The readability of selected reading texts

Patricia A. Gopperton
California State College
San Bernardino

THE READABILITY OF SELECTED READING TEXTS

A Project Proposal Submitted to
The Faculty of the School of Education
In Partial Fulfillment of the Requirements of the Degree of
Master of Arts
in
Education: School Administration Option

By

Patricia A. Gopperton, M.A.
San Bernardino, California
1978

APPROVED BY:

Advisor

Committee Member
TABLE OF CONTENTS

I. INTRODUCTION ........................................... 1

II. REVIEW OF THE LITERATURE ............................... 4
   What is Readability? ...................................... 4
   Some Readability Characteristics ..................... 5
   Reliability and Validity ................................ 6
   Criticisms of Readability ............................... 9
   The Future of Readability ............................... 11
   Limitations and Implications ........................... 13

III. OBJECTIVE OF REPORT .................................. 18

IV. PROCEDURE .............................................. 19

APPENDIX .................................................. 21

SELECTED BIBLIOGRAPHY ................................. 22
INTRODUCTION

No longer does a teacher's library consist of a horn-book, New England Primer, and Bible. Quite the contrary, a teacher is now faced with a deluge of books of innumerable sizes, types and subjects. And with the contemporary urge to individualize, a teacher's dealings with books is even greater. Which book, which series, which strand and which publisher to select are decisions a teacher must make as she or he begins to motivate and direct the reluctant reader, the non-reader, the so-called average reader, the above grade level reader and all the other kinds of readers that are in between or a part of those classifications.

Can a teacher trust the publisher when he says that book $X$ is a high interest, low level reader? Will the basal reader reach the "average" readers in the classroom? If the books are readable, will the student's interests be met so that he will read and comprehend them? These are some of the questions that plague educators as they look at the multitude of books available and try to decide on a suitable program of reading for the year.

Perhaps one way to alleviate some of the guesswork is to test the readability of the books commonly used in the classroom. Since teachers are becoming more directly
be readable by the children within that grade. In order to
effectively use and teach from any material, it is necessary
that the material be of such a nature that the intended
population be able to read and comprehend it to some useful
degree. Readability formulas can be a part of the solution
to the problem of matching reader with appropriate material.
REVIEW OF THE LITERATURE

What is Readability?

The definitions of readability are almost as copious as the formulas for finding readability. Harris says readability refers to a type of material that is easy to comprehend and at the same time generates interest.\(^1\) Gilliland also thinks readability is related to interest and the fluency with which a book can be read.\(^2\) And Klare states that readability has three characteristics; these are: legibility of handwriting or typography, the ease with which material can be read due to its interest-value, and/or the ability to understand the material written.\(^3\)

In summation, most investigators believe readability is the ability to read and comprehend, at least to some degree, the material read. They feel that interest, style of writing and legibility are factors that can influence whether one piece of material is more readable than another to a particular reader.

\(^1\)Albert J. Harris, Some New Developments in Readability (Vienna, Austria: ERIC Document Reproduction Service, ED 094 344, 1967).


Some Readability Characteristics

The characteristics of readability are many. No one readability formula purports to use all of them and, in fact, some characteristics have not been tested through readability formulas. Although certain factors are thought to be characteristics of readability, the methods for measuring their influence in or on readability have not been refined enough to validate or even wisely guess their significance.

The most widely accepted characteristics are word length (polysyllabic and monosyllabic), vocabulary, sentence length, number of affixes, and number of personal pronouns. Most readability formulas use at least one, if not all, of the previous criteria in their readability formulas.

Other characteristics now being studied are the use of abstract words, idea density, sentence structure and complexity, content, style, reader interest, letter frequency, independent clauses and spelling patterns. The relationship between readability and visibility has been raised. Does type of paper, size and kind of type have anything to do with readability? Is legibility a determining factor in readability? These questions and the relationships of the above listed characteristics still require further study before being validated or correlated in readability formulas.
Reliability and Validity

Reliability studies have been almost nonexistent on most readability formulas. Reliability in readability measurement concerns two areas: analyst reliability and sample reliability.

The Flesch revised formula was tested for analyst-to-analyst reliability as was the Flesch reading-ease formula and the simplified Farr-Jenkins-Paterson formula. All three gave evidence of reliability at .85 and above.\(^1\)

Sample reliability was tested by Dolch when he counted the number of different hard words in each of three books and in three different samples from every tenth page of the books. He concluded that any sampling of vocabulary of a book is an overestimation of its total difficulty.\(^2\)

Bertha V. Leifeste did a study on the representativeness of a sample for the entire book. She found that the longer the sample the closer its agreement with the grade-level of the whole book. Samples from every tenth page were within .5 of a grade-level of the whole book.\(^3\) Bormuth tested the correlations between scores on cloze tests and on tests of different types of comprehension and the

---

\(^1\) Jeanne S. Chall, Readability: An Appraisal of Research and Application (Columbus, Ohio: The Ohio State University, 1958), p. 59

\(^2\) Ibid., p. 62.

\(^3\) Ibid., p. 64.
reliability of each test and found that all were significant at the .01 level.¹

These few studies do not present a strong case for reliability but they do indicate the need for research in reliability of readability formulas.

The validity of readability formulas is concerned with readability as related specifically to word length, syllables, affixes, and sentence length. It is concerned with the magnitude of the correlation found between formula scores or results and the estimates of readability arrived at in some way other than by formula (i.e. teacher judgment).

In a study by Flesch the correlation between word length and affix counts was found to be $r = .87$; correlations between affix counts and abstract words was found to be $r = .78$; and correlations between sentence length and sentence complexity came out to be $.78$ and $.72$.² Bormuth found correlations of $.51$ at the word level of analysis, $.67$ with independent clause level, and $.68$ at the sentence level.³


Harris stated that the per cent of unfamiliar words had the highest correlation with readability (.87) and the per cent of words with more than five letters generated a correlation coefficient of .80 with readability, and the average number of letters developed a correlation coefficient of .74 with readability.¹

There have been many studies investigating concurrent validity by correlating the scores resulting from the application of two or more formulas to the same samples of reading material. One must remember though, that different formulas measure readability by different criteria. Fry found that the concurrent validity between his test and the Dale-Chall equaled .94, and the concurrent validity between Fry and Botel equaled .78. In the same study Fry found correlations ranging from .86 to .96 when the Fry, Spache, cloze formulas and oral reading scores were compared.²

The studies of validity between the readability formulas and the outside criteria of readability have also resulted in significant correlations. For example, Bormuth reports significant correlations between cloze tests and measures of comprehension ability.³ In another investigation

¹Harris, Some New Developments in Readability, p. 3.


he reported a correlation of .83 between cloze readabilities of passages and the combined subjective ratings of the passages by three judges.¹ Dale and Chall reported a correlation of .92 and one of .90 between their formula and the judgments of readability experts and the reading grades of children and adults.²

Criticisms of Readability

With the aforementioned validity and reliability statistics at hand, one may wonder that there are any criticisms about readability and think it the panacea for knowing where to place a student and with what book in order to achieve the most gains. However, the first characteristics accepted and used as criteria for readability, word and sentence length, have been attacked. Critics state that because some words are more familiar, i.e. grandfather, they will be more readable even though longer than a short, unfamiliar word, i.e. sen. They further state that a short sentence of difficult structure could be more difficult to a reader than a longer, uncomplicated sentence. They go on to add that longer sentence structure might be closely related to a reader's memory span rather than to readability.

Although authors of readability formulas often state that their particular readability formula is more apt to be accurate on a stated population such as primary readers, college educated, or other defined groups, readability has been attacked because, as Spache points out, each formula is applicable only to the kind of material the formula was tested upon.¹ Some critics point out that few readability formulas are applicable to primary levels of reading.

Coleman wrote that in investigations involving language characteristics, many investigators made generalizations to entire populations when only a few experimenters, out of thousands of studies, performed correct statistical tests that would allow generalizations to be made on a larger population.²

Other criticisms involve the fact that results from readability formulas were based on different criteria. For example, a formula may require $C_{50}$, which means that only fifty per cent comprehension is required, whereas another formula may require $C_{75}$, which means seventy-five per cent comprehension is necessary to attain a particular readability level.


Critics cite examples of readability formulas being used to adjust the reading difficulty of material. However, authors of the formulas have not encouraged this and also feel that using shortened sentences or making substitutions to easier words is not an acceptable way of lowering a readability reading.

Most critics are quick to point out that readability formulas do not take into consideration the maturity, experience, motivation or interest of the reader. Both Shnayer and Bernstein found that interest was significantly related to comprehension.¹

The Future of Readability

The future of readability is found in many ways to relate to the criticisms of readability. Investigators, experimenters and authors of readability formulas are aware of the criticisms aimed at readability and are seeking more refined methods to overcome these criticisms.

Most research points to a wide field in readability as concerns the linguistic variables, i.e. word depth, letter redundancy, independent clause frequency, letter counts, and parts of speech. Already spelling patterns

have been correlated with readability and efforts have been made to measure linguistic and syntactic difficulty of sentences.

Although Flesch has developed somewhat of a human interest aspect in his readability formula, it is far from the perfection needed to really measure interest as applied to readability. Shnayer states the future of interest in readability when he writes that research might be more profitable in the area of measuring interest rather than readability because of the studies done which show that low interest leads to difficulty in comprehension.\textsuperscript{1}

With so many variables being related to readability, or at least thought to be related to readability, it is no wonder that computers are becoming an integral part of readability research, because they speed up the computations of so many variables.

Until recently there has been a lag in readability research. Klare thinks this lag is due to the lack of effective methods of measurement of readability and the deficiency of an organized body of research and theory from which to draw information.\textsuperscript{2}

As it is, teachers, editors, publishers, speakers and virtually anyone connected with written or oral

\textsuperscript{1}Shnayer, \textit{Some Relationships Between Reading Interest and Reading Comprehension}, p. 3.

\textsuperscript{2}Klare, \textit{Measurement of Readability}, p. 183.
communication sees the necessity for readability formulas. Although readability formulas may not be as refined or perfect as one would wish, they are the best we have to work with at present. At their worst, readability formulas are better than uneducated guesses, a random handing out of books, or comments such as: "Well, try to read it."

Limitations and Implications of Formulas to be Used

Authors of readability formulas would be the first to warn people using the formulas to be aware of the limitations inherent in the formula used. Thus, once a grade level of difficulty has been established it should serve as a mere departure point for further analysis.

In considering a readability formula, one may be faced with contradictory needs. A person may want a formula that has high predictive accuracy and one that can be given with speed. Research has not yet indicated that quick-assessment readability formulas like those of Fry¹ and McLaughlin² are as highly predictive in accuracy as the formulas which are more time-consuming such as the well-known Dale-Chall.³


The amount of time available determines whether the forty-five minutes it takes to arrive at a Dale-Chall score is a limitation or not.

Since the Dale-Chall formula has been used for many years, there is an abundance of research indicating its validity. The recalculated Dale-Chall formula by Powers, Sumner, and Kearl was found to have a validity coefficient of .71. Thus, Klare states that such a study indicates that the original formula had values that were not "flukes" of chance. Klare writes more high intercorrelations have involved the Dale-Chall scores than any other formula which had a similar number of comparisons made. The Dale-Chall formula was found to have the best agreement of six formulas (Flesch, Lewerenz, Lorge, Winnetka, and Yoakam) with grade levels assigned by children's librarians to twelve popular juvenile books. Guckenheimer explains that the Dale-Chall formula and seven judges' (four teachers of social studies and three experts in analysis of reading materials) estimates

---


of grade-level difficulty of a representative sampling of pamphlets on international affairs correlated in the following manner: 0.67, 0.78, 0.84, 0.85, 0.91, and 0.76.¹

Yet the Dale-Chall formula is not without critics. Margaret Kerr criticized the Dale list when she pointed out that the hyphenated words, of which both parts are included in the Dale list, are considered familiar (i.e. right-thinking, loose-leaf) and compound words, of which both parts are included in the list, are not familiar (i.e. cattleman, steamship). She further states that no allowances are made for the definition of words used in context. For example, "mascot" would be considered a "hard" word when used in a sentence like, "Often groups of soldiers like to take to war with them a pet, or mascot." Kerr concludes her criticisms by stating that the appearance of a word on a list does not guarantee its being understood in all its meanings. For example, in the sentences, "The women skinned and cleaned the game. They dressed the skins and tanned them," "game," "dress," and "tan" are all on the Dale list, but each is used in a special connotation.²


Russell and Fea point out that Dale's list does not include any provisions for words used in a slang sense (i.e. You'll wait a "blame" long time). They criticize the lack of provisions for counting partial sentences.¹

The SMOG formula has not been used in enough research studies to establish credible validity and reliability findings. The fact that it is an indiscriminate counting of words with three or more syllables, irrespective of the familiarity of the words, leads one to question its predictive accuracy. However, McLaughlin, in his reply to Pauk's study of sample passages from twenty-different articles, states that the SMOG grade correlated .63 with the Dale-Chall predictions and .62 with the Fry predictions. He goes on to state that his SMOG grade is one at which a book can be completely understood in comparison to the Dale-Chall C₅₀ criteria.²

Neither the Dale-Chall formula nor the SMOG quick-assessment formula takes into consideration the motivation or interest of the reader.

Yet, the author will use the Dale-Chall formula in the proposed study because it is a highly predictive test that has been proven to have validity in a number of different studies. On the other hand, the SMOG formula

¹Russell and Fea, "Validity of Six Readability Formulas as Measures of Juvenile Fiction," p. 142.

is not as well validated, though it is an easier and speedier assessment of readability. Both formulas have been used with reading material which make their use relevant to the study the writer will do.

Yet we must remember that no readability score is the complete answer to the proper evaluation of textbooks, but if conscientiously used they can provide helpful information regarding the difficulty of books being considered.
OBJECTIVE OF PROJECT

The objective of this project is to find the readability levels of three California mandated series in popular use today on the fourth, fifth, and sixth grade levels by using the Dale-Chall readability formula and the SMOG quick-assessment formula. The readability levels thus determined will be compared with the readability levels suggest by the publishers of the books.

The books to be used are the Scott, Foresman and Company's series for disadvantaged, low ability readers:

- Seeking Adventure, grade 4
- Discovering Treasure, grade 5
- Exploring Afar, grade 6

The Harper and Row series for the "average" readers:

- Dreams and Dragons, grade 4
- Moccasins and Marvels, grade 5
- Time and Tigers, grade 6

The Macmillan Company's series for the above average readers:

- The Magic Word, grade 4
- Bold Journeys, grade 5
- Into New Worlds, grade 6
PROCEDURE

In using the Dale-Chall formula, one-hundred words will be taken from every tenth page beginning with the first page numbered ten. The count will begin with the first new paragraph on the page. If there is no new paragraph on an assigned page, the reader will go to the succeeding page to find the beginning of a new paragraph and then begin the count. If there are no words (i.e. a picture) on the assigned page, the reader will go to the succeeding page to begin the count. The rest of the readability procedure will be carried out according to the Dale-Chall formula.¹

In using the SMOG quick-assessment formula, the reader will begin with the page numbered thirty and count ten consecutive sentences beginning with the first new paragraph on that page and if necessary continuing to the next page to complete the count. Another count will be taken thirty pages after the mathematically determined middle point of each book and another count will be taken on the page which is thirty pages previous to the ending of the last story page in the book. If words do not occur on the assigned pages the first succeeding page with words

will be used for the count and always beginning with the first new paragraph on the page. The rest of the procedure will be carried out according to the SMOG formula.¹

The appendix contains an example of the kind of chart to be used to categorize the information obtained through the readability formulas.

Figure 1. A sample graph of the readability levels determined for *Seeking Adventure*, grade 4.

-recommended grade level by publisher
- - - Dale-Chall readability level for each sample
□ □ SMOG quick-assessment formula
- - - average Dale-Chall level for book


