

4-1-1987

Readability Formulas: A Necessary Evil?

Kathy Everts Danielson
University of Nebraska

Follow this and additional works at: https://scholarworks.wmich.edu/reading_horizons



Part of the [Education Commons](#)

Recommended Citation

Danielson, K. E. (1987). Readability Formulas: A Necessary Evil?. *Reading Horizons*, 27 (3). Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol27/iss3/4

This Article is brought to you for free and open access by the Special Education and Literacy Studies at ScholarWorks at WMU. It has been accepted for inclusion in Reading Horizons by an authorized editor of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.



READABILITY FORMULAS: A NECESSARY EVIL?

KATHY EVERTS DANIELSON

UNIVERSITY OF NEBRASKA-LINCOLN

Introduction

The word "readable" can be defined in at least three different ways. Klare (1971) suggested that it could mean interest-value, legibility, or comprehensibility. Klare's third meaning, comprehensibility, is the one which readability formulas address.

A readability formula is usually a mathematical equation that strives to relate the comprehension of the reader and the linguistic characteristics of the text. The purpose of this paper is to reflect upon some readability formulas and to propose that they be used only in perspective. Although these formulas are often considered a 'necessary evil,' they need not receive as much attention as they do. Readability formulas certainly should not be thought of as the ultimate instruments in measuring texts for grade levels. Too many factors which readability formulas do not address are more crucial to reading comprehension.

Readability Formulas

Many readability formulas exist, developed from research through two generations. Vogel and Washburne (1928) proposed the first readability formula by estimating the grade placement of children's reading materials. They made classifications of books based upon children's likes and dislikes, and the reading ability of those same children. The grade placements of these books then were representative of the children who read and enjoyed the books. Washburne and Vogel then set out to relate their grade placements to some characteristics of the book. They found that the following four factors were useful indices of grade placements: 1) the number of different words per one thousand words of text; 2) the number of uncommon words; 3) the number of simple sentences in seventy-five successive sentences; and 4) the number of prepositions per one thousand words. Vogel and Washburne did not intend for this formula

to be used for any one book or to appraise overall reading difficulty, although they stated that any book used in the elementary grades could be similarly analyzed "so far as structural difficulty is concerned" (p. 380).

Gray and Leary (1935) suggested a regression formula based upon the number of different hard words, the number of first-, second-, and third-person pronouns, average sentence length, percentage of different words, and the number of prepositional phrases. This formula failed to show differences in readability beyond a certain level of difficulty, however.

Lorge (1944) developed his readability formula using these variables: number of words in the sample, number of sentences in the sample, number of prepositional phrases per sample, and the number of hard words in the sample (using the Dale List of 769 Easy Words).

Flesch (1949) presented his Reading Ease formula which simply required counting the number of words per sentence (in one hundred word sample) and the number of syllables in one hundred words. The formula for this process is : $\text{Reading Ease} = 206.835 - .846 \text{ wl} - 1.015 \text{ sl}$ (where wl = the number of syllables per one hundred words and sl = the average number of words per sentence). To assess interest, Flesch advocated the use of his Reading Interest formula. To find this score, a selection of one hundred words is evaluated in terms of personal words, including all first-, second-, and third-person pronouns, words that have masculine or feminine natural gender, and group words like "people". Also computed is the number of personal sentences. Personal sentences are spoken sentences (quotations), questions or comments addressed directly to the reader, exclamations, and sentence fragments whose meanings must be inferred from the context of the section. The formula for this process is: $\text{Reading Interest} = 3.635 \text{ pw} + .314 \text{ ps}$ (where pw = the number of personal words and ps = the number of personal sentences). No other formulas attempt to measure interest level.

Spache (1953) created a readability formula for primary-grade reading materials, since most of the readability formulas in use were applicable only for reading levels of grade 4 and beyond (Flesch, 1943; Lorge, 1944; Dale-Chall, 1948). Two formulas that had been created for use with

primary-level reading materials (Lewerenz [1935] and Wilkinson [1936] were deemed too lengthy and complicated by Spache. Spache's formula was this: Grade level of textbook = $.141$ average sentence length per one hundred words + $.086$ words outside the Dale Easy Word List of 769 words + $.839$.

Fry's Readability Formula (1968) was designed to "save time." It called for the selection of three one-hundred word passages in the beginning, middle, and end of a book, using no proper nouns. The number of sentences is counted in each passage and averaged among the three selections. Syllables are then counted and again averaged among the three selections. Plotting these average numbers on Fry's graph gives an indication of the reading difficulty level. Fry created his graph grade levels from plotting publishers' graded readers and "smoothing the curve." Maginnis (1969) extended Fry's graph to include primer and preprimer levels, using books of commonly used basal reading series.

McLaughlin (1969) proposed a SMOG formula as another readability formula. (The title SMOG was McLaughlin's tribute to his birthplace, London, where the word 'smog' was coined). The SMOG formula is approached in this way: "1. Count ten consecutive sentences near the beginning of the text to be assessed, ten in the middle, and ten near the end. 2. In the thirty selected sentences count every word of three or more syllables. 3. Estimate the square root of the polysyllabic words counted by taking the square root of the nearest perfect square. 4. Add 3 to the approximate square" (p. 639). The number that is derived is the reading grade level of the book.

Elley (1969) suggested that grade levels for readability could be assessed using a mean noun frequency level. Three passages are selected from a text using at least twenty different nouns. These nouns are then evaluated using the NZCER (New Zealand Council for Educational Research) Spelling List of Levels of Difficulty and the frequency levels of the noun are written down. The mean of these frequency levels is then computed and checked with his readability scale. For example, a mean noun frequency level of 3 - 3.59 corresponds to Elley's suitable age level of seven to eight years.

Alternatives to Formulas Tedious counting is involved

in the formulas described in the previous section. Some alternatives to these formulas per se have been developed within the past thirty years.

Bormuth (1968) proposed the cloze readability procedure as a possible predictor of readability. The steps to this process are: 1) Randomly select six to twelve passages from a book. Each passage should begin at the normal beginning of a paragraph and should consist of at least two hundred and fifty words. Every fifth word should be deleted, allowing for fifty blanks. 2) Tests are then given to the grade level in question (twenty-five students are suggested for reliable results). 3) The mean of each test is determined and then an overall mean is calculated. 4) The test whose mean is closest to the overall mean score is then used for cloze readability. 5) Scores of 44-57% indicate instructional level, while scores above 57% would imply independent level reading materials. Advantages of this method are that the student him/herself manipulates the text and no abstract scale or model is applied.

Bormuth's proposal pointed out that although the cloze test is similar to a conventional test (i.e., fill-in-the-blank), there are three underlying differences. First, in a cloze readability test one word is deleted and in conventional tests whole phrases may be deleted. In addition, a cloze test allows for the deletion of structural words (conjunctions, prepositions, etc.) as well as lexical words (nouns and verbs). Secondly, cloze readability tests are made directly from the text, while conventional tests may use derived interpretations of the text. Thirdly, the cloze test is used before the reader has read the selection and thus a more valid indication of potential readability is attained.

Endicott (1973) proposed a scale of syntactic complexity using a unit of language analysis called the "T-unit" (Watts, 1948; Loban, 1963). His model considered the "extent of a child's syntactic organizational ability" (p. 16) and the complexity within the basic T-unit. The transformation of a child's sentences and embedding clauses allow for a complexity ratio which could be applied to readability formulas. Instead of a formula then, this scale looks at sentence structure.

Singer (1975) proposed simply matching paragraphs of unknown reading difficulty to paragraphs of known reading

difficulty and then attaching that reading difficulty score to the unknown scored paragraph. Since this technique involves visual comparisons of paragraphs it has been called the "Singer Eyeball Estimate of Readability" or SEER. Reliability and validity of this approach were arrived at through the use of thirty-two judges whose readability levels deviated less than an average of 1.0 grade levels when "eyeballing" and matching these paragraphs.

Carver (1975-6) described a similar technique for determining readability called the Rauding Scale. It involves the subjective rating of passages by qualified experts who used a set of established rating passages for references in their measurement of the concept difficulty of a passage. Duffelmeyer (1982) compared this scale and the SEER non-formula readability values to those determined by use of the Spache and Dale-Chall formulas. He found that the Rauding grade levels were closer to the formulated grade levels. Froese (1979), however, found the SEER scores to be more accurate than the Rauding scale.

Irwin and Davis (1980) proposed the use of a readability checklist as another alternative to readability formulas. The basic categories of this checklist are understandability and learnability. Understandability involves consideration of text information and the reader's conceptual and experimental background. Concept development, syntax, clearly stated main ideas, etc., are also factors of understandability. Learnability is based on the organization of the text, reinforcement of the text (including aids and elaboration), and motivation. Teachers can use the checklist to analyze a textbook and then can develop the necessary complementary materials to enhance comprehension.

Wheeler and Sherman (1983) suggested the following alternatives to using readability formulas with nonfiction texts. Considering the organization and content of the book is important, as well as looking for pictures or charts that enhance comprehension. Technical or new vocabulary should be printed in boldface print or in italics. Another critical factor is student reaction to the text. This judgment of students is very often overlooked in the area of readability, yet the students are the intended audience and thus should be a part of the selection process.

Comparison of Formulas Some interesting studies com-

paring readability formulas have emerged. Brown (1965) found that his seventh- and eighth-grade students comprehended a science textbook that the Dale-Chall formula rated at the twelfth-grade level. He then questioned the validity of the Dale list of 3,000 familiar words as the vocabulary load. In addition, Walker (1966), again using the Dale-Chall formula, investigated textbook grade-level placement and found that the publishers' grade levels did not match the Dale-Chall computations. The Dale-Chall formula consistently placed textbooks at higher grade levels than did the actual textbook authors.

Froese (1975) compared the Dale-Chall formula and mean cloze scores among sixth-grade science textbook passages and found that the Dale-Chall formula was not a valid measure of materials when the cloze procedure was used as a criterion. Indeed, the cloze texts proved to be more congruently valid with reading vocabulary and comprehension sources.

Computations of the mean readability of fifty-three Newbery Awards books using the formulas of Dale-Chall, Flesch, Fry, and Lorge were made by Guidry and Knight (1976). They found that the Fry method showed lower graded levels than average and the Dale-Chall predicted higher-than-average grade levels. No two formulas were in agreement about grade levels for every book.

Readable Writing

The authors of readability formulas often offer advice for writing in a more readable manner. This consideration of the audience in the writing process has been addressed differently by various authors.

Flesch (1951) said the writer should focus on the reader and the purpose for writing. Organization, breaking up sentences into shorter sentences, using simpler words, using lots of punctuation, and being brief are all seen as important factors to Flesch.

Klare (1963) suggested that for more readable writing the writer should use "words learned in early life, short words, words of Anglo-Saxon derivation, nontechnical words, words familiar in writing (for instance "pshaw" is used mostly in speech), words used in common meaning, and concrete or definite words, rather than abstract words" (p.19).

In addition, Klare proposed the use of few prepositional phrases and the use of simpler, less complex sentences.

These suggestions for more readable writing often focus on changing the text to make it easier (i.e., shorter words and sentences). This process, however, can make the text more difficult to understand.

Perera (1980) addressed this issue of word and sentence difficulty. She pointed to the fact that there are many difficult one-syllable words (like "adze, "carse", "gneiss") and many children do not know a simple word like "toy" when it is not used as a noun. In the area of sentence difficulty, Perera advised that readability formulas do not take syntax into account. Thus a short sentence with short words would receive a low readability score even if it were completely nonsensical. In addition, some syntactical variations are very difficult to read. Children often comprehend subject-verb-object or subject-verb-adverbial sentences, but adverbial-verb-subject, or object-subject-verb forms, which are much more difficult to read, would receive the same readability score. For instance, "Wagons rumbled down the city streets" is more comprehensible than "Down city streets rumbled the wagons." Both sentences would be scored the same in readability formulas, but the second version is markedly more difficult in terms of comprehension.

Cohesive long sentences may be easier to understand than short choppy ones. For instance, Reid (1972) found that given the following two sentences: "Mary's dress was neither new nor pretty" and "Mary's dress was not new and it was not pretty", students comprehended the second longer sentence much better than the first shorter one. This topic is also relevant to ESL students. Blau (1982) found in her study of ESL students that complex sentences with relational clues (i.e., "therefore") yielded higher comprehension scores than short simple sentences and complex sentences without clues. She suggested that these short simple sentences may actually impede comprehension. "Choppy, unnatural sentences are difficult to read and the relationships and meaning revealed by the formation of complex sentences are apparently lost" (p. 525).

Other Factors in Readability

Readability formulas are not the sole indicators of

readable texts. Klare (1976) reviewed thirty-six experimental studies of the ability of readability variables to predict comprehension and found that nineteen studies were statistically significant, eleven were not statistically significant, and six were mixed (had some significance). These inconclusive reports cannot be used to advocate the use of readability formulas, in fact, these results show that readability formulas cannot be used with a great deal of confidence about their success in predicting reading comprehension.

Many variables other than those mentioned in readability formulas account for whether or not a text is comprehended by a reader. Lorge (1949), himself an advocate of readability formulas, noted two great weaknesses of readability formulas --namely: "they do not directly evaluate conceptual difficulty, nor do they consider the way in which a text is organized" (p. 91). Certainly cohesion of text (Halliday and Hasan, 1976) is another critical factor in comprehension, yet readability formulas do not address it.

Content considerations may also have an effect on readability. New information, interest or value, the nature of the content (be it easy or difficult), and the maturity level of the content as it relates to reader maturity are all factors related to readability (Klare, 1976).

Some reader competence factors such as knowledge of subject matter, reading skill level, and intellectual level are also important. If the reader's knowledge of one subject is high, readability formulas often overestimate the text difficulty. Intellectual factors should be considered above and beyond the formulas as well.

Readability and relevance are two factors to consider together. "The alert teacher may begin an analysis of a work with a quick application of a readability formula, but will realize that the formula falls short when applied to literary prose. Hence, the teacher must consider qualities intrinsic to the work as well as the mental characteristics and interests of students to whom the work will be assigned (Beck, 1984, p. 49).

Readability formulas do not measure such critical factors as motivation, format, illustrations, and adult assistance. Manzo (1970) argued that these formulas are not

"people-oriented" (p. 963) and that there is "no measure of idea load and esthetic differences" (p. 962). Indeed, one scene may appear very different to different readers with diverse experiences and backgrounds, other variables that cannot be calculated with a formula.

Manzo (1970) has listed the following drawbacks to readability formulas: "1) readability formulas have certain inherent problems which make them extremely variable in their predictions; 2) the thrust of present research continues to be one-dimensional, i.e., linguistic; 3) even the best readability formulas are divorced from the influence of reader purpose and experience; 4) there are very few things which even the best readability formula can do which cannot be done almost as well without one; and 5) new insights might occur if researchers could back away long enough to reduce the present level of involvement from causing what appears to be tunnel vision" (p. 964).

Conclusion

The dangers of readability formulas are many. Often the comprehension of materials is based upon the reader him/herself and the quality of the ideas expressed in the text. The interpretation of the text is based upon the motivations, experiences, and interests of the reader. Readability formulas may be useful as references but they should not be used exclusively as the decision-making device guiding classroom uses of reading materials. They may be a necessary evil in some situation such as textbook selection, but they should not be used as the underlying structure of a reading program. Reading is more than decoding easy words and short sentences. Linking the text to the reader's knowledge and experience so that meaning may be derived is the real goal.

REFERENCES

- Beck, P. (1984). Readability and relevance: The quartzite trip. Teaching English in Two Year College, 11, 46-51.
- Blau, E. (1982). The effect of syntax on readability for ESL students in Puerto Rico. TESOL Quarterly, 517-527.
- Brown, W.R. (1965). Science textbooks and the Dale-Chall formula. School Science and Mathematics, 65, 164-167.

- Bormuth, J.R. (1968). Cloze readability procedure. Elementary English, 45, 429-436.
- Carver, R.P. (1975-6). Measuring prose difficulty using the Rauding scale. Reading Research Quarterly, 11, 660-685.
- Dale, E. & Chall, J. (1948). A formula for predicting readability. Educ. Research Bulletin, 27, 22-20, 37-54.
- Duffelmeyer, F. (1982). A comparison of two noncomputation readability techniques. The Reading Teacher, 36, 4-7.
- Elley, W.B. (1969). The assessment of readability by noun frequency counts. Reading Research Quarterly, 4, 411-42.
- Endicott, A. (1973) A proposed scale for syntactic complexity. Research in the Teaching of English, 7, 5-10.
- Flesch, R. (1948). A new readability yardstick. Journal of Applied Psychology, 32, 22-3.
- (1951) How to Test Readability. NY: Harper & Bros.
- (1943). Marks of a Readable Style. NY: Tchrs College.
- (1949) The Art of Readable Writing. NY: Harper
- Froese, V. (1975). Cloze readability versus the Dale-Chall formula. In B. Smith Schulwitz (Ed.), Teachers, Tangibles Techniques: Comprehension of Content in Reading. Newark, DE: International Reading Association.
- (1979). Rauding and SEERing: Judging global readability. Presented--NRC, San Antonio, Nov. '79.
- Fry, E. (1968). A readability formula that saves time. Journal of Reading, 11, 513-516, 575-578.
- Gray, W. & Leary B. (1935) What Makes a Book Readable? Chicago, IL: University of Chicago Press.
- Guidry, L.J. & Knight, D.F. (1976). Comparative readability: Four formulas and Newbery books. Journal of Reading, 19, 552-556.
- Halliday, M.A.K. & Hasan, R. (1976). Cohesion in English. London, England: Longman Group Limited.
- Irwin, J.W & Davis, C.A. (1980). Assessing readability; The checklist approach. Journal of Reading, 24, 124-130.
- Klare, G. (1976). A second look at the validity of readability formulas. Jrnl of Rdg Behavior, 8, 129-152.

- Klare, G. (1974) Assessing readability. Reading Research Quarterly, 10, 62-101.
- (1963). The Measurement of Readability. Ames, IA: Iowa State Press.
- (1968). The role of word frequency in readability. Elementary English, 45, 12-22.
- Lewerenz, A. (1935). A vocabulary grade placement formula. Journal of Experimental Education, 3, 236-240.
- Loban, W.D. (1963). The Language of Elementary School Children. Champaign, IL: Nat. Council of Tchrs of Eng.
- Lorge, I. (1944). Predicting readability. Teachers College Record, 45, 404-419.
- (1949) Readability formulae--an evaluation. Elementary English, 26, 86-95.
- Maginnis, F. (1969). The readability graph and informal reading inventories. The Reading Teacher, 22, 516-518.
- Manzo, A. (1970). Readability: A postscript. Elementary English, 47, 962-965.
- McLaughlin, G. H. (1969). SMOG grading--a new readability formula. Journal of Reading, 12, 639-646.
- Perera, K. (1980). The assessment of linguistic difficulty in reading material. Educational Review, 32, 151-161.
- Reid, J.F. (1972). Children's comprehension of syntactic features found in extension readers. In J.F.(Ed.) Reading Problems and Practices(pp. 394-403) London:Ward Lock.
- Singer, H. (1975). The SEER technique: estimating readability. Jrnl of Rdg Behavior, 7, 255-267.
- Spache The Elementary School Jrnl (1953) 410-413.
- Vogel & Washburne The El. Sch. Jrnl 1928, 373-381.
- Walker, W.L. (1966) Tchrs College Jrnl, 37, 179-181.
- Watts, A.F. (1948) Language and Devel. of Children, Heath.
- Wheeler & Sherman (1983) Science and Children, 20, 38-40.
- Wilkinson, M.S., Weedon, V., & Washburne, C. (1936). The Right Book for the Right Child. New York: John Day