DOCUMENT RESUME

ED 023 565 RE 001 477

By-Fry, Edward B.

The Readability Graph Validated at Primary Levels.

Harvard Univ., Cambridge, Mass. Graduate School of Education.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Report No-TR-1
Grant-OEG-0-8-085762
Note-13p.
EDRS Price MF-$0.25 HC-$0.75
Descriptors-*Cloze Procedure, Informal Reading Inventory, Oral Reading, *Predictive Validity, *Primary Grades, *Readability, *Reading Level, Reading Materials

The validity of Fry's Readability Graph for determining grade level readability scores was compared with the Spache Formula, the cloze technique, and oral reading in the case of seven primary-level books. Descriptions of these four indicated that to determine grade level, Fry's Readability Graph plots the total number of syllables with the total number of sentences for a 100-word passage. The materials used for comparative analysis were selected cloze passages read aloud by 30 primary grade children. Percent of errors was recorded for reading of the words not deleted, and grade level readability scores were computed by the Readability Graph and the Spache Formula. Rank order correlations showed highly consistent correlations for all four methods. The Readability Graph yielded about the same level scores as the Spache Formula. The cloze method was judged to be the most accurate and the most capable of making fine distinctions; however its use is limited because it requires a group of subjects to read the selections for evaluation at a given time. Tables and references are included. This research was funded under Title III of ESEA. (CM)

Technical Report No. 1, U.S. Office of Education Title III Contract No. OGE-0-8-085762 (056) Computer Assisted Instruction Laboratory, Graduate School of Education, Harvard University.

THE READABILITY GRAPH VALIDATED AT PRIMARY LEVELS

by

Edward B. Fry

This paper reports a downward extension of the validity of the Readability Graph that appeared in the Journal of Reading (Fry, 1968). In that article the Readability Graph was compared with readability scores obtained by using the Dale-Chall, Botel, Flesch and SRA formulas and with a set of comprehension scores of 10th graders. The books ranked were at the 5th grade through high school levels of difficulty. This article reports grade level scores, or ranking, for seven books, mainly within the primary levels of difficulty. Rank order correlations are also reported between the Readability Graph and the Spache formula, an oral cloze technique, and errors in an oral reading of the words present.

Background

The rationale for "one more" readability formula is based on simplicity. The author feels that the Readability Graph is both simple and fast. Indeed it is often lack of simplicity that has kept other readability formulas from more widespread use. The Readability Graph seems to work as well at the lower levels as it does at the upper levels, although it still needs further validation by investigators other than the author.

Cloze. For those who are not familiar with the use of the cloze technique as a method for determining readability, a word of explanation is in order. The cloze technique is a procedure in which words are omitted (blanked out). The student is then asked to read the passage, in this case orally, and fill in the blanks with the missing words. One group of students all read the same set of passages and then the passages on which the most errors are made are judged the hardest. Considerable research has shown that this is a good measure of readability. For those readers interested in further information on the cloze technique, articles by Taylor (1953), Rankin (1965), Bormuth (1963), Coleman (1966) and Gallant (1965) are especially recommended.

Spache Formula. The Spache Readability Formula is perhaps the best known primary level readability formula. It has a range from about 1.5 to 4th grade. The first step in this formula is to count off approximately 100 words and then to count the number of sentences in this sample. Next, every word in the sample is checked against the Stone revision of the Dale List of 769 words. Words not on the list are "hard." The percent of hard words is computed, and then the average sentence length is computed. The average sentence length is multiplied by a three-decimal-place constant and the percent of hard words word length is multiplied by a different three-decimal-place constant. These two products are then added to a third constant, producing a grade level estimate. Spache suggests that five to ten samples per book be averaged together. However,

it is not quite that simple. There are ten other rules for what to do in counting verb forms, plurals, possessive endings, adjectival and adverbial endings, hyphenated words, contractions and a few other things.

The computational part of this procedure has been simplified by Daniel Safier, who has developed a table that incorporates the constants and eliminates several steps (Spache, 1966).

Readability Graph. The procedure for using the Readability Graph is somewhat simpler: (1) count the total number of syllables in a 100 word passage; (2) count the total number of sentences in the passage; (3) plot the two scores on the Readability Graph for the grade level of the passage. There is no computation or grammar rules about such things as adverbial endings to worry about. For books and longer articles I recommend an average of three sets of sentence length and syllable counts; in the case of great variability, add a few more samples. The Readability Graph does have one grammatical rule: skip all proper nouns.

Oral Reading. Oral reading has long been used to judge the reading ability of a student. A number of students are asked to read one passage. Those who make the fewest errors are judged to be the best readers. To determine readability, one group of students is asked to read several passages. The passages on which the most errors are made are judged to be the most difficult, or to have the hardest readability.

Previous Correlational Studies. Spache (1966) reports that his formula correlates .86 with the grade levels used by publishers and that Ralph Staiger found a rank order correlation of .70 between the Spache formula

and a scaling of pupil performance based on oral reading errors and comprehension. A Journal of Reading article (Fry, 1968) reported the following rank order correlations for the upper end (grade levels 5-10) of the Readability Graph: Dale-Chall formula, .94; Flesch formula, .96; SRA Reading Ease Calculator, .98; Botel formula, .78; and average score of a group of 10th graders on multiple choice comprehension tests, .93. The Botel Formula does not include sentence length or any measure of grammatical complexity.

A fair amount of research has been done on the validity of using such factors as sentence length and number of syllables or word length to determine readability. Typical of these is the factor analytic study by Stolurow and Newman (1959) of 44 objective elements in writing. They found a high correlation between easy words and monosyllables and reading ease (.90) and, conversely, a low correlation for polysyllables and difficult or unknown words (.91). Average sentence length correlated .66 with difficulty. They concluded that "any yardstick which gave primary weight to the so-called word factor and a lesser but almost equal weight to the sentence factor would account for a good deal of the variance in readability." Brinton and Danielson, after a factor analysis of 20 language elements, came to a similar conclusion that "confirms the importance of word length and sentence length."

Stolurow and Newman suggest that the relative predictive values of these factors change with variations in ability levels of readers, and

Bormuth (1966), in analyzing a number of factors in readability, suggests curvilinearity. It is possible that the curve on the Readability Graph takes this variation into account. The curve is drawn so that in the lower levels sentence length plays a major role in readability while, at the upper levels, word length accounts for most of the variability.

Since readability formulas have importance for those teachers, writers and editors who are preparing material to be read by children, a word of caution is in order. Work from high-frequency word lists and try for a simple style with many, but not all, short sentences. Simply cutting long sentences in half and making word substitutions, like "ox" for "bullock", might not have the desired effect. Use the Readability Graph after the passage has been written.

Method

The purpose of this report is to validate the lower end of the Readability Graph by showing the ranking of passages from primary-level books by scores obtained from the Readability Graph, the Spache Readability Formula, a cloze procedure, and oral reading error scores. Rank order correlations are computed for the four measures.

The work reported on cloze passages is part of a larger project being conducted by Douglas Porter and Helen Popp at the Harvard University Office of Programmed Instruction. Thirty 2nd and 3rd grade pupils were asked to read aloud cloze passages from the seven books, supplying the missing words. The number of errors made on words deleted was recorded for each passage.

Porter and Popp used a cloze procedure which deleted about every fifteenth word in a 150 word passage if it was a noun, verb, adverb or adjective. If not one of those classes of words, the word nearest to the fifteenth word was deleted. After a try-out, a small modification was made: if the word deleted was "impossible" for six adult readers another word was chosen for deletion. The score obtained is the percent of errors made on the passage by 30 pupils.

They also obtained the percent of errors the 30 pupils made in oral reading of the word not deleted.

The same passages used by Porter and Popp were also used to compute grade level readability scores by both the Readability Graph and the Spache Readability Formula. This work was done by two teachers from the Boston Public Schools, Frank Galvin and Alice Healey, who were working with the author on a Title III project, sponsored by the Boston Public Schools, at the Harvard University Computer Aided Instruction Laboratory.

Results and Discussion

All four methods, the Readability Graph, the Spache Readability Formula, the cloze procedure, and oral reading errors ranked the difficulty level of the passages quite well (see Table 1). In my opinion the cloze method was the most accurate and made the finest distinctions. The cloze % of error scores ranged from 12 to 80. Were it not for the enormous amount of time this method takes, cloze procedure would be an excellent way to determine readability. In addition, to the time it takes to make the cloze passages, a number

of different passages must be tested at the same time on the same group of children. You cannot return to the same group of children several months later, for their reading abilities will have changed and the cloze error scores will not be comparable. As a research tool the method is excellent but for practical purposes it is all but impossible to use.

The method using oral reading errors suffers from some of the same limitations. Basically, you can rank a batch of materials at only one time with a group of children. The time needed to construct the cloze passages is saved, and the administration is somewhat faster, but it is still very time-consuming to hear 30 primary grade children read what, for them, are rather long passages. We also feel that the oral reading scores are neither as accurate nor as fine grain as the cloze scores. Table 1 shows that two books, What Spot? and Mississippi Possum, both received the same oral reading score, yet the three other methods all agree that there is a definite difference in readability. Nevertheless, the use of oral reading scores is an interesting method of judging readability, and one not often used. It has the advantage of being an objective, independent, and different validation procedure.

Readability formulas are often validated on such non-objective criteria as subjective judgment or publishers' recommendations. Or they are validated by comparing them with other formulas. These methods are not wrong, but we must continually keep in mind that the real basis for readability is whether a child can read the material. Therefore, validity measures that use children should receive high priority. For this reason

I would like to see oral reading errors used increasingly in research to validate readability formulas, although the time factor limits its use for practical purposes.

The Spache formula ranked the passages fairly well, but it did have some faults. It reported a readability level for Charlie and the Chocolate Factory which was inconsistent with the results of the three other methods. It also reported a readability level of 4.2 for Orlando, The Brave Vulture, a point beyond the formula's range so we had to use another formula (Flesch) to determine grade level.

Hence, one difficulty that we see is that the Spache formula cannot rank books above 4.0 while the Flesch, Dale-Chall and others cannot rank books below 4.0. Those using these formulas in mid-elementary levels must use one formula for the upper level and another for the lower level, sometimes without good articulation.

The Spache formula gives grade level to a tenth of a grade (one decimal place) but this appearance of accuracy could be misleading. Spache reports a probable error of estimate in predicting grade level of a book, using his formula, of 3.3 months. This means that half the time the true score of a book lies within a 6.6 months' band centered around the score obtained by working the formula, and that half the time the real grade level lies outside the 6.6 months' band. Anyone who thinks he is judging the grade level of a passage to within one tenth of a grade level by using the Spache formula is kidding himself.

The Readability Graph ranked the passages quite well and yielded about the same grade level scores as the Spache formula. The Graph did not misrank any books, but it did not make the fine distinctions of the cloze procedure. The rank order correlations in Table 2 show a high degree of consistency between the Graph and the other methods. It is interesting that both the Spache formula and the Readability Graph correlate about as well with the cloze passages. In studying the data I found that the Spache formula gained some consistency by making finer distinctions among the first three books, but lost some by misplacing one book. However, both the Spache formula and the Readability Graph had very high correlations with the cloze method and satisfactory correlations with the oral reading method and each other.

Book	Fry Grade Level	Spache Grade Level	Cloze Errors %	Oral Reading Errors --- %
Come & Have Fun	1	1.4	12	6.4
Green Eggs & Ham	1	1.8	16	5.4
Let's Get Turtles	1	2.1	37	10.3
What Spot?	2	2.7	42	11.7
Charlie & The Chocolate Factory	3	2.5	49	16.7
Mississippi Possum	3	3.6	58	11.7
Orlando, The Brave Vulture	7	6*	80	21.1

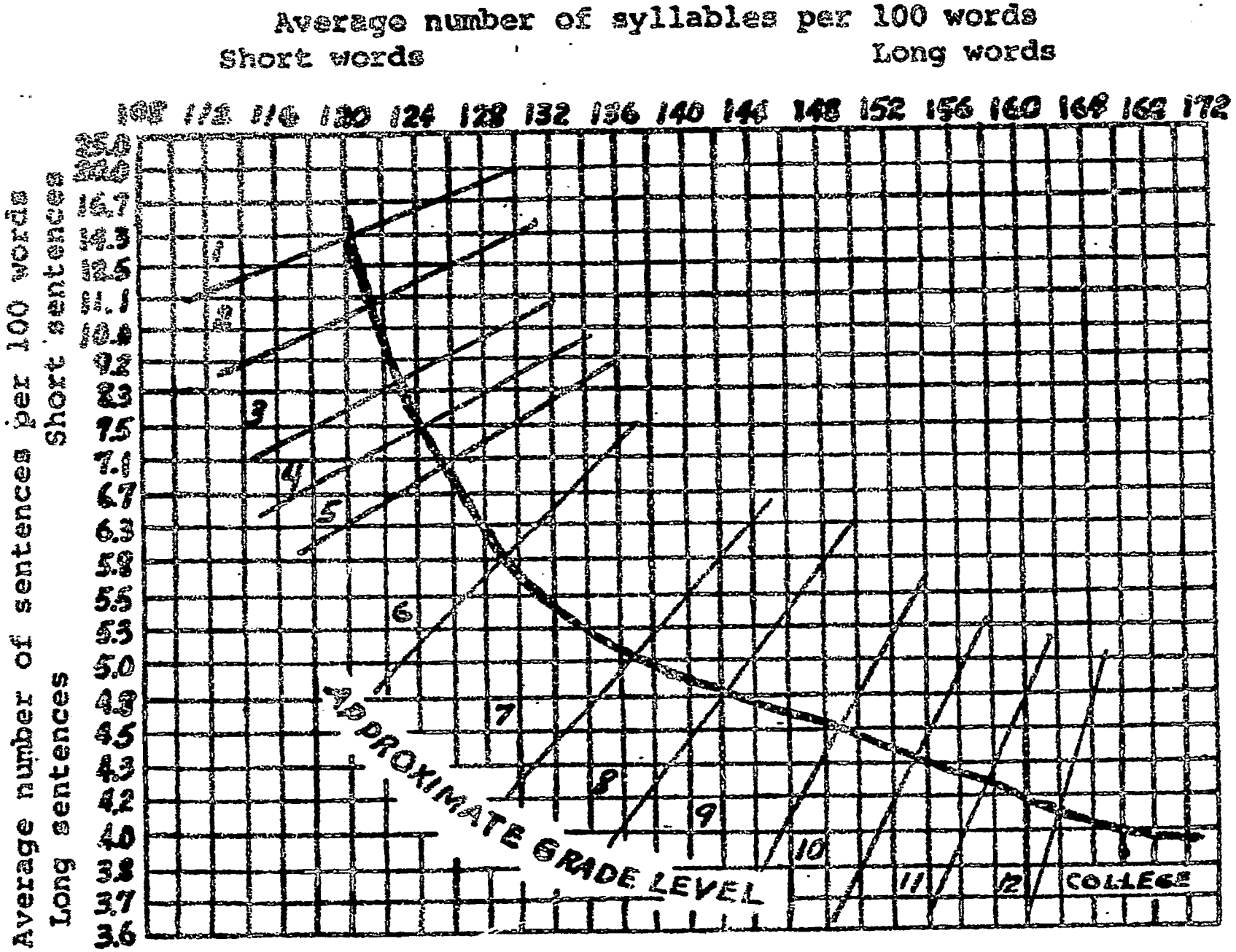
Table 1. A comparison of mean readability grade level scores on the Spache & Fry formulas and cloze error and oral reading error scores by 30 students on 7 books.

	Fry	Spache	Cloze	Oral Reading
Fry	-	.90	.95	.90
Spache	.90	-	.96	.86
Cloze	.95	.96	-	.86
Oral Reading	.90	.86	.86	-

Table 2. Rank order correlations between the Spache Formula, Fry Readability Graph, cloze errors and oral reading errors.

Graph for Estimating Readability

by Edward Fry, Rutgers University Reading Center



DIRECTIONS: Randomly select 3 one hundred word passages from a book or an article. Plot average number of syllables and average number of words per sentence on graph to determine area of readability level. Choose more passages per book if great variability is observed.

REFERENCES

Bormuth, J. R. Cloze as a measure of readability. Proceedings of the 8th Annual Convention, International Reading Association. Newark, Delaware: International Reading Association, 1963. Pp. 131-134.

Bormuth, J. R. Readability: a new approach. Reading Research Quarterly, 1966, 1(3), 79-132.

Brinton, J. E. & Danielson, W. A. A factor analysis of language elements affecting readability. Journalism Quarterly, 1958, 35, 420-426.

Coleman, E. B. Developing a technology of written instruction: some determiners of the complexity of prose. Symposium on Verbal Learning Research and Technology of Written Instruction. Columbia University, 1966. Unpublished manuscript.

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

Fry, E. B. A readability formula that saves time. Journal of Reading, 1968, 11, 513-516.

Gallant, R. Use of cloze tests as a measure of readability in the primary grades. Proceedings of the 10th Annual Convention, International Reading Association. Newark, Delaware: International Reading Association, 1965. Pp. 286-287.

Rankin, E. F. The cloze procedure -- a survey of research. In E. L. Thurston & L. E. Hafner (Eds.), Fourteenth Yearbook of the National Reading Conference, 1965, 14. Pp. 133-150.

Spache, G. M. Good books for poor readers. Champaign, Ill.: Garrard Publishing Co., 1966.

Stolurow, L. M. & Newman, J. R. A factorial analysis of objective features of printed language presumably related to reading difficulty. Journal of Educational Research, 1959, 52, 243-251.

Taylor, W. L. Cloze procedure. A new tool for measuring readability. Journalism Quarterly, 1953, 30, 415-433.