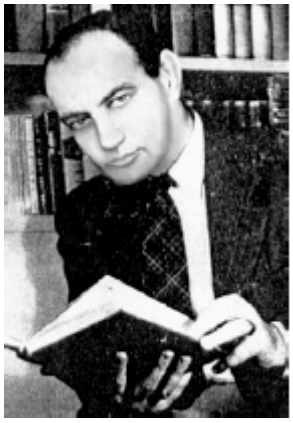


Fig. 6. Rudolf Flesch. The first edition of The Art of Plain Talk in 1946 was a best seller. The readability formulas it featured started a revolution in journalism and business communication.

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 1948, 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:

Score = 206.835 – (1.015 x ASL) – (84.6 x ASW)

Where:

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

ASL = average sentence length (the number of words divided by the number of sentences).

ASW = average number of syllables per word (the number of syllables divided by the number of words).

This formula correlates .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 (1949, p. 149), described his Reading Ease scale in this way:

Reading Ease Score	Style Description	Estimated Reading Grade	Estimated Percent of U.S. Adults (1949)
0 to 30:	Very Difficult	College graduate	4.5
30 to 40:	Difficult	13th to 16th grade	33
50 to 60:	Fairly Difficult	10th to 12th grade	54
60 to 70:	Standard	8th and 9th grade	83

70 to 80:	Fairly Easy	7th grade	88
80 to 90:	Easy	6th grade	91
90 to 100:	Very Easy	5th grade	93

Table 3. Flesch's Reading Ease Scores

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:

New Reading Ease score = 1.599 nosw – 1.015 sl – 31.517

Where: nosw = number of one-syllable words per 100 words;

sl = average sentence length in words

This formula correlates better than .90 with the original Flesch Reading Ease formula and .70 with 75% comprehension of 100-word samplings of the McCall-Crabbs reading lessons. In 1976, a study commissioned by the U.S. Navy modified the Reading Ease formula to produce a grade-level score, This popular formula is known as the Flesch-Kincaid formula, the Flesch Grade-Scale formula or the Kincaid formula (See "The Navy Readability Indexes" below).

In 1949, Flesch published the results of a 10-year study of the editorial content of several magazines. He found that:

- About 45% of the population can read *The Saturday Evening Post.*
- Nearly 50% of the population can read *McCall's*, *Ladies Home Journal*, and *Woman's Home Companion*.
- Slightly over 50% can read *American Magazine.*
- 80% of the population can read *Modern Screen*, *Photoplay*, and three confession magazines.

Flesch (1949, pp. 149-150) compared the reading scores of popular magazines with other variables:

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

Table 4. Flesch's1949 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.

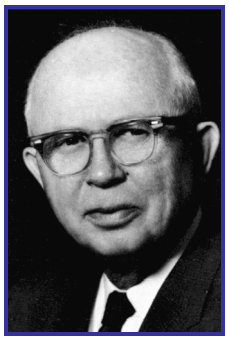


Fig. 6. Edgar Dale, a leading figure in communications, stressed the importance of vocabulary in assessing readability.

The Dale and Chall Original Formula Edgar Dale, for 25 years a professor of education at Ohio State University, 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 lists. He claimed it failed to measure the familiarity of words accurately. He subsequently developed new lists that were later used in readability formulas.

One of these was a formula he developed with Jeanne Chall, the founder and director for 20 years of the Harvard Reading Laboratory. She had 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.*

The original Dale-Chall formula (1948) was developed for adults and children above the 4th grade. They designed it to correct certain shortcomings in the Flesch Reading Ease formula. It uses a sentence-length variable plus a percentage of "hard words"–words not found on the Dale-Chall "long list" of 3,000 easy words, 80 percent of which are known to fourth-grade readers.

To apply the formula:

1. Select 100-word samples throughout the text (for books, every tenth page is recommended).
2. Compute the average sentence length in words.
3. Compute the percentage of words outside the Dale list of 3,000 words.
4. Compute this equation:

Score = .1579PDW + .0496ASL + 3.6365

Where: Score = reading grade of a reader who can answer one-half of the test questions on a passage.

PDW= Percentage of Difficult Words (words not on the Dale-Chall word list)

ASL = Average Sentence Length in words.

Dale and Chall also published the following chart for correcting the grade-level scores at the higher grades.

Formula Score	Corrected Grade Levels
4.9 and below	Grade 4 and below
5.0 to 5.9	Grades 5-6
6.0 to 6.9	Grades 7-8
7.0 to 7.9	Grades 9-10
8.0 to 8.9	Grades 11-12
9.0 to 9.9	Grades 13-15 (college)
10 and above	Grades 16 and above (college graduate)

Table 5. Dale-Call grade-correction chart.

Of all the formulas produced in the early classic period, validations of this formula have produced the most consistent, as well as some of the highest correlations. It correlated .70 with the multiple-choice test scores on the McCall-Crabbs reading lessons. You can find a computerized version of this original formula online at:

http://www.interventioncentral.org/htmdocs/tools/okapi/okapi.shtml

Those interested in manually applying this formula can find the original 1948 Dale-Chall easy word list online at:

http://www.interventioncentral.org/htmdocs/tools/okapi/okapimanual/dalechalllist.shtml

Fig. 8. Robert Gunning.

Robert Gunning and the technique of clear writing Robert Gunning was a graduate of Ohio State University. In 1935, he entered the field of textbook publishing. In the mid-1930s, educators were beginning to see high school graduates who were not able to read. Gunning realized that much of the reading problem was a writing problem. He found that newspapers and business were full of "fog" and unnecessary complexity.

Gunning was among the first to take the new readability research into the workplace. In 1944, he founded the first consulting firm specializing in readability. During the next few years, he tested and worked with more than 60 large city daily newspapers and the popular magazines, helping writers and editors write to their audience.

In *The Technique of Clear Writing*, Gunning (1952) published his own readability formula developed for adults, the **Fog Index**, which became popular because of its ease of use. It uses two variables, average sentence length and the number of words with more than two syllables for each 100 words.

Grade Level = .4 (average sentence length + hard words)

Where:

Hard words = number of words of more than two syllables

Gunning developed his formula using a 90% correct-score with the McCall-Crabbs reading tests. This gives the formula a higher grade criterion than other formulas except for McLaughlin's SMOG formula, which is based on a 100% correct-answer criterion. The grade-level scores predicted by these two formulas tend to be higher than other formulas.

Gunning found that popular magazines were consistent in their reading levels over time. He published these correlations between reading levels of different classes of magazines and their total circulation (p. 35). See Table 6.

Group	Approx. Total Circulation	Average Sentence Length	Percentage of Hard Words	Total	Fog Index
Class	Fewer than 1 million	20	10	30	12
News	About 3 million	16	10	26	10
Reader's Digest	8 million	15	7	22	9
Slicks	More than 10 million	15	5	20	8
Pulps	More than 10 million	15	3	16	6

Table 6. Gunning's analysis of the readability of adult reading materials.

The validation of the original Fog formula has never been published. According to this author's calculations, however, it correlates .93 with the normed reading texts of Chall et al. (1996), a figure which may account for its popularity.

Sumner, and Kearl (1958) recalculated the Fog formula using the McCall-Crabbs reading lessons. The recalculated Fog formula, shown here, correlates .59 with the reading passages.

> Grade level = 3.0680 + .0877 (average sentence length) + .0984 (percentage of polysyllables)

The publication of the Flesch, Dale-Chall, and Gunning formulas conveniently marks the end of the first stretch of readability development. The authors of these formulas brought the issue of readability to public attention. They stimulated new consumer demands for documents in plain language. Finally, they stimulated new studies, not only on how to improve the formulas, but also on the other factors affecting reading success.

The New Readability Studies

The new readability was a period of consolidation and deeper study. Investigators sought to learn more about how the formulas work and how to improve them.

In the 1950s, several other developments accelerated the study of readability. The challenges of Sputnik and the demands of new technologies created a need for higher reading skills in all workers. While the older manufacturing industries had little demand for advanced readers, new technologies required workers with higher reading proficiency.

The New Readability studies were characterized by these features:

- **A community of scholars.** The periodical summaries of the progress of readability research (Klare 1952, 1963, 1974-75, 1984, Chall 1958, and Chall and Dale 1995) revealed a community of scholars. They were interested in how and why the formulas work, how to improve them, and what they tell us not only about reading, but also about writing.
- **The cloze test.** The introduction of the cloze test by Wilson Taylor in 1953 opened the way for investigators to test the properties of texts and readers with more accuracy and detail.
- **Reading ability, prior knowledge, interest, and motivation.** A number of studies looked at the manner in which these reader variables affect readability.
- **Reading efficiency.** While other studies looked at the effects of readability on comprehension, these studies looked at the effects on reading speed and persistence.
- **The measurement of content.** The influence of cognitive psychology and linguistics in the 1980s stimulated renewed studies of cognitive and structural factors in the text and how they can be used to predict readability.
- **Text leveling.** Cognitive and linguistic theory revived interest in the qualitative and subjective assessment of readability. With training, leveling can be effective in assessing the elements of texts not addressed by the formulas.
- **Producing and transforming text.** Several studies examined the effectiveness of using the formula variables to write and revise texts. When writers attend to content, organization, and coherence, using the

readability variables can be effective in producing and transforming a text to a required reading level.

- **New readability formulas**. Extensive studies of readability by John Bormuth and others looked at the reliability of a wide range of measurable text variables. They produced an empirical basis for criterion scores and criterion texts for the development of new formulas and reworking of old ones.
- **Formula discrepancy** A look a the discrepancy between the results of different formulas and how writers can benefit from it.

A community of scholars

Two notable features of readability research were a community of scholars and a long research base. The recognized bibliographer of that effort was George R. Klare, now Distinguished Professor Emeritus of Psychology and former Dean of the College of Arts and Sciences at the Ohio University. Formerly the Dean of the Department of Psychology, his field was psychological statistics and testing as well as readability measurement. He not only reviewed readability research (1963, 1974-75, 1984), but he also directed and participated in landmark studies and took the results of research to the public. His reviews established the validity of the formulas and their proper use not only in English, but also in many other languages. Among Klare's many important publications were:

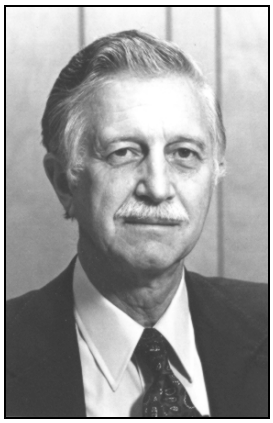


Fig. 7. George Klare. After serving as a navigator for the U.S. Air Force in WWII (in which he was shot down and captured by the Germans), Klare became a leading figure in readability research.

- *Know Your Reader: The Scientific Approach to Readability*, which he wrote with Byron Buck (1954).
- *The Measurement of Readability* (1963).
- "Assessing Readability*"* in the *Reading Research Quarterly* (1974-75). The Institute for Scientific Information recognized it as a Citation Classic, one of the scientific works most frequently cited in other studies—with well over 125 citations so far.
- "A Second Look at the Validity of the Readability Formulas" in *The Journal of Reading Behavior* (1976).
- "Readable Technical Writing: Some Observations" in *Technical Communication* (1977), which won "Best of Show" in the International Conference of the STC in Dallas in 1978.
- *A Manual for Readable Writing* (1975).
- *How to Write Readable English* (1980).
- *"*Readability" in *Encyclopedia of Educational Research* (1982).
- "Readability" in *The Handbook of Reading Research* (1984).
- "Readable Computer Documentation" in the *ACM Journal of Computer Documentation* (2000), which covered the latest research in readability.

Critics of the formulas (e.g., Redish and Selzer 1985) have complained that the readability formulas were developed for children and they never were never formulated or tested with technical documents. The record shows, however, that popular formulas such as the Flesch Reading Ease and the Kincaid formulas were developed mainly for adults and have been tested extensively on adult materials. For example, Klare (1952) tested the Lorge, Flesch Reading Ease, and Dale-Chall formulas against the 16 standardized passages of the Ojemann tests (1934) and the 48 passages of Gray and Leary (1935) tests, all developed for adult readers.

As we will see, several extensive studies (Klare et al. 1955a, Klare et al. 1957, Klare and Smart 1973, Caylor et al. 1973, Kincaid et al. 1975, Hooke et al. 1979) used materials developed for technical training and regulations in the military to formulate and test several of today's most popular formulas such as the Flesch-Kincaid grade-level formula.

The cloze test

Wilson Taylor (1953) of the University of Illinois published "Cloze Procedure: A New Tool for Measuring Readability." Taylor cited several difficulties with the classic readability formulas such as the Flesch and Dale-Chall. He noted, for instance, that Gertrude Stein's works measured much easier on the readability scales than expected.

Taylor argued that words are not the best measure of difficulty but how they relate to one another. He proposed using deletion tests called **cloze tests** for measuring an individual's understanding of a text. Cloze testing is based on the theory that readers are better able to fill in the missing words as their reading skills improve.

A cloze test uses a text with regularly deleted words (usually every fifth word) and requires the subjects to fill in the blanks. The percentage of words correctly entered is the **cloze score**. The lower the score, the more difficult the text. Because even advanced readers cannot correctly complete more than 65% of the deleted words correctly in a simple text, texts for assisted reading require a cloze score of 35% or more. Texts for unassisted reading need a higher score. Cloze scores line up with scores from multiple-choice tests in the following manner:

Purpose	Cloze	Multiple-Choice
Unassisted reading	50-60%	70-80%
Instructional, assisted reading	35-50%	50-60%
Frustration level	Below 35%	Below 50%

Table 7. Comparison of cloze and multiple-choice scores.

For the origins of these scores, see "The Problem of Optimal Difficulty" below.

A cloze test uses a text with selected words deleted and replaced with underlines of the same length. Having at least 50 blanks in the reading selection increases the reliability of the test.

To score a cloze test, use the percentage of all the words that are correctly entered, that is, the right words in the right form (no synonyms), number, person, tense, voice, and mode. Do not count spelling.

It greatly increases the accuracy of the test to test all the words by using different versions of the text. If you delete every 5th word, there are five possible versions, each one with a different first deleted word. Divide the subjects into as many groups as you have versions and give each group a different version.

Here is a sample cloze test:

> The potential for two-way _____ is very strong on _____ Web. As a result, _____ companies are focused on _____ Web's marketing potential. From _____ marketing point of view, _____ virtual worlds can attract _____curious Web explorers, and _____database engines can measure _____ track a visitor's every _____.

See the answers at the end of this article. Note that the standard cloze test does not provide the a list of the correct words to choose from as some online cloze programs do.

Cloze testing became the object of intensive research, with over a thousand studies published (Klare 1982). It quickly became popular as a research tool, and tended to complement not the formulas as expected but conventional reading tests. Unlike multiple-choice tests, cloze tests can provide suggestive information about individual sentences, clauses, phrases, and words. Cloze tests are suitable for intermediate and advanced readers. Cloze testing opened the way for much more intensive studies of the readability formulas, beginning with Bormuth in 1966 (see below).

Reading ability, prior knowledge, interest, and motivation

The interest factors affecting the readability of children's literature was taken up by Gates (1930) and Zeller (1941). One of the interest factors that Gates mentioned for children was reading ease. Flesch's early formula for adults (1949) included interest factors for measuring readability. The new research would establish that, along with vocabulary and sentence structure, the reader's reading ability, prior knowledge and motivation are powerful contributors to text readability.

Prior knowledge and retention A series of studies in the military (Klare et al. 1955a) examined how prior knowledge as well as the text variables affect the retention and the acceptability of technical documents.

The studies were conducted at Sampson Air Force Base in New York and Chanute Air Force Base in Illinois using 989 male Air Force enlistees in training with different versions of the same texts. They used the Flesch Reading Ease, Dale-Chall, and the Flesch Level-of-Abstraction formulas to rate the texts as Easy (grade 7), Present (12^{th} grade), and Hard (16^{th} grade).

While simplifying documents and changing the style, they retained all technical terms and used technical experts to assure that they did not change the content.

This study found the more readable versions resulted in:

- Greater and more complete retention.
- Greater amount read in a given time.
- Greater acceptability (preference).

The study found that, "…while style difficulty appears to affect immediate retention of subjects who are naïve regarding material, subjects who have considerable knowledge of the material may profit little if any from an easier style of material" (p. 294).

Duffy (1985) criticizes the results of this study. He states that the 8% percent improvement in comprehension, achieved by dropping the reading level of the texts eight grades (from the 16^{th}+ grade to the 7^{th}-8^{th} grades—1% improvement for each grade dropped) is not large enough to justify the effort required.

Duffy underestimates the difficulty of demonstrating the comprehension gained by changing any textual variable while carefully controlling the other variables done in the study. Most researchers are very happy to get any non-chance improvements in comprehension, the holy grail of reading research.

The difficulty arises from the complexity of reading comprehension and the means we have of testing it, which are all indirect. Researchers, for example, are not sure exactly what the results of reading tests are telling them. Do they reveal comprehension of the text or other artifacts such as prior knowledge, memory, or the difficulty of the questions?

Studies of the effects of textual variables and writing strategies on comprehension are very often inconsistent, inconclusive, or non-existent. Examples include: the use of illustrations (Halbert 1944, Vernon 1946, Omaggio

1979; Felker et al., 1981), schemas (Rumelhart 1984), structural cues (Spyridakis, 1989, 1989a), highlighting (Klare et al. 1955b, Felker et al.), paragraph length (Markel, et al., 1992), typographic format (Klare 1957), syntax simplification (Ulijn and Strother 1990), prior knowledge (Richards 1984), nominalizations, diagrams, parallelism, white space, line graphs, and justified margins (Felker et al,), "whiz deletions" (Huckin et al. 1991), writer guidelines (McLean 1985), and coherence and cohesion (Freebody and Anderson 1983, Halliday and Hasan 1976).

No one would say that any of these items are not helpful or do not affect comprehension. These studies show, however, how difficult it is to detect and measure the effect on comprehension of any reading variable. Even small gain in comprehension that is significant can be important over time and suggests further study. In this regard, the formulas do very well. See "Producing and transforming texts" below.

Career preferences, aptitudes, and test scores. A further investigation by the same authors (Klare et al. 1955c) looked into the effect of career aptitude and preferences on immediate retention. As expected, the subjects with higher degree of mechanical and clerical aptitude showed consistently higher retention on test scores. There were no significant relationships, however, between career preferences and retention.

Interest, Prior Knowledge, Readability and Comprehension A study (Klare 1976) of the experiments on the effects of using formulas to revise texts showed how different levels of motivation and reading ability can skew the results. It also indicated that the readability of a text is more important when interest is low than when it is high. The study by Fass and Schumacher (1978) supports this claim.

Woern (1977) later showed that prior knowledge and beliefs about the world affected comprehension significantly. Pearson, Hanson, and Gordon (1979) discovered significant effects of prior knowledge on the comprehension of children reading about spiders. Spilich, Vesonder, Chiesi, and Voss (1979) found that subjects having more knowledge about baseball remembered more information about a baseball episode. Chiesi, Spilich, and Voss (1979) found that high-knowledge subjects had better recognition, recall, and anticipation of goal outcomes than did low-knowledge subjects.

Entin and Klare (1985) took up the interaction between the readability of the text and the prior knowledge and interest of the readers. The study used 66 students enrolled in introductory psychology courses at Ohio University. They were first tested with the Nelson-Denny Reading Test to determine reading skills. They were then given a questionnaire on their interest in selected topics and a questionnaire on their prior knowledge of the terminology used in the test passages. For test passages, they used 12 selected passages from the World Book Encyclopedia, six high-interest passages, and six low-interest ones. The passages were re-written and normed by judges for content at the 12^{th} and 16^{th}-grade levels, resulting in 24 passages for the experiment. Then, two cloze tests were made of each passage, resulting in 48 test passages

This study confirmed that easier readability of a text has more benefits for those of less knowledge and interest than those of more. Advanced knowledge of a subject can "drown out" the effects of an otherwise difficult text.

This study also suggested that when reader interest is high, comprehension is **not** improved by writing the material below, rather than at, the grade level of the readers. When interest is low, however, comprehension **is** improved by writing the materials below, rather than at, the reading level of the readers.

Comprehension was improved when the materials are written *at* the reading levels of all readers rather than above those levels.

Reading efficiency

While early studies used reader comprehension as a measure of readability, new studies were looking at other measures such as:

- Readership
- Reading persistence (or perseverance)
- Reading efficiency

Readability and reader persistence Several studies in the field of journalism found a significant relationship between reader persistence and readability. Some used split runs of newspapers to see the effects of improved readability on wide audiences.

Donald Murphy (1947), the editor of *Wallace's Farmer,* used a split run with an article written at the 9th-grade level on one run and on at the 6th-grade level on the other run. He found that increasing readability increased readership up of the article 18 percent. In a second test, he took great care not to change anything except readability, keeping headlines, illustrations, subject matter and the position the same. He found readership increases of 45% for an article on nylon and 60% for an article on corn.

Wilbur Schramm (1947) showed that a readable style contributes to the readers' perseverance, also called depth or persistence, the tendency to keep reading the text.

Charles E. Swanson (1948) showed that better readability increases reading perseverance as much as 80 percent. He developed an easy version of a story with 131 syllables per 100 words and a hard version with 173 syllables and distributed each to 125 families. A survey of readers taken 30 hours after distribution showed a gain in the easier version over the hard version of 93% of total paragraphs read, 83% in mean number of paragraphs read, and 82% in the number of correspondents reading every paragraph.

Bernard Feld (1948) grouped 101 stories from the *Birmingham News* into those with high Flesch scores, requiring 9th-grade education or more and those with low scores, requiring less than 9th-grade education. He found readership differences of 20 to 75 percent favoring the low-score versions. Feld's findings indicated that even a small actual percentage gain for a large-circulation paper greatly increased the number of readers.

Reading efficiency Klare, Shuford, and Nichols (1957) followed up these studies with a study of the reading efficiency and retention of 120 male aviators in a mechanics course at Chanute Air Force Base in Illinois. They used two versions of technical training materials, hard (13th-15th grade) and easy (7th-8th grade).

They measured reading efficiency with an eye-movement camera with which they could determine the number of words read per second and the number of words read per fixation. A strong "set-to-learn" was stimulated by allowing the subjects to re-read the text and giving them a pre-test before the experimental test.

The study showed that the easy text significantly improved both reading efficiency and retention. The results also indicated that a strong "set to learn" improved scores.

Hardyck and Petrinovich (1970) showed the connection between readability and both comprehension and muscle activity in the oral area (subvocalization).

Rothkopf (1977) showed the connection between readability and how many words a typist continues to type after the copy page is covered (functional chaining).

Readability and course completion Publishers of correspondence courses are understandably concerned when large numbers of students do not complete the courses. They often suspect the materials are too difficult for the students. Working with Kim Smart of the U. S. Armed Forces Institute, Klare (1973) applied the Flesch Reading Ease formula to thirty sets of printed correspondence courses used by the military.

They found that two of the high school courses and five of the college courses were too difficult for readers of average or below average reading skill.

They then compared their reading analysis to the completion records of the 17 courses that had been in use over two years. They found a Spearman rank-order correlation of .87 between the readability score and the probability of students completing the course. There was a Pearson product-moment correlation of .76.

These results showed the importance of readability for unassisted reading where pressure to complete a course of study is low and competition from distractions is high.

The measurement of content

For hundreds of years, writers and teachers have used and taught the cognitive and structural factors in text such as organization and coherence. Researchers in readability also addressed the effects of these factors on comprehension:

- Image words, abstraction, predication, direct and indirect discourse, types of narration, and types of sentences, phrases, and clauses (Gray and Leary 1935).
- Difficult concepts (Morriss and Holverson 1938, Chall 1958).
- Idea density (Dolch 1939).
- Human interest (Flesch 1949, Gunning 1952)
- Organization (Gunning 1952, Klare and Buck 1954, Chall 1958).
- Nominalization (Coleman and Blumenfeld 1963; Coleman, 1964)
- Active and passive voice (Gough 1965, Coleman 1966, Clark and Haviland 1977, Hornby 1974).
- Embeddedness (Coleman 1966).

The cognitive theorists and linguists, beginning in the 1970s, promoted the idea that reading was largely an act of thinking. Among the ideas they promoted were:

1. Meaning is not in the words on the page. The reader **constructs meaning** by making inferences and interpretations.
2. Information is stored in long-term memory in organized "knowledge structures." The essence of learning is linking new information to **prior knowledge** about the topic, the text structure or genre, and strategies for learning.
3. A reader constructs meaning using **metacognition**, the ability to think about and control the learning process (i.e., to plan, monitor comprehension, and revise the use of strategies and comprehension); and **attribution**, beliefs about the relationship among performance, effort, and responsibility (Knuth and Jones 1991).

The cognitive theorists, aware of the limitations of the readability formulas, set about to supplement them with ways to measure the content, organization, and coherence of the text. Their studies reinforced the importance of these variables for comprehension. They did not, however, come up with any practical method for measuring or adjusting them for different levels of readers.

The following sections summarize a few of these efforts.

Walter Kintsch and coherence Beginning in 1977, Walter Kintsch and his associates studied the cognitive and structural issues of readability. Kintsch proposed to measure readability by measuring the number of propositions in a text. A proposition consists of a predicate and one or more arguments. An argument can be a concept or another argument. A concept is the abstract idea conveyed by a word or phrase.

In the early part of his work, Kintsch (Kintsch and Vipond 1979) was quite critical of the readability formulas. He said they are not based on modern linguistic theory and they overlook the interaction between the reader and the text.

Over four years, however, he and his associates revised this position. He eventually admitted that "these formulas are correlated with the conceptual properties of text" and that vocabulary and sentence length are the strongest predictors of difficulty (Kintsch and Miller 1981, p. 222).

While Kintsch and his colleagues did not come up with any easily used formula, they did contributed to our understanding of readability, including the central role of **coherence** in a text. Kintsch found out that lack of coherence affects lower-grade readers much more than upper-grade ones. The upper-grade readers, in fact, feel challenged to reorganize the text themselves. They may require more opportunities for solving problems, while lower-grade readers require more carefully organized texts.

The Air Force transformational formula. Perhaps the most ambitious attempt to quantify the variables of the cognitive theorists and put them in a formula was the project of Williams, Siegel, Burkett, and Groff (1977). Working for the Air Force Human Resources Laboratory, they examined new variables, produced a new formula, and presented supporting data. The variables they included were:

- Four psycholinguistic variables such as Yngve word depths, transformational complexity, center embedding, and right branching.
- Four Structure of Intellect variables including cognition of semantic units, memory for semantic units, evaluation of symbolic implications, and divergent production of semantic units.

For a criterion, they used cloze scores on 14 passages of about 600 words each taken from the Air Force career-development course. They deleted each tenth word in the cloze test and used only one version out of a possible ten on 51 Air-Force subjects. The computerized formula produced a correlation of 0.601 with text difficulty.

Susan Kemper and the reader's mental load Following Kintsch, Susan Kemper (1983) sought to explain comprehension in terms of underling cognitive processes. She developed a formula designed to measure the "inference load" based on three kinds of causal links:

- Physical states
- Stated mental states
- Inferred mental states

The Kemper formula measures the density of the propositions and embedded clauses. It takes considerable time and effort in comparison to the readability

formulas. It has a correlation of .63 with the McCall-Crabbs tests (the original Dale-Chall formula has a correlation of .64).

Kemper (p. 399) commented: "..sentence length and word familiarity do contribute to the comprehension of these passages…. These two different approaches to measuring the grade level difficulty of texts are equivalent in predictive power."

Kemper admitted that her formula, like all readability formulas, is better at predicting problems than fixing them. For writing, both formulas are best used as a general guide.

Bonnie Meyer and organization Bonnie Meyer and others worked on using the organization of larger units of texts as a possible measurement of readability. She claimed that a text that follows a topical plan is more efficient (saves effort) and more effective (gets more results). She wrote:

> That is, people remember more and read faster information which is logically organized with a topical plan than they do when the same information is presented in a disorganized, random fashion…. Thus the plan of discourse can be considered apart from content, and deserves separate consideration from researchers, as from those who are planning a composition (Meyer 1982, p. 38).

Among Meyer's observations are the following:

- A visible plan for presenting content plays a key role in assessing the difficulty of a text.
- A plan incorporates a **hierarchy** showing the dependencies of the facts to one another:
  - The **antecedent/consequent** plan shows causal relationships in "if/then" logic.
  - The **comparison** plan presents two opposing views that give weight to both sides.
  - The **adversative** plan clearly favors one side over the other (political speeches).
  - The **description** plan describes the component parts of an item (newspaper articles). This plan is the least effective for remembering and recall.
  - The **response** plan gives answers to remarks, questions, and problems (science articles).
  - The **time-order** plan relates events chronologically (history texts).

Better readers tend to share the same plan as authors of the material they are reading. Readers who use a different plan other than the authors may be at a disadvantage.

There are two types of **highlighting** for showing the relationships between items:

- **Subordination**, used to connect the main idea with supporting text as in a hierarchical structure.
- **Signaling**, explicit markers to clarify relationships such as:

  "On the one hand…On the other hand…"

  "Three things have to be stressed here."

  "Thus," "consequently," and "therefore"

  "Nevertheless," "all the same," "although," "but," and "however"

Signaling can also clarify how larger blocks of content are related, for example: "For example," "For further details," "summary," "abstract," "conclusion," and "preview." For more on signaling, see the studies by Jan Spyridakis (1989, 1989a).

Besides reducing the difficulty of the text, Meyer wrote that strategy training can also help older adults deal with the difficulties they encounter in reading.

Bonnie Armbruster and textual coherence Also concerned with larger units of text, Bonnie Armbruster (1984) found that the most important feature for learning and comprehension is **textual coherence**, which comes in two types:

- **Global coherence**, which integrates high-level ideas across an entire section, chapter, or book.
- **Local coherence**, which uses connectives to link ideas within and between sentences.

Armbruster found that recalling stories from memory is superior when the structure of the story is clear. She also noted the close relationship between global content and organization. Content is an aspect of structure, and organization is the supreme source of comprehension difficulty.

For local coherence, Armbruster stressed the highlighting that carries meanings from one phrase, clause, or sentence to another:

- Pronoun references to previous nouns
- Substitutions or replacements for a previously used phrase or clause (sometimes called "resumptive modifiers"), for example: "These results [previously listed] suggest that…"
- Conjunctions
- Connectives

Finally, Armbruster supported Kintsch's finding that coherence and structure are more important for younger readers than older ones, simply because they have less language and experience.

Calfee, Curley, and the familiar outline R.C. Calfee and R. Curley (1984) built on the work of Bonnie Meyer. They stressed making the structure of the text clear to upper-grade readers. The content can be simple, but an unfamiliar underlying structure can make the text unnecessarily difficult.

They proposed that the teacher, researcher, and student all need to reach a mutual understanding of the type of outline being used for the text under discussion.

Most students are familiar with the narrative structure, but not with other forms. Calfee and Curley present a graduated curriculum that enables students to progress from simpler structures to ones that are more difficult:

1. Narrative—fictional and factual
2. Concrete process—descriptive and prescriptive
3. Description—fictional, factual particular, and factual general
4. Concrete topical exposition
5. Line of reasoning—rational, narrative, physical and relational cause-and-effect
6. Argument—dialogue, theories and support, reflective essay
7. Abstract exposition

The lessons of content, organization, and coherence Organization and coherence highlight the relationships between words, sentences, paragraphs, and

larger sections of text. They enable readers to fit new items of information into their own cognitive systems of organization.

The cognitive studies of readability also showed other problems that texts can reveal or create, such as:

- Unfamiliar life experiences and background
- The need for time to digest illustrations and new material
- The need for multiple treatments of difficult material
- The need for learning aids to overcome textual difficulty
- The need for learning aids to help readers of different levels of skill.

Generally, however, the cognitive researchers failed to translate their theories into practical and objective methods for adjusting the difficulty of texts for different levels of reading skill.

Critics of the formulas (e.g., Manzo 1970, Bruce et al. 1981, Selzer 1981, Redish and Selzer 1985, Schriver 2000) rightly claim that the formulas use only "surface features" of text and ignore other features like content and organization. The research shows, however, that these surface features—the readability variables—with all their limitations have remained the best predictors of text difficulty as measured by comprehension tests (Hunt 1965, Bormuth 1966, Maxwell 1978, Coupland 1978, Kintsch and Miller 1981, Chall 1984, Klare 1984, Davison 1984 and 1986, Carver 1990, Chall and Conard 1991, Chall and Dale 1995).

Text leveling

An important byproduct of the cognitive and linguistic emphasis was the renewed interest in text leveling. This involves a subjective analysis of reading level that examines vocabulary, format, content, length, illustrations, repetition of words, and curriculum. Text leveling is perhaps the oldest method of grading a text. The McGuffey readers were graded by leveling, and their success is an indication of its validity.

Leveling recently became popular largely due to the work of the New Zealand Department of Education. In the U.S., Marie Clay's (1991) Reading Recovery system uses leveling in tutoring of children with reading problems. In this system, teachers use leveling to find books with closely spaced difficulty levels, particularly at the first-and second-grade levels. Most traditional readability formulas are not particularly sensitive at those levels (Fountas and Pinnell, 1999).

For that same reason, readability experts have long encouraged the use of subjective leveling along with the readability formulas. Leveling can spot the items that the formulas do not measure (Klare 1963, pp. 137-144; Chall et al. 1996; Fry 2002).

R. P. Carver (1975-1976) introduced a method of using **qualified raters** to assess the difficulty of texts. Raters become qualified when accurately judging the difficulty of five passages using his "Rauding Scale," consisting of six passages representing grades 2, 5, 8, 11, 14, and 17. Carver claimed his method was slightly more accurate than the Dale and Chall and Flesch Reading Ease formulas and provides grade-level scores through grade 18.

H. Singer (1975) created a method called SEER, "Singer Eyeball Estimate of Readability." It involves the use of one or two accurate SEER judges matching a sample of text against one of two scales, each consisting of eight rated passages. Singer claims his method is as accurate as the Fry graph.

The problem, of course, is that it takes some effort to learn how to do leveling accurately. Advanced readers often fail to recognize how difficult texts can be for others. Leveling also becomes more effective and accurate as the number of experienced judges increases (Klare, 1984).

Jeanne Chall and her associates (1996) published *Qualitative Assessment of Text Difficulty, A Practical Guide for Teachers and Writers*. It uses graded passages, called "scales," from published works along with layouts and illustrations for leveling of texts. You can assess the readability of your own documents by comparing them to these passages and using the worksheet in the book. The 52 passages are arranged by grade level and by the following types of text:

- Literature
- Popular fiction
- Life sciences
- Physical sciences
- Narrative social studies
- Expository social studies

The scale passages were selected on the basis of the following grade-related requirements for the reader:

1. Knowledge of vocabulary
2. Familiarity with sentence structure
3. Subject-related and cultural knowledge
4. Technical knowledge
5. Density of ideas
6. Level of reasoning

The selections were then tested by:

1. Evaluation by several groups of teachers and administrators
2. Evaluation by students of corresponding grades
3. Cloze testing of students of corresponding grades
4. Readability formulas (Dale-Chall and Spache)

The book also describes at length the various characteristics of each type of text that can contribute to difficulty. An added section features samples of the design and illustrations of books appropriate for the first four grades.

The following are three samples of the scales taken from the book.

Reading Level 3

> The stars, like the sun, are always in the sky, and they are always shining. In the daytime the sky is so bright that the stars do not show. But when the sky darkens, there they are.
>
> What are the stars, you wonder, and how do they twinkle?
>
> Stars are huge balls of hot, hot gas. They are like the sun but they look small because they are much, much farther away. They are trillions and trillions of miles away, shining in black space, high above the air.
>
> Space is empty and does not move. Stars do not twinkle there, but twinkling begins when starlight hits the air. The air moves and tosses the light around.
>
> —From *The Starry Sky: An Outdoor Science Book* (Wyler 1989, pp. 15-16)

Reading Level 5-6

> Black holes are probably the weirdest objects in space. They are created during a supernova explosion. If the collapsing core of the exploding star is large enough—more than four times the mass of our sun—it does not stop compressing when it gets as small as a neutron star. The matter crushes itself out of existence. All that remains is the gravity field—a black hole. The object is gone. Anything that comes close to it is swallowed up. Even a beam of light cannot escape.
>
> Like vacuum cleaners in space, black holes suck up everything around them. But their reach is short. A black hole would have to be closer than one light-year to have even a small effect on the orbits of the planets in our solar system. A catastrophe such as the swallowing of the Earth or the sun is strictly science fiction.
>
> —From *Exploring the Sky* (Dickinson 1987, p. 42)

Reading Level 7-8

> As we have seen, a neutron star would be small and dense. It should also be rotating rapidly. All stars rotate, but most of them do so leisurely. For example, our Sun takes nearly one month to rotate around its axis. A collapsing star speeds up as its size shrinks, just as an ice-skater during a pirouette speeds up when she pulls in her arms. This phenomenon is a direct consequence of a law of physics known as the conservation of angular momentum, which holds that the total amount of angular momentum in a system holds constant. An ordinary star rotating once a month would be spinning faster than once a second if compressed to the size of a neutron star.
>
> In addition to having rapid rotation, we expect a neutron star to have an intense magnetic field. It is probably safe to say that every star has a magnetic field of some strength.
>
> —From *Discovering the Universe (Faufmann 1990, p. 290)*

Producing and transforming text

While the formulas were originally created to help educators select texts for different audiences, writers also use the formula variables to produce texts and transform (re-write) them into simpler versions. The evidence on how effective this is has been mixed. As both the supporters of the formulas and their critics have warned, if you just chop up sentences and use shorter words, the results are not likely to improve comprehension. You have to look at the many other factors that affect reading at the level for which you are writing.

Early evidence on the effects of using the formula variables to transform text was negative. Klare (1963) reported that, of the six readability studies involving the controlled manipulation of words or sentences, only one had a positive effect, and this involved simplifying vocabulary.

In a later study, Klare (1976) took a careful look at 36 studies that examined the effects on comprehension of using the readability formula variables in re-writing texts. He grouped them by their results:

- 19 studies had positive results (readability variables had a significant effect on comprehension and/or retention)
- 6 studies had mixed results
- 11 studies had negative results (no measurable effect).

In seeking the reasons for the differences, Klare looked carefully at 28 situational factors in which each experiment was conducted. The situational factors fell into these groups:

- The readability and content of the material.
- The competence and motivation of the subjects.

- The instructions given the subjects during the experiment.
- The details of the test situation.

Klare found that differences in readability were often overridden by other factors in the test situation such as:

- The instruction given to the subjects of the test.
- The presence of threats or rewards.
- The time allowed for reading and testing.
- The presence or absence of the text during the test.

Klare wrote that the performance of the subject in such tests is a function not only of the difficulty of the material, but also in critical degrees, a function of the test situation (time, place, etc.), the content of the material and the competence and motivation of the reader. Scores will be better, for instance, when the readers love the subject matter or if they are highly motivated (e.g., paid).

Klare concluded that in the studies that showed increased comprehension, transforming text requires attending to other problems besides word and sentence length. "The best assumption, it seems to me," he wrote, "is that the research workers, probably with considerable effort, managed to change basic underlying causes of difficulty in producing readable versions" (p. 148). Klare then listed the following word-and-sentence variables that affected comprehension:

Word characteristics:

1. Proportion of content (functional) words.
2. Frequency, familiarity, and length of content words.
3. Concreteness or abstractness.
4. Association value.
5. Active vs. nominalized verb constructions.

Sentence characteristics:

1. Length (esp. clause length).
2. Active vs. passive.
3. Affirmative vs. negative.
4. Embedded vs. non-embedded.
5. Low depth vs. high depth (branches).

Since Klare's 1976 study, there have been other studies showing the positive effects of using formula variables to improve comprehension (Ewing 1976, Green 1979, C. C. Swanson 1979).

In the many studies of before-and-after revision of the text, a negative result does not prove that there is no improvement in comprehension. They show instead that improvement has not been detected. There is a saying in statistics that you cannot prove a negative.

Studies reporting a negative result may result from failing to control the reading ability, prior knowledge, interest, and motivation of the subjects. They can also result from failing to control elements of the text such as organization, coherence, and design. The great difficulty of properly conducting such an experiment is seen in following two studies.

The Duffy and Kabance study Critics worry that technical communicators can too easily misuse the formulas, making documents more difficult, not less

(Charrow 1977, Kern, 1979, Selzer 1981, Lange 1982, Duffy 1985, Redish and Selzer 1985, Connaster 1999, Redish 2000, Schriver 2000). These writers offer little or no evidence of such misuse, however, widespread or otherwise. If unscrupulous or careless writers choose to cheat by "writing to the formula" and not attending to other textual issues, careful editors and reviewers easily spot the misuse. The study by Thomas Duffy and Paula Kabance (1981) is a case in point. Because formula critics (e.g., Redish and Selzer 1985; Redish 2000) often refer to this study, it deserves some attention.

The Duffy and Kabance study consisted of four experiments that examined the effects of changing only word and sentence length on comprehension. It used a "reading to do" task and a "reading to learn" task. The study used four versions of the text:

1. The original version (a narrative or expository passage from the 1973 Nelson-Denny reading tests).
2. One with vocabulary that they simplified using *The Living Word Vocabulary*
3. One with only simplified and shortened sentences.
4. One with both vocabulary and sentences simplified.

The effect was a 6-grade drop in reading level of the changed passages from the 11th to the 5th grade.

Following Klare's research protocols (1976), they attempted to maximize the readability effects by using readers who were low motivated, unfamiliar with the topic, and widely varying in reading skills.

Using the Nelson-Denny reading tests, they tested the reading ability of the 1,169 subjects, male Navy trainees between 17 and 20 years old, of which 80% were high-school graduates. They divided them into two groups, one with a median reading grade of 8.7 and the other 10.3. The experiments took place in groups of 40 to 70.

In the first two experiments, they simulated a "reading-to-do" situation. In the first experiment, they first showed the questions, then had the subjects read the text. After that, they were shown the questions again, which they answered. In the second experiment, they limited the reading time but let the subjects have access to the text while answering the questions. The third experiment was a standard cloze test. The fourth experiment was a standard multiple-choice test with the subjects first reading the text and then answering the questions without the text.

The first three experiments showed no significant improvements. The fourth experiment resulted in significant improvement but only with the low-ability group using the changed-vocabulary text, an improvement of 13 percent. The authors concluded that simplifying the text made no difference to the advanced readers. This is not a surprising result, when we consider the reading ability of the advance group was at grade 10.3 while the difficult text was at 11th grade.

The vocabulary variable is significant for the low-ability group, they stated, but only in reading-to-learn tasks but not reading-to-do tasks, where memory is less important. This correlation was also suggested by Fass and Schumacher (1978).

Duffy and Kabance concluded that the increased readability is not required for technical documents, in which the emphasis is on "reading to do" and memory is not required.

This is sometimes true. At other times, serious errors have taken place because of memory failure. Many, if not most, technical tasks involve learning a skill that can be repeated, as Redish (1988) emphasizes. Besides reading-to-learn and

reading-to-do tasks, she writes, many technical tasks require "reading to learn to do." Technical texts may require more memory than do most other kinds of literature such as magazines, newspapers, or fiction.

When we look at the methods of these experiments, difficulties appear that explain their inconsistent results. In their report, Duffy and Kabance provide four sample passages used in the study. The re-written passages appear disjointed and stilted, not what one would expect of a 5th-grade text (See Fig. 11). If these studies are representative of the other passages, we must assume that judges were not used to control for the coherence and content of the text.

This was the also the conclusion of Leslie Olsen and Rod Johnson (1989), who wrote: "In their study, Duffy and Kabance were trying to directly manipulate the understanding of the words and the syntax of the sentences. However, it seemed to us that they were also unintentionally altering other aspects of the text—in particular, the cohesive structures of the text."

In their paper, Olsen and Johnson defined "sensed cohesion" as the strength of the textual topicality and the sense of givenness. The strength of textual topicality is related to the persistence of what the text is about. The sense of givenness is the recognition that the reader has seen a particular noun phrase before.

In analyzing the passages of the Duffy and Kabance study, Olsen and Johnson found that long sentences were broken up into short sentences. In the process, they introduced new subjects. The original focus on the Spaniards was lost, making it difficult to know what the text is about. They analyzed the cohesiveness of the text and concluded, "the intended and the unintended effects of the revisions cancelled one another out," bringing the results of the study into question.

Original (11th Grade)	**Sentences and Vocabulary Revised (5th Grade)**
The night was cloudy, and a drizzling rain, which fell without intermission, added to the obscurity. Steadily, and as noiselessly as possible, the Spaniards made their way along the main street, which had so lately resounded to the tumult of battle. All was now hushed in silence; they were only reminded of the past by the occasional presence of some solitary corpse, or a dark heap of the slain, which too plainly told where the strife had been the hottest. As they passed along the lanes and alleys which opened into the great street, they easily fancied they discerned the shadowy forms of their foe lurking in ambush, ready to spring upon them. But it was only fancy; they city slept undisturbed even by the prolonged echoes of the tramp of the horses, and the hoarse rumbling of the artillery and baggage trains. At length, a lighter space beyond the dusky line of buildings showed the van of the army that it was emerging on an open causeway. They might well have congratulated themselves on having thus escaped the dangers of an assault in the city itself, and that a brief time would place them in comparative safety on the opposite shore.	The night was cloudy. A sprinkling rain added to the darkness. It fell without a break. The Spaniards made their way along the main street. They moved without stopping and with as little noise as possible. The street had so recently roared to the noise of battle. All was now hushed in silence. The presence of a single dead body reminded them of the past. A dark heap of the slain also reminded them. Clearly, the battle had been worse there. They passed along the lanes and alleys opening into the great street. They easily fancied the shadows of their enemy lying in wait. The enemy looked ready to spring upon them. But it was only fancy. The city slept without being bothered by the rough rumbling of the cannons and baggage trains. Even the constant sound of the tramp of horses did not bother the city. At length, there was a bright space beyond the dark line of the buildings. This informed the army look-out of their coming out onto the open highway. They might well have rejoiced. They had thus escaped the dangers of an attack in the city itself. A brief time would place them in greater safety on the opposite shore.

Fig. 8. Original and revised samples of the passages used in the Duffy and Kabance study of 1981. Lack of attention to coherence and other important variables can cancel out the effects of rewriting the text using the readability-formula variables.

The Charrow and Charrow study Critics of the formulas (e.g., Bruce et al. 1981, Redish and Selzer; Redish 2000) also refer to the elaborate study of oral jury instructions by attorney Robert Charrow and linguist Veda Charrow (1979). They claimed that simplifying text did not make verbal instructions more comprehensible.

The authors did not use the readability variables in re-writing jury instructions but simplified the instructions using a list of common legal "linguistic constructions." These were: nominalizations, unusual prepositional phrases, misplaced phrases, whiz deletions (use of participles instead of verbs), deletions of "that" or "which" beginning dependent clauses, technical legal terms, imperative terms, negatives, passive voice, word lists, organization, and dependent clauses.

The first experiment used 35 persons called for jury duty in Maryland using 14 jury instructions taken from California's standard civil jury instructions. The purpose of the study was mainly to see if it was the complexity of the legal issues that made the instructions difficult or the difficulty of the language used. A group of attorneys were asked to rate the instructions according to their perceived complexity.

The experimenters tested each person individually by playing each instruction twice on a tape recorder. After hearing each instruction, the subject then verbally paraphrased the instruction, which was also recorded. The results showed, contrary to the attorneys' expectations, it was not the complexity of the ideas that caused problems in comprehension, but the difficulty of the language.

The second experiment used 48 persons chosen for jury duty in Maryland. For this experiment, they re-wrote the instructions, paying close attention to the legal constructions noted above. They divided the group into two. Using 28 original and modified instructions, they gave seven original instructions and seven modified instructions to each group. They used the same protocols in playing the instructions twice and asking the subjects to paraphrase them.

There was a significant improvement of the mean scores in comprehension in nine of the fourteen instructions. They concluded that the subjects understood the gist of the original only 45% of the time and the simpler ones 59% of the time.

This is not good enough, according to Professor Robert Benson (1984-85) of Loyola Law School in Los Angeles. He wrote, "…none of us would care to be tried by jurors who understood only 59% of their instructions."

Benson went on to say that the Charrows own data was leading them to a conclusion that they were unable to draw: that juries are never likely to understand oral instructions adequately. Elwork, Sales, and Alfini (1982) reach the same conclusion and recommend giving all jurors written as well as oral instructions.

To prove his point, Benson included three of the Charrows' re-written instructions in his own study on legal language using 90 law students and 100 non-lawyers. Using cloze tests, he found that, while the Charrows had reported 59% comprehension, the readers understood the written instructions almost fully (p. 546).

As to the claim that paraphrasing is better than other testing techniques, Benson claims that it has its own limitations, depending as it does on the subjects' ability to orally articulate what they understand. The Charrows had avoided asking the subjects to paraphrase in writing because "subject's writing skills would confound the results." Unfortunately, they ignored similar possible differences in their listening and their oral skills (Benson, p. 537).

The Charrows state that sentence length does not cause reading difficulty. "Although readability formulas are easy to use," they write, "and certainly do indicate the presence of lengthy sentences, they cannot be considered measures of comprehensibility. Linguistic research has shown that sentences of the same length may vary greatly in actual comprehensibility" (p. 1319).

Benson answered by writing that extremely long sentences such as those found in legal language are known to cause difficulty, probably because of memory constraints. He also found that the Charrows' revised instructions had actually shortened sentences by 35 percent. The shorter sentences "may well have played a role in improved comprehension" (pp. 552-553).

A number of studies show that, in the average, as a sentence increases in length it increases in difficulty (e.g., Coleman, 1962, Bormuth 1966). Average sentence length has long been one of the clearest predictors of text difficulty.

New readability formulas

Critics of the formulas and formula developers questioned the reliability of the criterion passages, criterion scores, and the reading tests on which the formulas had been developed and validated. The arrival of cloze testing stimulated the development of new criterion passages, new formulas, manual aids, computerized versions, and the continued testing of text variables.

The Coleman formulas Edmund B. Coleman (1965), in a research project sponsored by the National Science Foundation, published four readability formulas for general use. The formulas are notable for their predicting mean close scores (percentage of correct cloze completions).

Coleman was also the first to use cloze procedures as a criterion rather than the conventional multiple-choice reading tests or rankings by judges.

The four formulas use different variables as shown here:

C% = 1.29w – 38.45

C% = 1.16w + 1.48s – 37.95

C% = 1.07s + 1.18s + .76p – 34.02

C% = 1.04w + 1.06s + .56p – .36prep – 26.01

Where: C% = percentage of correct cloze completions;

w = number of one-syllable words per 100 words

s = number of sentences per 100 words

p = number of pronouns per 100 words

prep = number of prepositions per 100 words

Coleman found multiple correlations of .86, .89, .90, and .91, respectively, for his formulas with cloze criterion scores. The use of cloze scores as criterion consistently provides higher validation coefficients than does use of the multiple-choice scores. This may be a partial reason for the high correlations shown here.

The Bormuth studies Recognizing the problems of having more reliable criterion passages, John Bormuth conducted several extensive studies, which gave a new empirical foundation for the formulas. His first study (1966) provided evidence of just how much changes in a number of readability variables beside just vocabulary and sentence length can affect comprehension. Cloze testing made it possible to measure the effects of those variables not just on the difficulty of whole passages but also on individual words, phrases, and clauses.

His subjects included the entire enrollment of students (675) in grades 4 through 8 of Wasco Union Elementary School district in California. Their reading levels went from the 2nd through the 12th grade. He used 20 passages of 275 to 300 words each, rated on the Dale-Chall formula from the 4th to the 8th-grade levels of difficulty. He used five cloze tests for each passage, with the fifth-word deletions starting at different words.

Reading researchers recognized that beginning readers relate differently to word variables than do better readers. For this reason, special formulas have been developed for the earliest primary grades such as the Spache formula (1953) and the Harris-Jacobson primary readability formula (1973).

Bormuth's study confirmed the **curvilinearity** of the formula variables. That means their correlation with text difficulty changes in the upper grades, producing a curve when plotted on a chart. Dale and Chall (1948) included an adjustment for this feature in their formula-correction chart. This adjustment was also included in the SMOG formula (McLaughlin 1968), the Fry Graph (Fry 1969), the FORCAST formula (Caylor et al. 1973), Degrees of Reading Progress (Koslin et al. 1987), and the ATOS formula (Paul 2003).

Some critics of the formulas (Rothkopf 1972, Thorndike 1973-74, Selzer 1981, Redish and Selzer 1985) claim that decoding words and sentences is not a problem for adults. Bormuth's study, however, shows that the correlation between the formula variables and comprehension do not change as a function of reading ability (p. 105). Empirical studies have confirmed that, in adult readers, difficulty in reading is linked to word recognition (Stanovich 1984) and decoding of sentences (Massad 1977). We cannot assume that adults are better learners than children of the same reading level. In fact, they are often worse (Russell 1973, Sticht 1982).

Bormuth's next project (1969) was a study of the readability variables and their relationship to comprehension. His subjects included 2,600 fourth-to-twelfth-grade pupils in a Minneapolis school district.

The method consisted first in rating the reading ability of all the students with the California 1963 Reading Achievement test. It used 330 different passages of about 100 words each to confirm the reliability of 164 different variables, many of them never examined before such as the parts of speech, active and passive voice, verb complements, and compound nouns.

The five cloze tests used for each passage (resulting in 1,650 tests) gave him about 276 responses for each deleted word, resulting in over 2 million responses to analyze.

With this data, Bormuth was able to develop 24 new readability formulas, some of which used 14 to 20 variables. These new variables, he found, added little to the validity of the two classic-formula variables and were eventually dropped. The study divided the students of each reading level into two groups, one that was given a multiple-choice test and the other a cloze test of the same material.

Since Thorndike's (1916) recommendation, educators and textbook publishers had used 75% correct scores on a multiple-choice test as the criterion for optimum difficulty for assisted classroom learning, and 90% for independent reading. These criterion scores, however, were based on convention and use, not on scientific study.

This Bormuth study validated the equivalencies of 35%, 45%, and 55% correct cloze criterion scores with 50%, 75%, and 90% correct multiple-choice scores. It also showed that the cloze score of 35% correct answers indicates the level of difficulty required for maximum information gain.

Finally, this study produced three different formulas, one is for basic use, one for machine use, and one for manual use. All three formulas predict the difficulty of texts for all grade levels using a 35%, 45%, 55%, or a mean-cloze criterion.

The Bormuth Mean Cloze formula This formula uses three variables: number of words on the original Dale-Chall list of 3,000, average sentence length in words, and average word length in letters. This formula was later adapted and used in the Degrees of Reading Power used by the College Entrance Examination Board in 1981 (see below). The original Bormuth Mean Cloze formula is:

$$R = .886593 - .083640\ (LET/W) + .161911\ (DLL/W)^3 - 0.021401\ (W/SEN) + .000577\ (W/SEN)^2 - .000005\ (W/SEN)^3$$

$$DRP = (1 - R\ ) \times 100$$

Where: R = mean cloze score
LET = letters in passage X
W = words in passage X
DLL = Number of words in the original Dale-Chall list in passage X
SEN = Sentences in passage X
DRP = Degrees of Reading Power, on a 0-100 scale with 30 (very easy) to 100 (very hard)

The findings of Bormuth about the reliability of the classic variables were confirmed by MacGinitie and Tretiak (1971) who said that the newer syntactic variables proposed by the cognitive theorists correlated so highly with sentence length that they added little accuracy to the measurement. They concluded that average sentence length is the best predictor of syntactic difficulty.

The Bormuth studies provided formula developers with a host of new criterion passages. Critics of the formulas claimed that the criterion passages used by formula developers were arbitrary or out-of-date (Bruce et al. 1981, Duffy, 1985). As new criterion passages became available, developers used them to create new formulas and to correct and reformulate the older ones (Bormuth 1966, 1969, Klare 1985). The new Dale-Chall formula (1995) was validated against a variety of criterion passages, including 32 developed by Bormuth (1971), 36 by Miller and Coleman (1967), 12 by Caylor et al. (1973) and 80 by MacGinitie and Tretiak (1971). Other formulas were validated against normed passages from military technical manuals (Caylor et al. 1973, Kincaid et al. 1975).

The Fry Readability Graph While Edward Fry (1963, 1968) was working as a Fullbright scholar in Uganda trying to help teachers teach English as a second language, he created one of the most popular readability tests that use a graph.

Fig. 9. Edward Fry's Readability Graph may be the most popular readability aid.

Fry would go on to become the director of the Reading Center of Rutgers University and an authority on how people learn to read.

Fry's original graph determines readability through high school. It was validated with comprehension scores of primary and secondary school materials and by correlations with other formulas.

Fry (1969) later extended the graph to primary levels. In 1977, he extended it through the college years (Fig. 11). Although vocabulary continues to increase during college years, reading ability varies much, depending on both individuals and the subjects taught. That means that a text with a score of 16 will be more difficult than one with a score of 14. It does not mean, however, that one is appropriate for all seniors and the other for all sophomores.

Directions:

1. Select samples of 100 words.
2. Find y (vertical), the average number of sentences per 100-word passage (calculating to the nearest tenth).
3. Find x (horizontal), the average number of syllables per 100-word sample.
4. The zone where the two coordinates meet shows the grade score.

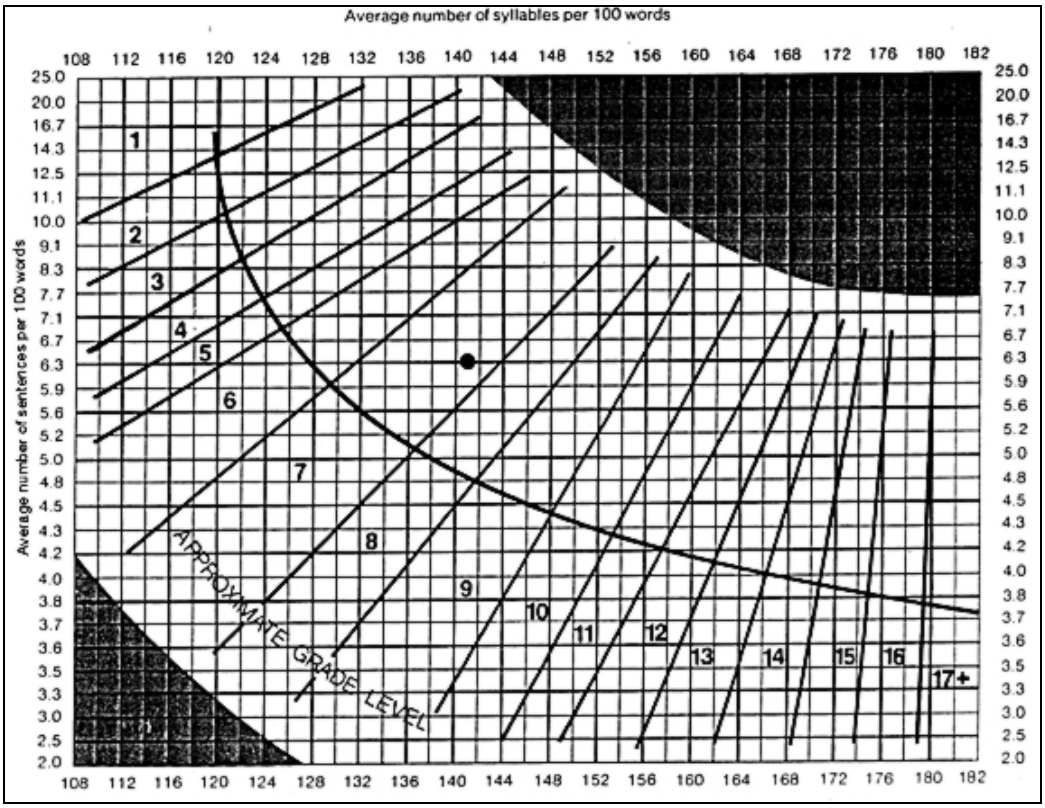


Fig. 10. The Fry Readability Graph as amended in 1977 with the extension into the primary and college grades. Scores that appear in the dark areas are invalid.

The Listening Formulas People have been concerned about the clarity of spoken language perhaps for a longer period than written language. Speech is generally much simpler than text. Because a listener cannot re-read a spoken sentence, it puts a greater demand on memory. For this reason, "writing like you talk" and reading text aloud have long been methods for improving readability. Studies of the correlations of listenability and readability have had mixed results (Klare 1963).

Some formulas have been developed just for spoken text. Rogers (1962) published a formula for predicting the difficult of spoken text. He used 480 samples of speech taken from the unrehearsed and typical conversations of students in elementary, middle, and high school as his data for developing his formula. The resulting formula is:

G = .669 I + .4981 LD – 2.0625

Where:

G = reading grade level

I = average idea unit length

LD = the average number of words in a hundred-word sampling that do not appear on Dale's long list (3,000 words).

Roger's formula has a multiple correlation of .727 with the grade level of his samples.

Irving Fang (1966-1967) used newscasts to develop his Easy Listening Formula (ELF), shown here:

ELF = number of syllables above one per word in a sentence.

An average sentence should have an ELF score below 12 for easy listenability. Fang found a correlation of .96 between his formula and Flesch's Reading Ease formula on 36 television scripts and 36 newspaper samples.

Subsequent research into listenability indicates that after the 8th grade, listening skills do not keep up with the improvement in reading skills. After the 12th-grade level, the same text may be harder to understand when heard than when read (Chall 1983b; Dale and Chall 1995; Sticht, Beck, et al. 1974).

The SMOG formula G. Harry McLaughlin (1969) published his SMOG formula in the belief that the word length and sentence length should be multiplied rather than added. By counting the number of words of more than two syllables (polysyllable count) in 30 sentences, he provides this simple formula:

SMOG grading = 3 + square root of polysyllable count.

McLaughlin validated his formula against the McCall-Crabbs passages. He used a 100 percent correct-score criterion. As a result, his formula generally predicts scores at least two grades higher than the Dale-Chall formula.

The FORCAST formula The Human Resources Research Organization studied the reading requirements of military occupational specialties in the U.S. Army (Caylor et al. 1973). In order to resolve professional questions about using a formula for technical material read by adults, the authors first undertook the creating of a readability formula that would be:

1. Based on essential Army-job reading material.
2. Adjusted for the young adult-male Army-recruit population.
3. Simple and easy for standard clerical personnel to apply without special training or equipment.

The researchers first selected seven high-density jobs and 12 passages that recruits are required to understand to qualify for them. They graded the passages with the modified Flesch formula, finding them to range from the 6th to the 13th grade in difficulty. They also selected 15 text variables to study for a new formula. They next tested the reading ability of 395 Army recruits, and then divided them into two groups, one with a mean-grade reading level of 9.40 and another 9.42.

They next tested the recruits with cloze tests made of the 12 passages. The 12 passages were then re-graded using the criterion of at least 50% of those subjects of a certain grade level being obtaining a cloze score of at least 35%. Results indicated that average subjects scored 35.1% on the text graded 9.1 and 33.5% on the text graded 9.6.

They next intercorrelated the results of the reading tests with the results of the graded cloze tests. Results showed usable correlations of .83 and .75 for the two groups of readers. Among the 15 variables they examined, the number of one-syllable words in the passage correlated highest (.86) and was selected for use in their new formula. Because they found that adding a sentence factor did not improve the reliability of the formula, they left it out. The resulting FORCAST formula is:

$$\text{Grade level} = 20 - (N \div 10)$$

Where N = number of single-syllable words in a 150-word sample.

The new formula correlated r = 9.2 with the Flesch Reading Ease formula, 9.4 with the original Dale-Chall formula with, and r = .87 with the graded text passages with. It is accurate from the 5th to the 12th grade.

They cross-validated the formula with a second study using another sample of 365 Army recruits at Ford Ord using another sample of reading passages scaled from grade 7 to grade 12.7 using the FORCAST formula. The results of this experiment correlated r = .98 with the Flesch formula, .98 with Dale-Chall, and .77 with the graded military passages. These figures were judged appropriate for the purpose of the formula.

Using the FORCAST formula, they tested the critical job-reading materials for readability. The results show the percentage of materials in each occupation written at the 9.9 grade level: Medical specialist, 24.4%; Light Weapons Infantryman, 18.3%; Military Policeman, 15.1%; General Vehicle Repairman, 13.4%; Amorer/Unit Supply Specialist, 10.8%; Ground Control Radar Repairman, 4.2%, and Personnel Specialist, 2.2%.

The study showed that materials for the different occupations all had texts above the 9th grade. This suggested the need for new quality-control measures for making materials more useful for the majority of personnel.

Fig. 13. Thomas Sticht. After participating in the military studies which resulted in the FORCAST readability formula, he went on to become a worldwide authority in adult literacy. He is shown here with UNESCO's Mahatma Gandhi Medal he received in 2003 for his contributions to that field. The citation reads:

> The UNESCO Mahatma Gandhi Bronze Medal has been awarded to Dr. Thomas Sticht (USA) in deep appreciation of his 25years of service as a member of UNESCO's International Jury for Literacy Prizes and in recognition of his devotion to the cause ofadult literacy, especially for his efforts to 'reach the unreached.

In a follow-up study, Lydia Hooke and colleagues (1979) validated of the use of the FORCAST formula on technical regulations for the Air Force. They also found that four of seven writers of the regulations underestimated the grade level of their materials by more than one grade.

In the main portion of the Hooke study, they administered cloze and reading tests to 900 AF personnel to determine the comprehension of each regulation by the user audience. Where there was no literacy gap (difficulty too high for the reader), they found that comprehension was adequate (at least 40% cloze score) in all cases. Where a literacy gap did exist, comprehension scores were below the criterion of 40% in three of four cases.

The FORCAST formula is very unusual in that it does not use a sentence-length measurement. This makes it a favorite, however, for use with short statements and the text in Web sites, applications, and forms. The Department of the Air Force (1977) authorized the use of this formula in an instruction for writing understandable publications.

The following are two of the scaled passages taken from training materials and used in the occupational specialty study for the development and validation of the FORCAST formula. Also shown are: 1. The scaled Reading Grade Level

(RGL), the mean reading grade level of the subjects who scored 35% correct scores on the cloze tests; and 2. The scores of the FORCAST, the Flesch, and the original Dale-Chall readability grade levels.

Passage 21

If you do not have a compass, you can find direction by other methods.

The North Star. North of the equator, the North Star shows you true north. To find the North Star—

Look for the Big Dipper. The two stars at the end of the bowl are called the "pointers." In a straight line out from the pointers is the North Star (at about five times the distance between the pointers). The Big Dipper rotates slowly around the North Star and does not always appear in the same position.

You can also use the constellation Cassiopeia. This group of five bright stars is shaped like a lopsided M (or W, when it is low in the sky). The North Star is straight out from the center star about the same distance as from the Big Dipper. Cassiopeia also rotates slowly around the North Star and is always almost opposite the Big Dipper.

Scaled RGL = 6. FORCAST = 8.6. Flesch = 7. Dale-Chall =7-8.

Passage 15

Adequate protection from the elements and environmental conditions must be provided by means of proper storage facilities, preservation, packaging, packing or a combination of any or all of these measures. To adequately protect most items from the damaging effects of water or water-vapors, adequate preservation must be provided. This is often true even though the item is to be stored in a warehouse provided with mechanical means of controlling the temperature and humidity. Several methods by which humidity is controlled are in use by the military services. Use is also made of mechanically ventilating and dehumidifying selected sections of existing warehouses. Appropriate consideration will be given to the preparation and care of items stored under specific types of storage such as controlled humidity, refrigerated, and heated. The amount and levels of preservation, packaging, and packing will be governed by the specific method of storage plus the anticipated length of storage.

Scaled RGL = 11.4. FORCAST = 12.1. Flesch = 13-16. Dale-Chall = 13-15.

The Army's Automated Readability Index (ARI) For the U.S. Army, Smith and Senter (1967) created the Automated Readability Index, which used an electric typewriter modified with three micro switches attached to cumulative counters for words and sentences.

The ARI formula produces reading grade levels (GL):

GL = 0.50 (words per **sentence**) + 4.71 (strokes per word) – 21.43.

Smith and Kincaid (1970) successfully validated the ARI on technical materials in both manual and computer modes.

The Navy Readability Indexes (NRI) Kincaid, Fishburne, Rogers, and Chissom (1975, Fishburne 1976) followed a trend by recalculating new versions of older formulas and testing them for use on Navy materials. The first part of the experiment aimed at the recalculation of readability formulas. The second part of the study aimed at validating the effectiveness of the recalculated formulas on Navy materials as measured by:

- Comprehension scores on Navy training manuals
- Learning time, considered being an important measurement of readability.

The first part of the study first determined the reading levels of 531 Navy personnel using the comprehension section of the Gates-MacGinitie reading test. At the same time, they tested their comprehension of 18 passages taken from Navy training manuals. The results of those tests were used in calculating the grade levels of the passages. They then used those passages to recalculate the ARI, Flesch, and Fog Count formulas for Navy use, now called the Navy Readability Indexes (NRIs). The recalculated grade-level (GL) formulas are:

ARI simplified:

GL = .4 (words per sentence) + 6 (strokes per word) – 27.4

Fog Count new:

$$GL = \frac{((\text{easy words} + 3\ (\text{hard words})) \ / \ (\text{sentences})) - 3}{2}$$

Where:

easy words = number of number of 1 and 2-syllable words per 100 words

hard words = number of words of more than 2 syllables per 100 words

sentences = number of sentences per 100 words

Flesch Reading Ease formula simplified and converted to grade level (now known as the Flesch-Kincaid readability formula):

New:

GL = (.39 x ASL) + (11.8 x ASW) – 15.59

Simplified:

GL = (.4 ASL) + (12 ASW) – 15

Where:

ASL = average sentence length (the number of words divided by the number of sentences).

ASW = average number of syllables per word (the total number syllables in the sample divided by the number of words).

The second part of the study looked at the relationship between readability and learning time. It monitored the progress of 200 Navy technical-training students through four modules of their course for both comprehension and learning time. The study was replicated with a secondary sample of 100 subjects performing on four additional modules.

The results of the comprehension test showed the highest percentage of errors in both the readers with the lowest reading grade levels and in the modules with the highest grade-levels of readability.

In the same manner, the learning time systematically decreased with reading ability and increased with the difficulty of the modules. The study confirms that learning time as well as reading ability are significant performance measures for predicting readability.

The new Flesch-Kincaid formula was able to predict significant differences between modules less than one grade level apart using both comprehension scores and learning times. The U.S. Department of Defense (1978) authorized this formula in new procedures for validating the readability of technical manuals for the Armed Services. The Internal Revenue Service, and the Social Services Administration also issued similar directives.

Both Kern (1979) and Duffy (1985) urge the military to abandon use of the formulas. They note that writers in the military often find the task of simplifying texts below the 10th grade “too difficult” and “not worth the trouble.”

Unfortunately, there are no practical alternatives to the skill hard work required to create simple language. When large numbers of readers are involved, even small increases in comprehension pay off.

The Hull formula for technical writing At the 1979 STC conference, Leon C. Hull (1979) argued that technical writing, with its increased use of difficult words, needs a special kind of formula. While acknowledging that the FORCAST and Kincaid formulas were developed precisely for that reason, he looked for a formula that does not use word length as a variable.

Basing his work on Bloomer (1959) and Bormuth (1969) as well as his own experience as a technical writer, Hull claims that an increase in the number of adjectives and adverbs before a noun lowers comprehension. His study indicates that the modifier load is almost as predictive as a syllable count, more causal, and more helpful for rewriting.

Hull devised four cloze tests of each of five criterion passages from the Kincaid study. The first test was the original passage. Each of the other tests increased one of three indicators of modifier load by at least 50%: density of modifiers, ambiguity of modifiers, and density of prepositions. The subjects were 107 science, engineering, and management students enrolled in a senior course in technical and professional communication at Rensselaer Polytechnic Institute.

The mean cloze scores on the five unaltered passages correlated (r =) 0.882 with the Kincaid reading-grade levels assigned to these passages. This result justified both the subject sampling and the use of the test results to produce a new formula. The test results confirm the negative effect (r = -0.664) of modifier density on comprehension. They also indicated that sentence length is a valid indicator for technical material, perhaps better than word difficulty (contrary to previous research).

Hull developed first formula with five variables, which accounts for (r^2 =) 68% of passage difficulty. Like others before him, he found that the difficulty of using a larger number of variables reduces the reliability of the formula and makes it impractical. He created a another formula, shown here, that uses only sentence length and the density of modifiers (called prenomial modifiers) and accounts for (r^2 =) 48% of passage difficulty. Though slightly less valid than the Kincaid formula, it is as accurate as many other popular formulas:

Grade level = 0.49 (average sentence length)
+ 0.29 (prenomial modifiers per 100 words) – 2.71

In the conclusion of his paper, Hull advises technical writers that using shorter sentences reduces their complexity and makes them easier to read. He also recommends eliminating strings of nouns, adjectives, and adverbs as modifiers. Instead, writers should use prepositional phrases and place adjectives in the predicate position (after the verb) rather than in the distributive position (before the noun).

Degrees of Reading Power (DRP) In 1981, the College Entrance Examination Board dropped its use of grade-level reading scores and adopted the Degrees of Reading Power (DRP) system developed by Touchstone Applied Science Associates (Koslin et al. 1987, Zeno et al. 1995).

The DRP uses the Bormuth Mean Cloze formula to predict scores on a 0 (easy) to 100 (difficult) scale, which can be used for scoring both text readability and student reading skills. The popular children's book *Charlotte's Web* has a DRP value of 50. Likewise, students with DRP test scores of 50 (at the independent level) are capable of reading *Charlotte's Web* and easier texts independently. The Board also uses this system to provide readability reports on instructional materials used by school systems.

Computerized writing aids Beginning in the 1980s, the first computer programs appeared that not only contained the formulas but also other writing aids. The Writer's Workbench, developed at Bell Laboratories became the most popular of these (Macdonald, Frase, Gingrich, and Keenan 1982). It contains several readability indexes, stylistic analysis, average lengths of words and sentences, spelling, punctuation, faulty phrases, percentages of passive verbs, a reference on English usage, and many other features.

Kincaid, Aagard, O'Hara, and Cottrell (1981) developed CRES, a computer readability editing system for the U.S. Navy. It contains a readability formula and flags uncommon words, long sentences, and offers the writer alternatives.

Today, popular word processors such as Microsoft Word and Corel WordPerfect include a combination of spell checkers, grammar checkers, and readability formulas to help in creating texts that are more readable. Note that the Flesch-Kincaid Grade Level in Word's Readability Statistics is defective in that it only goes to the 12th grade.

Lexile Framework At the height of the controversy about the readability formulas, the founders of MetaMetrics, Inc. (Stenner, Horabin, et al. 1988a) published a new system for measuring readability, Lexile Framework, which uses average sentence length and average word frequency found in the *American Heritage Intermediate Corpus* (Carroll et al. 1971) to predict a score on a 0–2000 scale. The AHI corpus includes five million words from 1,045 published titles to which students in grades three through nine are commonly exposed.

The cognitive theorists had claimed that different kinds of reading tests actually measure different kinds of comprehension. The studies of the Lexile theorists (Stenner et al. 1988b, Stenner and Burdick 1997) indicate that comprehension is a one-dimensional ability that subsumes different types of comprehension (e.g., literal or inferential) and other reader factors (e.g., prior knowledge and special subject knowledge). You either understand a passage or you don't.

The New Dale-Chall Readability Formula In *Readability Revisited: The New Dale-Call Readability Formula*, Chall and Dale (1995) updated their list of 3,000 easy words and improved their original formula, then 47 years old. The new formula was validated against a variety of criteria, including:

- 32 passages tested by Bormuth (1971) on 4th to 12th-grade students.
- 36 passages tested by Miller and Coleman (1967) on 479 college students.
- 80 passages tested by MacGinitie and Tretiak (1971) on college and graduate students.
- 12 technical passages tested by Caylor et al. (1973) on 395 Air Force trainees.

The new formula was also cross-validated with:

- The Gates-MacGinitie Reading Test
- The Diagnostic Assessments of Reading and Trial Teaching Strategies (DARTTS).
- The National Assessment of Reading Progress.
- The Spache Formula.
- The Fry Graph.
- Average judgments of teachers on the reading level of 50 passages of literature.

Fig 11. Jeanne S. Chall created the Harvard Reading Lab and directed it for 20 years.

The new formula correlates .92 with the Bormuth Mean Cloze Scores, making it the most valid of the popular formulas.

At the time of writing this, the new Dale-Chall formula is not yet available on the Internet. It was once available in a computer program, "Readability Master," but is hard to find. You can easily apply the formula manually, however, using the instructions, worksheet, word list, and tables provided in the book. The book also has several chapters reviewing readability research, the uses of the formulas, the importance of vocabulary, the readability controversies, and a chapter on writing readable texts.

The following are two of the sample passages in the book, with the difficult words not found on their new word list underlined (pp. 135-140). The right-hand column gives a few readability statistics, the New Dale-Chall mean cloze score, and reading grade level.

Grades 5-6

Eskimos of Alaska's Arctic north coast have hunted whales for centuries.

Survival has depended on killing the 80-foot-long bowhead whales that swim from the Bering Sea to the ice-clogged Beaufort Sea each Spring. The Eskimos' entire way of life has been centered around the hunt.

But now that way of life is being threatened by America's need for oil, say many Eskimos who hunt the whales.

Huge amounts of oil may be beneath the Beaufort Sea. And oil companies want to begin drilling this spring.

However, many Eskimos say severe storms and ice conditions make drilling dangerous…

From *My Weekly Reader, Edition 6*

Readability Data	
Number of Words in Sample	100
Number of Whole Sentences	6
Number of Unfamiliar Words	11
Cloze Score	42
Reading Level	5-6

Grades 9-10

The controversy over the laser-armed satellite boils down to two related questions: Will it be technically effective? And should the United States make a massive effort to deploy it?

To its backers, the laser seems the perfect weapon. Traveling in a straight line at 186,000 miles per second, a laser beam is tens of thousands of times as fast as any bullet or rocket. It could strike its target with a power of many watts per square inch. The resulting heat, combined with a mechanical shock wave created by recoil as surface layers were blasted away, could quickly melt…

From *Discover*

Readability Data	
Number of Words in Sample	100
Number of Whole Sentences	5
Number of Unfamiliar Words	23
Cloze Score	28
Reading Level	9-10

ATOS readability formula for books Researchers at School Renaissance Institute (1999, 2000, Paul 2003) and Touchstone Applied Science Associates produced the Advantage-TASA Open Standard (ATOS) Readability Formula for Books. Their goal was to create an "open" formula that would be available to the educational community free of charge, that would be easy to use, and that could be used with any nationally normed reading tests.

The project was perhaps the most extensive study of readability ever conducted. Formula developers used 650 norm-referenced reading tests, 474 million words representing all the text of 28,000 K–12 books read by real students with many published in the previous five years, an expanded vocabulary list, and the reader records of more than 30,000 students who read and tested on 950,000 actual books.

The readability formula was part of a computerized system to help teachers conduct a program of guided independent reading to maximize learning gains. Noting the differences in difficulty between samples and entire books, the developers claim this is the first readability formula based on whole books, not just samples.

They found that the combination of three variables gives the best account of text difficulty: words per sentence ($r^2 = .897$), the average grade-level of words ($r^2 = .891$), and characters per word ($r^2 = .839$). The formula produces grade-level scores, as they are easier for teachers to understand and use.

The formula developers paid special attention to the Zone of Proximal Development (ZPD) proposed by Vygotsky (1978), the level of optimal difficulty that produces the most learning gain. They found that, for independent reading below the 4th grade, maximum learning gain requires at least 85% comprehension. Advanced readers need a 92% score on reading quizzes. Those who exceed that percentage should be given material that is more challenging.

Other results of the studies indicate that:

- Maximum learning gain requires careful matching of book readability and reading skill.
- The amount of time spent reading correlates highly with gains in reading skill.
- Book length can be a good indication of readability.
- Feedback and teacher interaction are the most important factors in accelerated reading growth.

Formula applications

Many researchers outside the field of reading have recognized the value of the formulas. Edward Fry (1986) points out that articles on the readability formulas are among the most frequently cited articles of all types of educational research. The applications give researchers an objective means of controlling the difficulty of passages in their experiments.

The following is a sample of readability studies that used formulas: political literature (Zingman 1977), corporate annual reports (Courtis 1987), customer service manuals (Squires and Ross 1990) drivers' manuals (Stahl and Henk 1995), dental health information (Alexander 2000), palliative-care information (Payne et al. 2000), research consent forms (Hochhauser 2002; Mathew 2002; Paasche-Orlow et al. 2003), informed consent forms (Williams et al. 2003) online health information (Oermann and Wilson 2000), lead-poison brochures (Endres et al. 2002) online privacy notices (Graber et al. 2002) medical journals (Weeks and Wallace 2002), environmental health information (Harvey and Fleming 2003) and mental-health information (King et al. 2003).

Court actions and legislation Fry (1989a) points out that the validity of the formulas has been challenged in court and found suitable for legal purposes. The courts increasingly rely on readability formulas to show the readability of texts in protecting the rights of citizens to clear information. Court cases and legislation involving government documents and correspondence, criminal

rights, product labeling, private contracts, insurance policies, ballot measures, warranties, and warnings are some of the legal applications of the formulas.

In 1984, Joseph David of New York was upset by his inability to understand a letter of denial he received in response to his appeal for Medicare benefits. Legal Services went to court in behalf of David and other elderly recipients of Medicare in New York. They pointed out that 48% of the population over 65 had less than a 9th-grade education. Edward Fry testified in court that the denial letter was written at the 16th-grade level. As a result, the judge ordered the Secretary Heckler of the U.S. Department of Health and Social Services to take "prompt action" to improve the readability of Medicare communications (David vs. Heckler 1984).

A number of federal laws require plain language such as the Truth in Lending Act, the Civil Rights Act of 1964, and the Electronic Funds Transfer Act. In June 1998, President Clinton directed all federal agencies to issue all documents and regulations in plain language.

Beginning in 1975, a number of states passed plain-language laws covering such common documents as bank loans, insurance policies, rental agreements, and property-purchase contracts. These laws often state that if a written communication fails the readability requirement, the offended party may sue and collect damages. Such failures have resulted in court judgments.

States such as California also require plain language in all agency documents, including "any contract, form, license, announcement, regulation, manual, memorandum, or any other written communication that is necessary to carry out the agency's responsibilities under the law" (Section 6215 of the California Government Code). California defines plain language as "written or displayed so that the meaning of regulations will be easily understood by those persons directly affected by them" (Section 11349 of the Administrative Code).

Textbook publishers After 80 years, textbook publishers consider the grade level of textbooks as more important than cost, the choice of personnel, or the physical features of books. All of them use word-frequency lists. Eighty-nine percent of them use readability formulas in evaluating the grade-levels of texts, along with other methods of testing. Widely read children's publications such as *My Weekly Reader* and magazines published by National Geographic for children of different ages have used the formulas along with field-testing and other methods (Chall and Conard 1991).

Using the formulas

Formula discrepancies The discrepancy between the scores of different formulas has long been perplexing. For example, the scores for the following four paragraphs are:

Corrected Dale-Chall grade level: 13-15
Flesch-Kincaid grade level: 12.5
FORCAST grade level: 11.2
SMOG grade level: 14.5
Fog grade level: 16.3
Fry Readability Graph: 17+

Critics have often cited such discrepancies as indications of the lack of precision of the formulas. Kern (1979) argued that the discrepancies among the Kincaid and Caylor formulas deprive them of usefulness, and that the military should discard them. What Kern ignores in his review are the correlations of the formulas with comprehension tests. What is important is not how the formulas agree or disagree on a particular text, but their degree of consistency in predicting difficulty over a range of graded texts.

The most obvious causes of the discrepancies are the **different variables** used by different formulas. For example, some use the number of syllables per word and others use the number of letters per word. The FORCAST formula uses a sentence-length variable only, no word variable.

Another important difference is that formulas use **different criterion scores**. The formulas—like reading tests—simply do not have a common zero point (Klare 1982). The criterion score is the required level of comprehension as indicated by the percentage of correct answers on a reading test. For example, one formula might predict the level of reading skill required to answer correctly 75 % of the questions on a reading test based on a criterion passage. Another formula might predict the reading level of a class that can correctly answer 50% of the questions on a reading test.

The FORCAST formula and Dale-Chall formulas uses a 50% criterion score as measured by multiple-choice tests. The Flesch use a 75% score, Gunning Fog formula, a 90% score, and the McLaughlin SMOG formula, a 100% score. The formulas developed with the higher criterion scores tend to predict higher scores, while those the highest validity correlations (e.g., Gunning and Flesch) tend to predict lower scores.

The **different algorithms** used by different computer programs to count sentences, words, and syllables can also cause discrepancies—even though they use the same formula.

Finally, the range of scores provided by different formulas remind us that they are not perfect predictors. They provide probability statements or, rather, rough estimates (r^2 = .50 to .84) of text difficulty. That means the readability formulas account for 50 to 84 percent of the variance in text difficulty as measured by comprehension tests..

The problem of optimal difficulty Different uses of a text require different levels of difficulty. As we have seen, Bormuth (1969) indicated the 35% cloze score was the point of optimum learning gain (see Table 7 above) for assisted classroom reading.

Vygotsky (1978) supported Bormuth's findings that optimal difficulty should be slightly above their current level of development and not below. Using books that are at the reader's present level or below may increase fluency and rate, but not in the way of comprehension.

For this reason, experts advise that materials intended for assisted reading when an instructor is available should be somewhat harder than the readers' tested reading level. Materials for the general public, however, such as medicine inserts, instructions for filing tax forms, instructions for using appliances, and health information should, be as easy as possible (Chall and Dale 1995).

Paul (2003) found that independent reading requires at least an 85% comprehension on multiple-choice reading quizzes for readers below the 4th grade and 92% for advanced readers. He also recommends that advanced students who score better than 92% correct on quizzes should be given material that is more challenging.

From this evidence, we can tentatively conclude that for texts intended for classroom, training, and other forms of assisted reading, the Dale-Chall (50% correct criteria) is the preferable formula to use. For unassisted reading, especially where health information and safety issues are involved, then the Flesch (75%) and Gunning (90%) formulas may be more effective.

The formulas and usability testing Redish (2000) and Schriver (1991, 2000), promote the need for reading protocols and usability testing as an alternative to the formulas. They feel that usability testing eliminates the need for readability

testing. They fail to state, however, how to match the reading ability of subjects with that of the target audience.

Dumas and Redish (1999), in their work on usability testing, hardly mention reading comprehension. They have us assume that, if test subjects correctly perform a task, they have correctly understood the instructions. When problems arise, however, it is difficult to locate the source of the difficulty.

In both usability testing and reading protocols, some subjects are more skilled than others in articulating the problems they encounter. Do problems come from the text or from some other source? If they are located in the text, do they come from the design, style, organization, coherence, or content? We are often left with guesswork and trial-and-error cycles of revision and testing.

As experienced writers know, this gets expensive. In preparation for a test, it makes as little sense to neglect the readability of a document as it does to neglect its punctuation, grammar, coherence, or organization.

One cannot emphasize enough the importance of testing and of frequent contacts with members of the targeted audience before, during, and after the process of producing documents as urged by Schriver (1997) and Hackos and Redish (1998). Assessing both the reading ability of the audience and the readability of the text will greatly facilitate this process.

Conclusion

Today, the readability formulas are more popular than ever. There are readability formulas for Spanish, French, German, Dutch, Swedish, Russian, Hebrew, Hindi, Chinese, Vietnamese, and Korean (Rabin 1988).

The formulas have survived 80 years of intensive application, investigation, and controversy, with both their credentials and limitations remaining intact. The national surveys on adult literacy have re-defined our audience for us. Any approach to effective communication that ignores these important lessons cannot claim to be scientific. If we walk away from this research, others will one day rediscover it and apply it to our work as technical communicators.

The variables used in the readability formulas show us the skeleton of a text. It is up to us to flesh out that skeleton with tone, content, organization, coherence, and design. Gretchen Hargis of IBM (2000) states that readability research has made us very aware of what we "write at the level of words and sentences." She writes:

> Technical writers have accepted the limited benefit that these measurements offer in giving a rough sense of the difficulty of material.
>
> We have also assimilated readability as an aspect of the quality of information through its pervasiveness in areas such as task orientation, completeness, clarity, style, and visual effectiveness. We have put into practice, through user-centered design, ways to stay focused on the needs of our audience and their problems in using the information or assistance that we provide with computer products.

The research on literacy has made us aware of the limited reading abilities of many in our audience. The research on readability has made us aware of the many factors affecting their success in reading. The readability formulas, when used properly, help us increase the chances of that success.

References

Alexander, R. E. 2000. “Readability of published dental educational materials.” *Journal of America dental association* 7:937-943.

Armbruster, B. B. 1984. “The problem of inconsiderate text.” In *Comprehension instruction,* ed. G. Duffey. New York: Longmann, p. 202-217.

Barr, R. and R. Dreeben. 1984. “Grouping students for reading instruction.” In *Handbook of reading research,* ed. P. D. Pearson. New York: Longman, pp. 885-910.

Benson, R. W. 1984-1985. “The end of legalese: The game is over.” *Review of law and social change* 13, no. 3:519-573.

Betts, E. 1946. *Foundations of reading instruction*. New York: American Book Company.

Bloomer, R. H. “Level of abstraction as a function of modifier load.” *Journal of educational research* 52:269-272.

Bormuth, J. R. 1966. “Readability: A new approach.” *Reading research quarterly* 1:79-132.

Bormuth, J. R. 1969. *Development of readability analysis*. Final Report, Project no. 7-0052, Contract No. OEC-3-7-070052-0326. Washington, D. C.: U.S. Office of Education, Bureau of Research, U.S. Department of Health, Education, and Welfare.

Bormuth, J. R. 1971. *Development of standards of readability: Towards a rational criterion of passage performance.* Washington, D. C.: U.S. Office of Education, Bureau of Research, U.S. Department of Health, Education, and Welfare.

Bruce, B., A Rubin, and K. Starr. 1981. *Why readability formulas fail: Reading Education Report No. 28*. Champaign, IL: University of Illinois at Urbana-Champaign.

Buswell, G. 1937. “How adults read.” *Supplementary educational monographs Number 45.* Chicago: University of Chicago Press. In Sticht and Armstrong 1994, pp. 43-50.

Calfee, R. C. and R. Curley. 1984. “Structures of prose in content areas.” In *Understanding reading comprehension,* ed. J Flood. Newark, DE: International Reading Association, pp. 161-180.

Carroll, J. B. 1987. “The national assessments in reading: are we misreading the findings?” *Phi delta kappan* pp. 414-430.

Carroll, J. B., P. Davies, and B. Richman. 1971. *The American Heritage Intermediate Corpus.* New York: American Heritage Publishing Co.

Carver, R. P. 1975-1976. “Measuring prose difficulty using the Rauding scale.” *Reading research quarterly* 11:660-685.

Carver, R. P. 1990. “Predicting accuracy of comprehension from the relative difficulty of material.” *Learning and individual differences* 2: 405-422.

Caylor, J. S., T. G. Sticht, L. C. Fox, and J. P. Ford. 1973. *Methodologies for determining reading requirements of military occupational specialties*: *Technical report No. 73-5*. Alexandria, VA: Human Resources Research Organization.

Chall, J. S. 1958. *Readability: An appraisal of research and application.* Columbus, OH: Ohio State University Press. Reprinted 1974. Epping, Essex, England: Bowker Publishing Company.

Chall, J. S. 1967, 1983. *Learning to read: The great debate*. New York: McGraw-Hill.

Chall, J. S. 1983b. *Stages of reading development.* New York: McGraw-Hill.

Chall, J. S. 1984. "Readability and prose comprehension: Continuities and discontinuities." In *Understanding reading comprehension: Cognition, language, and the structure of prose,* ed. J. Flood. Newark, DE: International Reading Association.

Chall, J. S. 1988. "The beginning years." In *Readability: Its past, present, and future*, eds. B. L. Zakaluk and S. J. Samuels. Newark, DE: International Reading Association.

Chall, J. S., and S. S. Conard. 1991. *Should textbooks challenge students? The case for easier or harder textbooks*. New York: Teachers College Press.

Chall, J. S., G. L. Bissex, S. S. Conard, and S. Harris-Sharples. 1996. *Qualitative assessment of text difficult: A practical guide for teachers and writers*. Cambridge, MA: Brookline Books.

Chall, J. S. and E. Dale. 1995. *Readability revisited, the new Dale-Chall readability formula*. Cambridge, MA: Brookline Books.

Charrow, V. R. 1977. *Let the re-writer beware.* Arlington, VA: Center for Applied Linguistics.

Charrow, R. P. and V. R. Charrow. 1979. "Making legal language understandable: A psycholinguistic study of jury instructions." *Columbia law review* 79:1306-1347.

Chiesi, H. L., G. J. Spilich, and J. F. Voss. 1979. "Acquisition of domain-related information in relation to high and low domain knowledge." *Journal of verbal learning and verbal behavior* 18: 257-273.

Clark, H. H. and S. E. Haviland. 1977. "Comprehension and the given-new contract." In *Discourse Production and Comprehension,* ed. R. O. Freedle. Norwood NJ: Ablex Press, pp. 1-40.

Clay, M. 1991. *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.

College Entrance Examination Board. 1980 and later editions. *Degrees of reading power (DRP)*. Princeton, NJ: College Entrance Examination Board.

Coleman, E. B. 1962. "Improving comprehensibility by shortening sentences." *Journal of Applied Psychology* 46:131.

Coleman, E. B. 1964. "The comprehensibility of several grammatical transformations." *Journal of applied psychology* 48:186-190.

Coleman, E. B. 1965. "On understanding prose: some determiners of its complexity." NSF Final Report GB-2604. Washington, D.C.: National Science Foundation.

Coleman, E. B. 1966. "Learning of prose written in four grammatical transformations." *Journal of applied psychology* 49:332-341.

Coleman, E. B. 1971. "Developing a technology of written instruction: some determiners of the complexity of prose." In *Verbal learning research and the technology of written instruction*, eds. E. Z. Rothkopf and P. E. Johnson. New York: Teachers College Press, Columbia University, pp. 155-204.

Coleman, E. B. and P. J. Blumenfeld. 1963 “Cloze scores of nominalization and their grammatical transformations using active verbs.” *Psychology reports* 13:651-654.

Connaster, B. F. 1999. “Last rites for readability formulas in technical communication.” *Journal of technical writing and communication*. 29, no. 3:271-287.

Coupland, N. 1978. “Is readability real?” *Communication of scientific and technical information.* April:15-17.

Courtis, J. K. 1987. “Fry, Smox, Lix and Rix: Insinuations about corporate business communications.” *Journal of business communications.* 24, no. 2:19-27.

Dale, E. 1967. *Can you give the public what it wants?* New York: World Book Encyclopedia.

Dale, E. and R. W. Tyler. 1934. “A study of the factors influencing the difficulty of reading materials for adults of limited reading ability.” *Library Quarterly* 4:384-412.

Dale, E. and J. S. Chall. 1948. “A formula for predicting readability.” *Educational research bulletin* Jan. 21 and Feb. 17, 27:1-20, 37-54.

Dale, E. and J. S. Chall. 1949. “The concept of readability.” *Elementary English* 26: 23.

Dale, E. and J. O’Rourke. 1981. *The living word vocabulary: A national vocabulary Inventory*. Chicago: World Book–Childcraft International.

Davison, A. 1984. “Readability formulas and comprehension.” In *Comprehension instruction: Perspectives and suggestions,* eds. G. G. Duffy, L. R. Roehler, and J. Mason. New York: Longman.

Davison, A. 1986. *Readability: The situation today.* Reading education report No. 70. Champaign, IL: Center for the Study of Reading. University of Illinois at Urbana-Champaign.

Department of the Air Force. 1977. *Publications management: Writing understandable publications* (HQ Operating Instruction 5-2). Washington, DC: Headquarters U.S. Air Force. 25 March.

Department of Defense. 1978. *Manuals, technical: General style and format requirements.* Military Specification MIL-M-3784A (Amendment 5), 24 July.

Dickinson, T. 1987. *Exploring the night sky.* Ontario, Canada: Camden East.

Doak, C. C., L. G. Doak, and J. H. Root**.** 1996. *Teaching patients with low literacy skills*. Philadelphia: J. B. Lippincott Company.

Dolch, E. W. 1939. “Fact burden and reading difficulty.” *Elementary English review* 16:135-138.

Decina, L. E. and K. Y. Knoebel. 1997. *"*Child safety seat misuse patterns in four states*.*" *Accident analysis and prevention* 29: 125-132.

Duffy, T. M. 1985. “Readability formulas: What’s the use?” In *Designing usable texts*, eds. T. M. Duffy and R. M. Waller. New York: Academic Press, pp. 113-143.

Duffy, T. M. and P. Kabance. 1981. “Testing a readable writing approach to text revision.” *Journal of educational psychology* 74, no. 5:733-748.

Dumas, J. and J. Redish. 1999. *A practical guide to usability testing*. Exeter UK: Intellect.

Elwork, A., B. Sales, and J. Alfini. 1982. *Making jury instructions understandable*. Charlottesville, VA: Michie Company.

Endres, J., J. Montgomery, and P. Welch. "Lead poison prevention: A comparative review of brochures." *Journal of environmental health* 64, no. 6:20-25.

Entin, E. B. and G. R. Klare 1985. "Relationships of measures of interest, prior knowledge, and readability to comprehension of expository passages." *Advances in reading/language research* 3:9-38.

Ewing, M. J. 1976. *A comparison of the effects of readability and time on learning the contents of a state driver's handbook*. Unpublished doctoral dissertation, Florida State University.

Fang, I. E. 1966-1967. "The 'easy listening formula.'" *The journal of broadcasting* 11:63-68.

Farr, J. N., J. J. Jenkins, and D. G. Paterson. 1951. "Simplification of the Flesch Reading Ease Formula." *Journal of applied psychology* 35, no. 5:333-357.

Fass, W. and G. M. Schumacher. 1978. "Effects of motivation, subject activity, and readability on the retention of prose materials." *Journal of educational psychology* 70:803-808.

Faufmann, W. J., III. 1990. *Discovering the universe*, 2nd Ed. New York: W. H. Freeman.

Feld. B., Jr. 1948. "Empirical test proves clarity adds readers." *Editor and publisher* 81:38.

Felker, D. B., F. Pickering, V. R. Charrow, V. M. Holland, and J. C. Redish. 1981. *Guidelines for document designers*. Washington, DC: American Institutes for Research.

Fishburne, R. P. 1976 *Readability and reading ability as predictors of comprehension and learning time in a computer managed instructional system*. Doctoral dissertation, College of Education, Memphis State University.

Flesch, R. 1943. "Marks of a readable style." *Columbia University contributions to education*, no. 897. New York: Bureau of Publications, Teachers College, Columbia University.

Flesch, R. 1946. *The art of plain talk*. New York: Harper.

Flesch, R. 1948. "A new readability yardstick." *Journal of Applied Psychology* 32:221-233.

Flesch, R. 1949 and 1974. *The art of readable writing*. New York: Harper.

Flesch, R. 1951. *The art of clear thinking*. New York: Harper.

Flesch, R. 1955. *Why Johnny can't read—And what you can do about it*. New York: Harper.

Flesch, R. 1964. *The ABC of style: A guide to plain English*. New York: Harper.

Flesch, R. 1979. *How to write in plain English: A book for lawyers and consumers.* New York: Harper.

Fountas, I.C., and G. S. Pinnell. 1999. *Matching books to readers: Using leveled books in guided reading*. Portsmouth, NH: Heinemann.

Freebody, P. and R. C Anderson. 1983. "Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension." *Reading research quarterly* 18: 277-294.

Fry, E. B. 1963. *Teaching faster reading.* London: Cambridge University Press.

Fry, E. B. 1968. "A readability formula that saves time." *Journal of reading* 11:513-516.

Fry, E. B. 1969. "The readability graph validated at primary levels." *The reading teacher*. 22:534-538.

Fry, E. B. 1977. "Fry's readability graph: Clarifications, validity, and extension to level 17." *Journal of Reading* 21, no. 3:242-252.

Fry, E. B. 1986. *Varied uses of readability measurement*. Paper presented at the 31st Annual Meeting of the International Reading Association, Philadelphia, PA.

Fry, E. B. 1988. "Writeability: the principles of writing for increased comprehension." In *Readability: Its past, present, and future,* eds. B. L Sakaluk and S. J. Samuels. Newark, DE: International Reading Association.

Fry, E. B. 1989a. *The legal aspects of readability.* Paper presented at the 34th Annual Meeting of the International Reading Association, New Orleans.

Fry, E. B. 1989b. "Reading formulas: maligned but valid." *Journal of reading* 32, no. 4:292-297.

Fry, E. B. 2002. "Readability versus leveling." *Reading Teacher* 56, no. 3:286-292.

Fry, E.B., J. E. Kress, and D. L. Fountoukidis. 1993. *The reading teacher's book of lists: Third edition*. West Nyack, NY: The Center for Applied Research in Education.

Gates, A. I. 1930. *Interest and ability in reading.* New York: Macmillan.

Gilliland, J. 1972. *Readability*. London: Hodder and Stoughton.

Gough, P. B. 1965. "Grammatical transformations and the speed of understanding." *Journal of verbal learning and verbal behavior* 4:107-111.

Graber, M. A., D. M. D'Alessandro, and J. Johnson-West. 2002. "Reading level of privacy policies on internet health web sites." *Journal of Family Practice* 51, no. 7:642-647.

Gray, W. S. and B. Leary. 1935. *What makes a book readable*. Chicago: Chicago University Press.

Green. M. T. 1979. *Effects of readability and directed stopping on the learning and enjoyment of technical materials*. Unpublished doctoral dissertation, University of South Carolina.

Gunning, R. 1952. *The technique of clear writing,* New York: McGraw-Hill.

Hackos, J. and J. Redish 1998. *User and task analysis for interface design*. New York: Wiley.

Hackos, J. and D. M. Stevens. 1997. *Standards for online communication*. New York: Wiley.

Halliday, M. and R. Hasan 1976. *Cohesion in English*. London: Longman.

Halbert, M. G. 1944. "The teaching value of illustrated books." *American school board journal* 108, no. 5:43-44.

Hardyck, C. D. and L. F. Petrinovich 1970. "Subvocal speech and comprehension level as a function of the difficulty level of reading material." *Journal of verbal leaning and verbal behavior* 9:647-652.

Hargis, G., A. K. Hernandez, P. Hughes, J. Ramaker, S. Rouiller, and E. Wilde. 1998. *Developing quality technical information: A handbook for writers and editors.* Upper Saddle River, NJ: Prentice Hall.

Hargis, G. 2000. "Readability and computer documentation." *ACM journal of computer documentation 24, no. 3:122-131.*

Harris, A. J. and M. D. Jacobson. 1973. *The Harris-Jacobson primary readability formulas.* Paper presented at the Annual Meeting of the International Reading Association, Denver, CO.

Harvey, H. D. and P. Fleming. 2003. "The readability and audience acceptance of printed health promotion materials used by environmental health departments." *Journal of environmental health*. 65, no. 6:22-29.

Hochhauser, M. 2002. "The effects of HIPAA on research consent forms." *Patient care management* 17, no. 5:6-7.

Hooke, L. R. et al. 1979. *Readability of Air Force publications: A criterion referenced evaluation.* Final Report. AFHRL-TR-79-21. Washington, D.C.: U. S. Air Force. Eric No: ED177512.

Horn, E. 1937. *Methods of instruction in the social studies.* New York: Charles Scribner's and Sons. (Report of the Commission on Social Studies, American Historical Association).

Hornby, P. A. 1974. "Surface structure and presupposition." *Journal of verbal learning and verbal behavior* 13:530-538.

Huckin, T. N., E. H. Curin, and D. Graham. 1991. "Prescriptive linguistics and plain English: The case of "whiz-deletions." In *Plain language: Principles and practice*, ed. E. Steinberg. Detroit: Wayne State University Press.

Hull, L. C. 1979. "Measuring the readability of technical writing." *Proceedings of the 26th International Technical Communications Conference*, Los Angeles: E79 to E84.

Hunt, K. 1965. *Grammatical structures written at three grade levels*. National council of Teachers of English Report No. 3. Urbana, IL: National Council of Teachers of English.

Johnson, W. 1946. *People in Quandaries*. New York: Harpers, Appendix.

Johnston, C., F. P. Rivara, and R. Soderberg. 1994. "Children in Car Crashes: Analysis of Data for Injury and Use of Restraints." *Pediatrics* 93, no. 6 pt 1: 960-965.

Kahane, C. 1986. *An evaluation of child passenger safety: the effectiveness and benefits of safety seats*. Washington, D. C.: National Highway Traffic Safety Administration.

Kemper, S. 1983. "Measuring the inference load of a text." *Journal of educational psychology*. 75:3, 391-401.

Kern, R. P. 1979. "*Usefulness of readability formulas for achieving Army readability objectives: Research and state-of-the-art applied to the Army's problem.* Fort Benjamin Harrison, ID: Technical Advisory Service, U.S. Army Research Institute. (NTIS No. AD A086 408/2).

Kincaid, J. P., J. A. Aagard, J. W. O'Hara, and L. K. Cottrell. 1981. "Computer readability editing system." *IEEE transactions on professional communications* March.

Kincaid, J. P., R. P. Fishburne, R. L. Rogers, and B. S. Chissom. 1975. *Derivation of new readability formulas (Automated Readability Index,*

Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. CNTECHTRA Research Branch Report 8-75.

King, M. M., A. S. W. Winton, and A. D. Adkins. 2003. "Assessing the readability of mental health internet brochures for children and adolescents." *Journal of child and family studies* 12, no. 1:91-100.

Kintsch, W. and E. Vipond. 1979. "Reading comprehension and readability in educational practice and psychological theory." In *Perspectives on memory research*, ed. L. G. Nilsson. Hillsdale, NJ: Erlbaum.

Kintsch, W. and J. R Miller. 1981. "Readability: A view from cognitive psychology." In *Teaching: Research reviews.* Neward, DE: International Reading Association.

Kitson, H. D. 1921. *The mind of the buyer*. New York: Macmillan.

Klare, G. R. 1952. "Measures of the readability of written communication: An evaluation." *The journal of educational Psychology* 43, no. 7:385-399.

Klare, G. R. 1957. "The relationship of typographic arrangement to the learning of technical training material." *Journal of applied psychology* 41, no. 1:41-45.

Klare, G. R. 1963. *The measurement of readability*. Ames, Iowa: Iowa State University Press.

Klare, G. R. 1968. "The role of word frequency in readability." *Elementary English*, 45:12-22.

Klare, G. R. 1974-75. "Assessing readability." *Reading research quarterly* 10: 62-102.

Klare, G. R. 1975. *A manual for readable writing.* Glen Burnie, MD: REMco. (revised 1980).

Klare, G. R. 1976. "A second look at the validity of the readability formulas." *Journal of reading behavior* 8:159-152.

Klare, G. R. 1977. "Readable technical writing: Some observations." *Technical communication* 24, no. 2:1-5.

Klare, G. R. 1980. *How to write readable English.* London: Hutchinson.

Klare, G. R. 1981. "Readability indices: do they inform or misinform?" *Information design journal 2:251-255.*

Klare, G. R. 1982. "Readability." *Encyclopedia of educational research* 3:1520-1531. New York: The Free Press.

Klare, G. R. 1984. "Readability." *Handbook of reading research*, ed. P. D. Pearson. New York: Longman, pp. 681-744.

Klare, G. R. 1985. "Matching reading materials to readers: The role of readability estimates in conjunction with other information about comprehensibility." In *Reading, thinking, and concept development,* eds. T. L. Harris and E. J. Cooper. New York: College Entrance Examination Board.

Klare, G. R. 2000. "Readable computer documentation." *ACM journal of computer documentation* 24, no. 3:148-168.

Klare, G. R. and B. Buck, 1954. *Know your reader, the scientific approach to readability.* New York: Hermitage House.

Klare, G. R., J. E. Mabry, and L. M. Gustafson. 1955a. "The relationship of style difficulty to immediate retention and to acceptability of technical material." *Journal of educational psychology* 46:287-295.

Klare, G. R., J. E. Mabry, and L. M. Gustafson. 1955b. "The relationship of patterning (underlining) to immediate retention and to acceptability of technical material." *Journal of applied psychology* 39, no. 1:40-42.

Klare, G. R., J. E. Mabry, and L. M. Gustafson. 1955c. "The relationship of immediate retention of technical training material to career preferences and aptitudes" *Journal of educational psychology* 46, no. 6:321-329.

Klare, G. R., E. H. Shuford, and W. H. Nichols. 1957. "The relationship of style difficulty, practice, and ability to efficiency of reading and retention." *Journal of applied psychology* 41:222-226.

Klare, G. R., P. P. Rowe, M. G. St. John, and L. M. Stolurow. 1969. "Automation of the Flesch reading ease readability formula, with various options." *Reading research quarterly* 4, no. 4:550-559.

Klare, G. R. and K. L. Smart. 1973. "Analysis of the readability level of selected USAFI instructional materials." *The journal of educational research* 67:176.

Knight, D. and J. D. Alcorn. 1969. "Comparisons of the performance of educationally disadvantaged adults and elementary children on selected measures of reading performance." *The nineteenth yearbook of the National Reading Conference.* Milwaukee, WI: Marquette University.

Knuth, R. A. and B.F. Jones. 1991. *What does research say about reading*. Oak Brook, IL: North Central Reading Educational Laboratory. http://www.ncrel.org/sdrs/areas/stw_esys/str_read.htm

Koslin, B. I., S. Zeno, and S. Koslin. 1987. *The DRP: An effective measure in reading*. New York: College Entrance Examination Board.

Lane, W. G., G. C. Liu and E. Newlin. 2000. "The Association between hands-on instruction and proper child safety seat installation." *Pediatrics* 2, no. 4: 924-929.

Lange, R. 1982. "Readability formulas: Second looks, second thoughts." *Reading Teacher* 35:858-861.

Lewerenz, A. S. 1929. "Measurement of difficulty of reading materials." *Los Angeles Educational Research Bulletin* 8:11-16.

Lewerenz, A. S. 1929a. "Objective measurement of the difficulty of reading materials." *Los Angeles Educational Research Bulletin* 9:8-11.

Lewerenz, A. S. 1935. "A vocabulary grade placement formula." *Journal of Experimental Education* 3:236.

Lewerenz, A. S. 1939. "Selection of reading materials by pupil ability and interest." *Elementary English Review* 16:151-156.

Lively, B. A. and S. L. Pressey. 1923. "A method for measuring the 'vocabulary burden' of textbooks. *Educational administration and supervision* 9:389-398.

Lorge, I. 1938. *The semantic count of the 570 commonest English words.* New York: Bureau of Publications, Teachers College, Columbia University.

Lorge, I. 1939. "Predicting reading difficulty of selections for children." *Elementary English Review* 16:229-233.

Lorge, I. 1944. "Predicting Readability" in the *Teachers College Record,* Vol. 45, 404-419.

Lorge, I. 1948. "The Lorge and Flesch Readability Formulae: A Correction." *School and Society*, Vol. 67, pp. 141-142.

McCall, W. A. and L. M. Crabbs. 1926, 1950, 1961, 1979. *Standard Test Lessons in Reading*. New York: Teachers College, Columbia University Press.

Macdonald, N. H., L. T. Frase, P. S. Gingrich, and S. A. Keenan. 1982. "The writer's workbench: Computer aids for text analysis." *Educational psychologist* 17:172-179.

MacGinitie, W. and R. Tretiak. 1971. "Sentence depth measures as predictors of reading difficulty." *Reading research quarterly* 6:364-376.

McLaughlin, G. H. 1968. "Proposals for British readability measures." *Third international reading symposium* eds. Brown and Downing. London: Cassell.

McLaughlin, G. H. 1969. "SMOG grading - a new readability formula," *Journal of reading* 22:639-646.

McLean, J. I. 1985. *Formative evaluation of the comprehensibility of written instruction.* Paper presented at the 36th Annual Meeting of the Conference on College Composition and Communication, Minneapolis, MN.

Manzo, A. 1970. "Readability: A postscript." *Elementary English* 47:962-965.

Markel, M. et al. 1992. "Effects of paragraph length on attitudes toward technical writing." *Technical communication* 39, no. 3:454-456.

Massad, C. E. 1977. *The role of decoding skills in adult functional reading.* Paper presented at the Annual Meeting of the American Educational Association, New York, April.

Mathew, J. 2002. "Readability of consent forms in schizophrenia research." *Australian and New Zealand journal of psychiatry* 36, no. 4:564-566.

Maxwell, M. 1978. "Readability: Have we gone too far?" *Journal of reading* 21:525-530.

Meyer, B. J. F. 1982. "Reading research and the composition teacher: The importance of plans." *College composition and communication* 33, no. 1:37-49.

Miller, G. R. and E. B. Coleman. 1967. "A set of thirty-six prose passages calibrated for complexity." *Journal of verbal learning and verbal behavior* 6:851-854.

Morriss, E. C. and D. Holversen. 1938. "Idea analysis technique." Unpublished manuscript. New York: Teachers College, Columbia University. In J. S. Chall. 1958.

Murphy, D. 1947. "How plain talk increases readership 45% to 60%." *Printer's ink* 220:35-37.

National Academy on an Aging Society. 2002. "Health literacy fact sheet." Accessed 12/06/02: http://www.agingsociety.org/healthlit.htm

National Assessment of Educational Progress (NAEP). 1985. *The Reading Report Card.* Princeton, NY: Educational Testing Service.

National Center for Education Statistics. 1993. *Adult literacy in America.* Washington, D. C.: U.S. Dept. of Education. http://www.nald.ca/fulltext/Report2/rep15-01.htm

National Center for Education Statistics. 1999. *Literacy in the work force*. Washington, D.C.: U.S. Dept. of Education. http://nces.ed.gov/pubs99/1999470.pdf

National Center for Education Statistics. 2001. *Adult literacy and education in America: Four studies based on the adult literacy survey*. Washington, D. C.: U.S. Dept. of Education. http://nces.ed.gov/pubs2001/2001534.pdf

National Center for Health Statistics. 2000. *National Mortality Data, 1998*. Hyattsville, MD: NCHS.

National Highway Traffic Safety Administration. 1996. *Observed Patterns of Misuse of Child Safety Seats*. Washington, D. C.: NHTSA.

Oermann, M. H. and F. E. Wilson. 2000. "Quality of health care information for consumers on the internet." *Journal of nursing care quality* 14, no. 4:45-55.

Ojemann, R. 1934. "The reading ability of parents and factors associated with the reading difficulty of parent education materials." In *Researchers in parent education* II. 19. (University of Iowa Studies, Studies in Child Welfare, VIII), ed. G. D. Stoddard. Iowa City, IA: University of Iowa.

Olsen, L. A. and R. Johnson. 1989. "Towards a better measure of readability: Explanation of empirical performance results." *Proceedings of the 36th International Technical Communication Conference.* Washington, D.C.: Society of Technical Communication.

Omaggio, A. C. 1979. *Pictures and reading: Some research results.* Paper presented at the Central States Conference on the Teaching of Foreign Languages. Minneapolis, MN, April 19-20, 1979.

Paasche-Orlow, M. K., H. A. Taylor, and F. L. Brancati. 2003. "Readability standards for informed-consent forms as compared with actual readability." *New England journal of medicine* 248, no. 8:721-727.

Patty, W. W. and W. I. Painter. 1931. "A technique for measuring the vocabulary burden of textbooks. *Journal of educational research* 24:127-134.

Paul, T. 2003. *Guided Independent Reading*. Madison, WI: School Renaissance Institute. http://www.renlearn.com/GIRP2008.pdf

Payne, S., S. Large, N. Jarrett, and P. Turner. 2000. "Written information given to patients and families by palliative care units: A national survey." *Lancet* 355, no. 9217:1792.

Pearson, P. D., J. Hansen, and C. Gordon. 1979. "The effect of background knowledge on young children's comprehension of explicit and implicit information." *Journal of reading behavior* 11:201-209.

Powers, R. D., W. A. Sumner, and B. E. Kearl. 1958. "A recalculation of four readability formulas." *Journalism quarterly* 49:99-105.

Rabin, A. T. 1988. "Determining difficulty levels of text written in languages other than English." In *Readability: Its past, present, and future*, eds. B. L. Zakaluk and S. J. Samuels. Newark, DE: International Reading Association.

Redish, J. C. 1988."Reading to learn to do." *The technical writing teacher*. 15, no. 3:323-333.

Redish, J. C. 2000. "Readability formulas have even more limitations than Klare discusses." *ACM journal of computer documentation* 24, no. 3:132-137.

Redish, J. C. and J. Selzer. 1985. "The place of readability formulas in technical communication." *Technical communication* 32, no. 4:46-52.

Richards, B. 1984. "Availability of prior knowledge and its effects on transfer of learning." *Proceedings of the Annual Meeting of the Association for Education Communications and Technology*, Dallas, TX, Jan. 20-24.

Rogers, J. R. 1962. "A formula for predicting the comprehension level of material to be presented orally." *The journal of educational research* 56:218-220.

Rothkopf, E. Z. 1972. "Structural text features and the control of processes in learning written materials," In *Language comprehension and the acquisition of knowledge,* eds. R.O. Freedle and J. B. Carroll. Washington, DC: Winston.

Rothkopf, E. Z. 1977. "Text difficulty predicted from functional chaining in short term memory." *Abstracts of papers and symposia.* American Educational Research Association annual meeting, Washington, D.C.

Rumelhart, D. E. 1984. "Understanding understanding." In *Understanding reading comprehension: cognition, language, and the structure of prose*, ed. J. Flood. Newark, DE: International Reading Association.

Russell, S. N. 1973. "Error pattern relationship of developmental readers and functionally illiterate adults." Ed.D. dissertation. Oklahoma State University.

School Renaissance Institute. 1999. *ZPD guidelines: Helping students achieve optimum reading growth.* Madison, WI: School Renaissance Institute, Inc.

School Renaissance Institute. 2000. *The ATOS readability formula for books and how it compares to other formulas.* Madison, WI: School Renaissance Institute, Inc.

Schramm, W. 1947. "Measuring another dimension of newspaper readership." *Journalism quarterly* 24:293-306.

Schriver, K. 1991. "Plain language through protocol-aided revision." In *Plain Language: Principles and practice*, ed. E. R. Steinberg. Detroit: Wayne State University Press.

Schriver, K. 1997. *Dynamics in document design.* New York: Wiley.

Schriver, K. 2000. "Readability formulas in the new millennium: What's the use?" *ACM journal of computer documentation* 24, no. 3:138-140.

Selzer, Jack. 1981. "Readability is a four-letter word." *Journal of business communication* 18, no. 4:23-34.

Sherman, A. L. 1893. *Analytics of literature: A manual for the objective study of English prose and poetry*. Boston: Ginn & Co.

Singer, H. 1975. "The SEER technique: A non-computational procedure for quickly estimating readability levels." *Journal of reading behavior* 3:255-267.

Smith, E. A. and J. P. Kincaid. 1970. "Derivation and validation of the automated readability index for use with technical materials." *Human factors* 12:457-464.

Smith, E. A. and R. J. Senter. 1967. *Automated readability index.* AMRL-TR-66-22. Wright-Patterson AFB, OH: Aerospace Medical Division.

Spache, G. 1953. "A new readability formula for primary-grade reading materials." *Elementary school journal*. 53:410-413.

Spilich, G. J., G. T. Vesonder, H. L. Chiesi, and J. F. Voss. 1979. "Text processing of domain-related information for individuals of high and

low domain knowledge. *Journal of verbal learning and verbal behavior* 18: 275-290.

Spyridakis, J. H. 1989. "Signaling effects: A review of the research-Part 1." *Journal of technical writing and communication.* 19, no. 3:227-240.

Spyridakis, J. H. 1989a. "Signaling effects: Increased content retention and new answers-Part 2." *Journal of technical writing and communication.* 19, no. 4:395-415.

Squires, P. and Ross, R. G. 1990. *Literacy requirements for customer service jobs.* Paper presented at the Annual Convention of the American Psychological Association, Boston, MA.

Stahl, N. A. and W. A. Henk. 1995. "Are drivers manuals understandable?" *Transportation quarterly* 49, no. 1:105-116.

Stanovich, K. E. 1984. "Word recognition: changing perspectives." In R. Barr, M. Kamil, P. B. Mosenthal, and P.D. Pearson, eds. *Handbook of reading research: Volume II.* New York: Longman.

Stenner, A. J. and D. S. Burdick. *The objective measurement of reading comprehension in response to technical questions raised by the California department of education technical study group*. Durham, NC: MetaMetrics, Inc.

Stenner, A. J., I. Horabin, D. R. Smith, and R. Smith. 1988a. *The Lexile Framework*. Durham, NC: Metametrics, Inc.

Stenner, A. J., I. Horabin, D. R. Smith, and R. Smith. 1988b. "Most comprehension tests do measure reading comprehension: A response to McLean and Goldstein." *Phi Delta Kappa*: 765-767.

Sticht, T. G. 1982. *Evaluation of the "Reading potential" concept for marginally literate adults: Technical report 82-2*. Washington, DC: Human Resources Research Organization.

Sticht, T. G. 1995. *The military experience and workplace literacy: A review and synthesis for policy and practice*. NCAL Technical Report TR94-01. http://www.literacyonline.org/products/ncal/pdf/TR9401.pdf

Sticht, T. G. and A. B. Armstrong. 1994. *Adult literacy in the United States: A compendium of quantitative data and interpretive comment.* http://www.nald.ca/fulltext/adlitUS/cover.htm.

Sticht, T. G., L. J. Beck, R. N. Hauke, G. M. Kleinman, and J. H. James. 1974. *Auding and reading.* Alexandria, VA: Human Resources Research Organization.

Swanson, C. C. 1979. *Readability and top-level structure: Effects on reading comprehension.* Unpublished doctoral dissertation, Arizona State University.

Swanson, C. E. 1948. "Readability and readership: A controlled experiment." *Journalism quarterly* 25:339-343.

Taylor, W. 1953. "Cloze procedure: A new tool for measuring readability." *Journalism quarterly* 30:415-433.

Thorndike, E. L. 1916. "An improved scale for measuring ability in reading." *Teachers college record* 17:40-67.

Thorndike, E. L. 1921. *The teacher's word book*. New York: Bureau of Publications, Teachers College, Columbia University.

Thorndike, E. L. 1932. *A teacher's word book of 20,000 words*. New York: Bureau of Publications, Teachers College, Columbia University.

Thorndike, E. L. and I. Lorge. 1944. *The teacher's word book of 30,000 words*. New York: Bureau of Publications, Teachers College, Columbia University.

Thorndike, R. L. 1973-1974. "Reading as reasoning." *Reading research quarterly* 9:135-137.

Ulijn, J. M. and J. B. Strother. 1990. "The effect of syntactic simplification on reading EST texts as L1 and L2." *Journal of research in reading* 13, no. 1:38-54.

Venezky, R. L. 1984. "The history of reading research." *Handbook of Reading Research*, ed. P. David Pearson. New York: Longmann.

Vernon, M. D. 1946. "Learning from graphic material." *British journal of psychology* 36:145-158.

Vogel, M. and C. Washburne. 1928. "An objective method of determining grade placement of children's reading material." *Elementary school journal, 28:373-381*.

Vygotsky, L. 1978. *Mind in Society*. M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, eds. and trans. Cambridge, MA: Harvard University Press.

Waples, D. and R. W. Tyler. 1931. *What people want to read about*. Chicago: University of Chicago Press.

Weeks, W. B. and A. E. Wallace. 2002. "Readability of British and American medical prose at the start of the 21st century." *British medical journal* 325, no. 7378:1451-1453.

Wegner, M. and D. Girasek. 2003. "How readable are child safety seat installation instructions?" *Pediatrics* 111:588-591.

Weiss, B. D. and C. Coyne. 1997. Communicating with patients who cannot read. *New England journal of medicine* 337, no. 4: 272-274.

Williams, A. R., A. L. Siegel, J. R. Burkett, and S. D. Groff. 1977. *Development and evaluation of an equation for predicting the comprehensibility of textual material.* AFHRL-TB-77-8. Brooks AFB, TX: Air Force Human Resources Laboratory, Air Force Systems Command.

Williams, B. F., J. K. French, and H. D. White. 2003. "Informed consent during the clinical emergency of acute myocardial infarction (HERO-2 consent substudy): a prospective observational study." *Lancet* 361 no. 9361:918-923.

Woern, Y. 1977. "On the relationship between knowledge of the world and comprehension of texts: Assimilation and accommodation effects related to belief structures." *Scandinavian journal of psychology*, 18: 130-139.

Wyler, R. 1989. *The starry sky: An outdoor science book.* Englewood Cliffs, NJ: Julian Messner (Prentice Hall).

Yngve, V. H. 1960. "A model and hypothesis for language structure." *Proceedings of the American Philosophical Association* 404:444-466.

Zeller, D. 1941. *The relative importance of factors of interest in reading materials for junior high school pupils.* New York: Bureau of Publications, Teachers College, Columbia University.

Zeno, S. M., S. H. Ivens, R. T. Millard, and R. Duvvuri. 1995. *The educator's word frequency guide.* Brewster, NY: Touchstone Applied Science Associates.

Zingman, D. E. S. 1977. *Readability and mass political literature: The 1976 presidential election Campaign.* M.Ed. Thesis, Rutgers, The State University of New Jersey.

Zipf, G. K. 1949. *Human behavior and the principle of least effort: An introduction to human ecology.* New York: Addison-Wesley (revised 1965).

Cloze Answers .

The potential for two-way communication is very strong on the Web. As a result, many companies are focused on the Web's marketing potential. From a marketing point of view, today's virtual worlds can attract the curious Web explorers, and interactive database engines can measure and track a visitor's every response.

Biosketch

William H. DuBay has worked as a technical communicator in several southern California firms for 20 years. He is a Senior Member of Orange County chapter and contributes frequently to the *TechniScribe* newsletter and *InterCom.* He serves on the Advisory Board of the Technical and Professional Writing Certificate Program of California State University Long Beach. He also belongs to the Plain Language Association International. Currently, he works as a readability consultant.

The author wishes to thank George Klare for his guidance and help in the development of this article.

Now in Paperback

Smart Language:
Readers, Readability, and the Grading of Text

By William H. DuBay

156 pages, Bibliography, Index
BookSurge ISBN: 1-4196-5439-X

Smart Language is language that fits the reading skill of the reader. This work tells the story of the readability formulas, standard tools used by education, government, and business to match texts with their audience.

Early in the 20th century, education in the U.S. was confronted by a changing school population, especially an increase in the second generation of immigrants in secondary schools. In 1935, surveys revealed that the average readers in the U.S. were adults of limited reading ability.

In response, there was a new reliance on scientific tools for studying and objectively measuring educational problems. This effort engaged the best minds of the age, including Edward Thorndike, William Gray, and Edgar Dale. Their work led to the first readability formulas.

Continuing research demonstrated how better readability increases comprehension, retention, and reading speed and persistence. Publishers of newspapers found that improved readability increased readership up to 65%.

Today, the readability formulas continue to benefit millions of readers throughout the world in many languages.

Writers, teachers, educators, and all those involved in communications will find this work an indispensable guide to the 80 years of research behind the readability formulas. It will give them new confidence in writing for audiences of different reading levels.

Available on Amazon.

Also in Paperback

Unlocking Language:
The Classic Readability Studies

William H. DuBay, Editor
246 pages
BookSurge ISBN: 1-4196-6176-0

Unlocking Language features reprints of 10 landmark studies in readability. They start with Edward L. Thorndike's introduction to his *Teacher's Word Book* in 1921 and end with Rudolf Flesch's introduction to his Reading Ease readability formula in 1949.

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

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

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

That story is briefly covered in a companion book, *Smart Language: Readers, Reading, and the Grading of Text*. The purpose of this book is to bring students of reading into contact with this introductory sample of the original articles, methods, and thinking of these educators.

In all of them, we see the urgency and pragmatism of the times. I hope that reading them in context will highlight their special place in the story of our remarkable language.

Available at Amazon

