The validity of three readability formulas in measuring juvenile fiction

Madelyn Carol Square
Atlanta University

Follow this and additional works at: http://digitalcommons.auctr.edu/dissertations

Part of the Education Commons

Recommended Citation
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Table of Formula Results for Comparison of Formulas and Selection-Aid Tool Readability Score</td>
<td>24</td>
</tr>
<tr>
<td>2.</td>
<td>Quick Reference Chart</td>
<td>32</td>
</tr>
<tr>
<td>3.</td>
<td>Dale-Chall Correction Table</td>
<td>34</td>
</tr>
<tr>
<td>4.</td>
<td>Dale-Chall Wordcheck List</td>
<td>37</td>
</tr>
<tr>
<td>5.</td>
<td>Lorge Wordcheck List</td>
<td>38</td>
</tr>
<tr>
<td>6.</td>
<td>Summary Table of Formula Elements</td>
<td>42</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>iii</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>I. <strong>INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>Rationale</td>
<td>1</td>
</tr>
<tr>
<td>Evolution of the Problem</td>
<td>5</td>
</tr>
<tr>
<td>Contribution to Educational Theory and Practice</td>
<td>7</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>7</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>7</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>8</td>
</tr>
<tr>
<td>Research Procedures</td>
<td>8</td>
</tr>
<tr>
<td>Materials and Instruments</td>
<td>9</td>
</tr>
<tr>
<td>Survey of Related Literature</td>
<td>13</td>
</tr>
<tr>
<td>II. <strong>COMPARISON OF READABILITY LEVELS, ANALYSIS, AND EVALUATION OF LORGE, FLESCH, AND DALE-CHALL FORMULAS</strong></td>
<td>20</td>
</tr>
<tr>
<td>Comparison of Readability Levels</td>
<td>22</td>
</tr>
<tr>
<td>Element Analysis of the Lorge Formula</td>
<td>26</td>
</tr>
<tr>
<td>Element Analysis of the Flesch Formula</td>
<td>29</td>
</tr>
<tr>
<td>Element Analysis of the Dale-Chall Formula</td>
<td>31</td>
</tr>
<tr>
<td>Comparison of Formula Elements in the Lorge and Dale-Chall Formula</td>
<td>33</td>
</tr>
<tr>
<td>Comparison of Formula Elements and Findings of the Flesch Formula</td>
<td>39</td>
</tr>
<tr>
<td>Evaluation of the Dale-Chall, Flesch and Lorge Readability Formulas</td>
<td>41</td>
</tr>
<tr>
<td>III. <strong>SUMMARY AND RECAPITULATION</strong></td>
<td>44</td>
</tr>
<tr>
<td>Summary of Related Literature</td>
<td>46</td>
</tr>
<tr>
<td>Findings of the Study</td>
<td>48</td>
</tr>
<tr>
<td>Conclusions</td>
<td>50</td>
</tr>
<tr>
<td>Implications</td>
<td>51</td>
</tr>
<tr>
<td>Recommendations</td>
<td>52</td>
</tr>
<tr>
<td><strong>BIBLIOGRAPHY</strong></td>
<td>54</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Rationale

The teacher of reading has a vital interest in materials needed in the teaching of reading, but his interest in this is not any greater than is his interest in the readability of the materials which are used. If the major aim of reading is the comprehension of meaning, the teacher must be interested in the measurement of comprehensibility of materials. He wants some means of quantifying his statements about the difficulty of material. It is not enough to say that reading is difficult or easy. He must have reference points or scales by which to judge printed materials.

Readability is not an easily defined concept. It involves an interaction between reader and book. Because communication between writer and reader seldom is perfect, readability rarely can be absolute. Usually it is a matter of degree. The teacher thus must make some practical decisions as to the degree of understanding that is necessary before a book may be considered readable by a child at a certain grade level. In short, he must determine how much the reader must get from the printed material before it becomes readable for him.

The concept of readability generally refers to the success that the average individual has with a book. Dale and Chall point out that read-
ability refers to those elements within printed material which affect the success that a group of readers have with it.

Unfortunately, the teacher cannot be completely satisfied with this concept. A book that is readable for one child may not be readable for another child even if he has the same general level of reading ability. The teacher cannot make prescriptions for the average individual. He must make decisions for a specific child.

One approach to the readability of books is to ask teachers or librarians to rate books according to difficulty. This has been tried and has been found to be inadequate. Teachers and librarians are not proficient in designating the grade-level for which a given book is written. It is doubtful that they exhibit a greater degree of accuracy in selecting the right book for the individual youngster.

Another approach was made by Vogel and Washburne. They surveyed 37,000 children and developed the Winnetka Book List. This list contained 700 books that children from the second to the seventh grades claimed to have read and enjoyed. The grade placement of each book was found by determining the average reading ability of the children who read and enjoyed the book.

Teachers frequently use an informal technique in helping children choose books. Books are selected somewhat below the pupil's estimated level of reading ability. Rough estimates of the appropriateness of a

---


book are obtained by having the child read orally. Two or three selections of 100 words each are chosen from the book for the pupil to read. Errors are noted as the child reads orally. If he misses more than five words out of a hundred and/or achieves less than 85 to 95 per cent comprehension, the child is directed to another book which has been assigned a lower grade level of difficulty.

Finally, it is possible to develop and apply a readability formula to the book. Since 1923, more than thirty such formulas have been developed. Generally, the authors of these have attempted to identify the factors most agreed upon are vocabulary, sentence structure, and the number of prepositional phrases.¹

The problem of the readability of books is not new. It has been the subject of discussion and experimentation for years. Interest in problems regarding readability has been heightened for a number of reasons. First, the trend to emphasize reading as the chief aid to learning appears to be on the increase. This trend continues in spite of significant production of professional literature on the use of other types of learning aids. The tempo of the reading approach to education continues unabated and the readability of instructional materials has become a first-order problem.²

Readability has become a problem of prime importance for a second reason which stems from a better professional understanding of the relationship between the readability of instructional materials and frustra-

¹Dechant, op. cit., p. 465.
tions in reading abilities at any grade level. Studies of differences as well as likenesses in the learning problems of children have directed attention to the wide range of reading abilities at any one age or grade level. The fact that at any one reading age and/or grade level there have been reports of discrepancies between grade scores achieved on standardized tests and the ability to read instructional materials designated for that grade level comprises a third reason for the increased focus upon readability problems.

Discrepancies between the readability of books with the same grade level designations is a fourth reason for current interest in a more scientific approach to readability.

Other reasons for research in readability have been: (1) the trend to reduce the vocabulary load of basal textbooks, especially in grades one and two, (2) the slow extension of practices in the direction of the experience approach to reading, and (3) the commercial values of both tradebooks and texts.¹

The idea behind readability measurement is the matching of reader and printed material. It assumes that readers differ in their ability to read and that the printed material in turn varies in the amount and kind of ability needed to read and understand it.²

In devising formulas for readability investigators have started with books or passages representing a range of reading difficulty. The books


²Jean Chall, "Measurement of Readability," Education Digest, XXLI (November, 1955), 44.
were ranked or graded by some method without regard for varying components or elements of the books.

Then the books were analyzed to account for this independently determined order of difficulty.

Factors which could be measured reliably and were significantly related to the difficulty of the passages were combined in a formula. The resulting formula could then be used to estimate the grade level of material similar to the kind on which the formula was standardized, without resorting to testing the prospective readers.¹

The readability formulas are valuable as yardsticks. They are the best means yet devised for measuring certain aspects of reading difficulty and for comparing various pieces of writing.

No one maintains that any readability formula is perfect. Research is still being conducted and doubtless will continue for years. In the meantime, these formulas can be extremely helpful in measuring reading difficulty on an objective basis. With the measure given by a formula, we have the means of comparing the difficulty of one book with another. With a book for children, this is particularly important if we are to develop and refine their interests in reading.

**Evolution of the Problem**

In the writer's short experience of working with children, she has seen how children leave one book on the shelf unused and wear out another with constant borrowing and reading. In many cases, the writer noticed that both books were attractively illustrated and well designed. When

¹Ibid.
both books were read aloud to a group of children, both held them spell-bound. Clearly both were good stories to listen to but one was not read by the children themselves. There seemed to be something about the way the words were chosen and combined into sentences that made one book more difficulty for a child "to read to himself."

On furthering her studies at Atlanta University in the area of reading, the writer came across what seemed to have been the answer to the query about children's books. The answer was in the form and knowledge of readability.

The writer found that there were in use formulas which propound to measure the difficulty level of printed materials. Three formulas in particular were thought to be the most outstanding in judging juvenile reading material.

The writer undertook this study, determined to find some valid and reliable means whereby "the right book for the right child may be selected."

Though the writer felt that these formulas are a step in the right direction of matching books and reader, she realized that the use of readability formulas as evaluators of reading levels of printed material is only partially solved today. After close examination of these formulas the writer began questioning their validity because the formula did not take into account the reader's reading ability, maturity, experiences, motivation, or interest and/or lack of it in the material itself.

Other methods of validation, such as the use of external criteria are needed in the effort to arrive at some measure of the validity of the different formulas. The use of library selection aids as described in
the proposed study is one of a number of methods employing external criteria.

**Contribution to Educational Theory and Practice**

A study of the science of measuring readability can give us many clues to the success and failure of certain books for children. The efforts of the classroom teacher and librarian could also be facilitated by the products of these studies in the form of readability indexes of books.

Some agency, perhaps a National Bureau of Readability Standards, could make it possible to provide comparable ratings on all instructional materials sold for use in schools. This type of service would do two things: (1) stimulate more research on this problem, and (2) give teachers dependable information on the relative difficulty of books and instructional materials.

**Statement of the Problem**

The problem of this study was to determine if the Lorge, Flesch, and Dale-Chall Readability Formulas were equally valid measures of the readability levels of award-winning and classical juvenile fiction.

**Purpose of the Study**

All of the methods for determining readability which have been devised up to this time have shown limitations. Yet in spite of these limitations such formulas have been used to assess the difficulty of books and other materials. The question arose: How valid, actually, are these formulas for measuring specific categories of fiction?

Specifically, these purposes were:
1. To determine which formula of the ones used in this study is more consistent with the readability levels assigned by specific library selection aids as a measure of juvenile fiction.

2. To determine which formula(s) can be used in measuring readability with the greatest efficiency.

3. To analyze elements which are counted in the computation of each of the formulas to be used:
   a. vocabulary load
   b. sentence length
   c. affixes
   d. prepositional phrases
   e. personal references

4. To compare formula results and elements based on an analysis of the factors listed above.

**Definitions of Terms**

For the purpose of this study terms which might need clarification are indicated below:

1. Award winning juvenile fiction—books of fiction which are considered as outstanding contributions to children's literature and, because of literary merit, have won acclaim.

2. Classical books—books which after years of reading still appeal to the young because of their literary quality.


4. Readability—the quality of a piece of reading matter that makes it interesting and understandable to those for whom it is written. In many respects, it is a parallel term to the more familiar "reading ability." The former refers to a characteristic of the book, the latter to a characteristic of the reader.

5. Readability Formulas—formulas devised to measure the reading levels of printed material.

**Research Procedures**

The investigation into the problem proceeded in the following
manner:

1. The writer first compiled a list of thirty books, randomly selected, from various book award lists.

2. Grade placements were then assigned to each book by library selection aid tools. Books in which library selection aid tools were in agreement on grade placement were selected for study. The number of books selected for the study was twelve.

3. Library selection aids were subjected to a thorough search for references to juvenile fiction selected for the study. Where such references were made, information given was carefully recorded and sequentially classified in terms of: (a) reading grade levels, (b) ages of children for which the books was suitable, and/or (c) additional indications of difficulty of the material contained in the book.

4. Each book selected for study was subjected to three measures of readability, the Dale-Chall, Flesch, and Lorge readability formulas.

5. Reading grade levels suggested by library aids for each book were compared with each of three computed readability measures and charted for purposes of comparison and evaluation.

6. A systematic and thorough study was made of each formula for the purposes of analyzing, comparing and evaluating the three readability formulas.

7. After collection and organization of research data a summary of findings, conclusions, recommendations and implications were presented.

Materials and Instruments

The book selection.--From the original books, the writer made a random selection of books to use in her study. Books which were suggested were selected from the John Newberry Medal Award List. This award is presented each year to the author of the most distinguished contribution to American Literature for children. Other books were selected from The Spring Book Festival Award List; Boy's Club of America Junior Book Award (this award is based upon children's choices of books); list of Best Books of the year selected by the editors of Junior Books Appraised, and from
the Children's Classics list published by The Horn Book Magazine. This magazine is the only magazine in the United States concerned exclusively with literature for children and young people. Selections were also taken from the Children's Catalog which gives a story summary and also grade level of book listings.

Books selected were:


Library selection aids.—The selected books were given grade placements according to four library selection aid tools. The selection aids
used were:

Bowker, 1966.

Arbuthnot, Mary, et al. *Children's Books Too Good to Miss.*
Cleveland, Ohio: The Press of Western Reserve University,
1966.

Eskin, Mary (ed.). *Good Books for Children.* Chicago: The
University of Chicago, 1966.

Wilson, 1951-66.

The twelve books on which the selection aid tools agreed most
closely on grade placement were then selected for analysis.

The readability formulas.--The readability formulas used in this
study were Rudolf Flesch's Yardstick for Measuring Readability, The Lorge
Formula for Estimating Difficulty of Reading Materials, Dale-Chall's
Formula for Predicting Readability. Procedures for using the formulas
are as prescribed by the authors:

In his formula, Dr. Flesch used three factors: average sentence
length, relative number of affixed morphemes (prefixes, suffixes, in-
flectional endings), and relative number of personal references. Flesch's
Formula appears below.

1. Systematically select samples of 100 words throughout the
material to be rated;

2. Compute average sentence length in words \(X_s\);

3. Count the number of affixes \(X_m\);

4. Count the number of personal references \(X_n\);

5. Average the results and insert in the formula:

\[
0.1338X_s + 0.0645X_m - 0.0659X_n - 0.7502
\]

\[1\]

Rudolf Flesch, *How to Test Readability* (New York: Harper and
Brothers, 1951).
The Lorge Formula was one of the first easy-to-apply readability formulas and is still among those widely used. By counting the relative number of prepositional phrases an index of readability, in terms of grade scores, can be obtained.

1. Compute average sentence length in words ($X_2$);
2. Compute number of prepositional phrases per 100 words ($X_3$);
3. Count number of different hard words per 100 words not on the Dale 796 word list ($X_4$);
4. Substitute in formula: 
   $$X_1 \text{ (grade placement)} = .07X_2 + .1301X_3 + .073X_4 + 1.6126$$

The Dale-Chall Formula is thought of as the second most frequently used formula. This formula was based on the assumption that: (1) a larger word list would predict better than the Dale 796 word list used by Lorge, particularly at the upper levels of difficulty, (2) the count of personal references as used by Flesch was unnecessary, and (3) a shorter, more efficient formula could be developed using only a word factor and a sentence factor.

1. Select 100-word samples throughout the material to be rated;
2. Compute the average sentence length in words ($X_2$);
3. Compute the percentage of words outside the Dale list of 3000 ($X_1$, or Dale score)
4. Apply in the formula: 
   $$X_{c50} = .1579X_1 + .0496X_2 + 3.6365$$

Survey of Related Literature

The history and development of readability formulae.—Reading materials vary in the ease or difficulty with which they can be used and enjoyed by readers in various stages of maturity. Every teacher who has observed children and helped them choose books has discovered that materials dealing with the same topic may vary widely in difficulty.

The idea of gradation in a series of texts for reading instruction was first introduced in the McGuffey readers. The selections were assigned to lower or higher levels largely on the basis of author's judgment. A check against pupils and books proves that grade for grade the McGuffey books are relatively difficult.

Formerly, teachers depended upon designation of texts for indications of relative difficulty. For trade books the author's and publisher's statements as to the audience for which the book was intended, supplemented by the expert opinions of librarians and other judges of children's books, have served as guides. Teachers have also relied on the judgments of the children themselves, and on observations of the children's behavior in using a particular book, but these ratings and observations are highly subjective.¹

In 1923, Lively and Pressey conducted one of the first experiments to find the elements that made for difficulty in reading materials. They had heard from many junior high school teachers that the textbooks in science contained so many technical words that more time was spent on the

study of scientific vocabulary than on scientific facts.

Lively and Pressey measured the vocabulary difficulty of books by assigning the Thorndike frequency number to each different word and taking the average of these numbers. Lively and Pressey consulted Thorndike's *The Teacher's Word Book* (a collection of words dealing with the frequency of their use) to determine the frequency of word usage. Thorndike's word book was of special importance because it became one of the bases not only for the teaching of vocabulary in the schools but also for the first measure of readability that can be considered a formula. Books that had a lower number were considered more difficult than books with higher numbers.¹ This measurement did not directly relate the books to reading comprehension. Using only one kind of measurement, it merely arranged books in relation to one another in terms of vocabulary difficulty of the material. While measurements found second-grade books to be simplest on the basis of vocabulary difficulty, a conceptual level was not determined, therefore there was no determination as to whether these books could be understood by children in the second grade.

It was not until Vogel and Washburne produced their Winnetka Formula in 1928 that a formula was created which related elements of difficulty in written materials to specific reading levels.

Dr. Carleton Washburne and teachers of the Winnetka, Illinois, schools, confronted with the problem of finding the right book for each child from the growing supply of attractive trade books in the school

library collections, undertook to grade an array of children's books according to levels of reading ease. They sought to determine a grade index for each book that corresponded with the reading achievement levels of average children, grades three to eight. In 1924 Dr. Washburne and his associates requested 800 teachers to find out what books were read and liked by elementary school children. The children were asked to tell how well they like the books and why. These responses indicated the books with the greatest appeal for each reading level.

Responses were received from 36,750 children. The ages and reading levels as measured by the Stanford Achievement Test were also obtained. This information was used to determine the median reading achievement grade of the children liking a particular book. This study eventuated in the Winnetka Graded Book List contained a list of 700 books on which 25 or more judgments were received. This publication gave median chronological age at which a particular book was liked by boys and girls reporting.

This experiment in rating the level of children's literature was followed a few years later by pioneering research by Washburne and his associates in determining a formula for rating the level of reading ease.1 The term "readability" was coined some years later.

Representative passages of known reading level in 152 books chosen from the Winnetka Graded Book List were analyzed for factors contributing to reading ease. The elements included in the formula were:

1. The number of different words occurring in a 1000-word passage.
2. The number of prepositions (including duplicates) occurring

---

1 Vogel and Washburne, op. cit., pp. 373-81.
in a 1000-word sample.

(3) The number of words in a 1000-word sample not found in the Thorndike Teacher's Word Book, first edition.

(4) The number of simple sentences in 75 sample sentences.

Scores by this formula were correlated with criterion scores. These were the average paragraph meaning scores on the Stanford Achievement Test for children who had liked a particular book. The Washburne-Vogel formula assessed the average grade level of each book (for example, grade 6.5 or 4.9, as the case might be). Each book listed received a grade index indicating the difficulty level. The Washburne-Vogel formula was used for furnishing the ratings given in a subsequent publication, The Right Book for the Right Child. In the introduction to this book is a description of how the formula was derived. Later the formula was revised to include just three elements: (1) number of different words, (2) number of different uncommon words, and (3) number of simple sentences. The authors also published a chart to facilitate the use of the formula.

Since this pioneering effort, many other workers have devised formulas based on the statistics of regression equation which are variations of the Washburne-Vogel technique.

A summary was made by Betts in 1949 which listed 142 references on


3Hildreth, op. cit., p. 373.
the subject of readability.\(^1\) Another summary by Cole\(^2\) the series of
articles edited by Dale\(^3\) and an earlier summary by Gray and Leary combine
with Betts bibliography to give a fairly complete picture of objective
attempts to determine the difficulty of printed materials. All references
mentioned here and above are concerned with testing of hypothesis rather
than with critical evaluation. They reflect an interest in exploration
rather than validation of readability measures.

The application of readability formulae to specific types of read-
ing material.—Kiare in 1950 estimated that 34 formulas or methods for
determining the reading difficulty of printed material had been devised.\(^4\)

The evaluative studies that have been done in readability have
proven to be contradictory to each other in certain instances.

Margret Kerr writes, "... none of the formulas (Flesch, Lorge,
Dall-Chall) is effective at the primary level. They were originally
designed for use with more difficult material." She goes on to point out,
"For general use evaluating books for use in grades IV-VIII the Dale and
Lorge formulas are easier for the average person to use.\(^5\)

\(^1\) Betts, op. cit.

\(^2\) Luella Cole, The Elementary School Subjects (New York: Holt,

\(^3\) Edgar Dale (ed.), Readability, A Publication of the Conference on
Research in English (Chicago: National Council of Teachers of English,
no date).

\(^4\) George Klare, "Evaluation of Quantitative Indices of Comprehen-
of Minnesota, 1950).

\(^5\) David Russell and H. Foa, "Validity of Six Readability Formulas as
Measures of Juvenile Fiction," Elementary School Journal, LXII (November,
1951), 136.
Russell's and Fea's research turned up a somewhat different finding. Russell did research on the validity of six readability formulas using them as measures of difficulty of juvenile fiction (Flesch, Lorge, Winnetka, Yoakam, Dale-Chall and Lewerenze). Sixty-three Children's Librarians were asked to rate certain books according to grade level of difficulty. The findings show that the Dale-Chall, Flesch, and Lorge formulas most closely approximate the ratings of the children's librarians while the Lewerenze, Winnetka, and Yoakam formulas average over a grade and one-half deviation from the librarian's mean rating of difficulty.¹

Spache in 1953 pointed out that an examination of certain formulas and studies revealed that interest had been focused largely on the evaluation of materials written for adults, therefore none of these formulas, the Flesch, Lorge or Dale-Chall is applicable to materials on the primary level.²

Other studies in the area of readability formulas have found the same results as listed:

1. Readability formulas are not universally applicable. They are especially ill adapted to poetry and plays.

2. Flesch, Russell and Fea, and Chall agree that readability formulas currently available cannot measure readability with a high degree of accuracy or effectiveness.

3. Existing readability ratings produce approximate assessments of readership levels on factual material. However, on creative material, they are relatively ineffective.

¹Ibid.

²George D. Spache, "New Readability Formula for Primary-Grade Reading Materials," Elementary School Journal, LXIII (March, 1953), 412.
4. No readability formula is effective at all levels of difficulty. In addition, each formula is most reliable when used with materials of the same level as that on which it was standardized. This further limits the use of each formula.

5. Analyzing difficulty levels in literature is an aid to teachers, publishers, librarians and parents. It enables them to match reading material to individual level of students.¹

Readability formulas are not a panacea for meeting comprehension problems. They frequently yield different results and fail to measure many elements considered important for readability. However, they are useful in that they give the user some idea of the difficulty of books, in determining the sequency to be followed in recommending books to a child, and in detecting the difficult words and sentences in the book.

CHAPTER II

COMPARISON OF READABILITY LEVELS, ANALYSIS, AND EVALUATION OF LORGE, FLESCH, AND DALE-CHALL FORMULAS

The problem of the study was to determine if the Lorge, Flesch, and Dale-Chall Readability Formulas are equally valid measures of the readability levels of award-winning and classical juvenile fiction.

Purposes for the study were to determine which formula could be used in measuring readability with the greatest efficiency and was consistent with readability levels assigned by specific library selection aids. The writer's last two purposes for the study were derived from the question, "What makes these formulas and the results obtained from these formulas differ?" Therefore, the writer undertook the last two purposes which were analyzing formula elements to determine what influence these factor counts (vocabulary load, sentence length, affixes, prepositional phrases and personal references) exerted on the formulas and comparing formula results and elements based on the analysis of formula elements.

Application of formulas had to be administered in order for purposes one and two to be accomplished. Formula application adhered to guidelines set forth by their authors. In order to determine, as stated in purpose two, which formula was consistent with readability levels assigned by specific library selection aids the writer consulted the following publications: *Best Books for Children*, *Good Books Too Good to Miss*, *Good Books for Children*, and *Children's Catalog*. These library selection aids
also acted as external criteria from which the validity of the different formulas was arrived. The usability of the formula itself determined the formula's efficiency and was the basis from which its comparison and analysis were drawn.

In order for the study to evolve, a list of thirty books randomly selected from various book award lists (John Newberry, Spring Book Festival, Boy's Club of America, Best Books of the Year, Horn Book Magazine and Children's Catalog), had to be compiled. Grade placements were assigned to each book by library selection aid tools. Books in which library selection aid tools were in agreement on grade placement were selected for the study. The number of books selected for the study was twelve.

Each book selected for the study was subjected to three measures of readability: The Dale-Chall, Flesch, and Lorge Readability Formulas. Reading grade levels suggested by library aids for each book were compared with each of three computed readability measures charted for the purposes of comparison and evaluation.

A systematic and thorough study was made of each formula for the purposes of analyzing, comparing and evaluating the three readability formulas. After collection and organization of research data a summary of findings, conclusions, recommendations and implications were presented.

Formula application was the major criterion against which each formula was compared and efficiency determined. The following guidelines were devised by the writer as a second basis on which to judge formula efficiency: (1) the length and number of samples one must select to apply formula, (2) the number of formula elements, and (3) the complexity of factor counts.
The results of each readability score was compared with grade placements received from the library selection aid tools. The comparison of readability scores with those taken from the selection aids leads the way for analysis of formula elements.

Chapter II will present the data found under the following major headings:

1. Comparison of Readability Levels of Juvenile Fiction as Computed by Formulas and Determined by Library Selection Aid Tools.

2. Element Analysis of the Flesch Formula.


5. Comparison of Formula Elements and Findings of the Flesch Formula with those of the Lorge and Dale-Chall Formula.


Comparison of Readability Levels of Juvenile Fiction as Computed by Formulas and Determined by Library Selection Aid Tools

Grade placement scores of selection aid tools and formula results were charted for the purpose of analysis.

Formula results showed a close approximation with library selection aid tools grade placement. However, there was a definite trend for grade placement scores of formulas to be a half grade to two or three grades higher than the lowest approximate grade assigned by selection aid tools. The writer refers the reader to Table 1 and specifically points out the results given for the following books: Alice in Wonderland, The Saturdays, Mary Poppins, Henner's Lydia, and Ringtail.
The Lorge Formula results produced slightly lower scores than the Dale-Chall and Flesch Formulas, although these scores were not lower than the selection aid tools scores. However, the Lorge formula did produce one grade placement lower than the selection aid scores; this score was computed for *Huckleberry Finn*.

In referring to Table 1 the reader will also note, after checking formula results for the Dale-Chall Formula, that three books are rated lower than the Lorge Formula: *Adventures of Pinocchio*, *Winnie-the-Pooh* and *Tom Sawyer*. These lower scores are contributed to Dale-Chall's Correction Table in which formula raw scores must be converted. Dale-Chall's Correction Table as related to the formula's raw scores and grade placements will be explained and discussed in depth further in the paper.

Flesch's formula results were also higher than Lorge formula results. A table for the purpose of converting raw scores in Flesch formula had to be applied. This table will be discussed in depth.

During the course of this paper it should be remembered that library selection aid tools were used as external criterion and that studies similar to this one might produce contradictory results if different external criterion is used.

Library selection aid tools attempt to classify children's books by various age or grade levels. It should be recognized, however, that this system and similar ones are only guides. In other words, when books of a particular grade and/or age level are assigned to children performing on that level, one must remember that different children of the same age often read at different maturity levels.

Discrepancies between selection tools and formulas were cited earlier. Discrepancies between formulas themselves were also stated under
### TABLE 1

**TABLE OF FORMULA RESULTS FOR COMPARISON OF FORMULAS AND SELECTION-AID TOOL READABILITY SCORE**

<table>
<thead>
<tr>
<th>Title of Book</th>
<th>Selection Tools</th>
<th>Lorge Formula</th>
<th>Dale-Chall Formula</th>
<th>Flesch Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peter Pan</td>
<td>4 - 7</td>
<td>4.85</td>
<td>5 - 6</td>
<td>5</td>
</tr>
<tr>
<td>Alice in Wonderland</td>
<td>4 - 7</td>
<td>5.4</td>
<td>5 - 6</td>
<td>6</td>
</tr>
<tr>
<td>Adventures of Pinocchio</td>
<td>3 - 6</td>
<td>4.68</td>
<td>1 - 4</td>
<td>5</td>
</tr>
<tr>
<td>The Saturdays</td>
<td>4 - 7</td>
<td>5.88</td>
<td>5 - 6</td>
<td>6</td>
</tr>
<tr>
<td>Winnie-the-Pooh</td>
<td>1 - 4</td>
<td>4.78</td>
<td>1 - 4</td>
<td>5</td>
</tr>
<tr>
<td>Cinderella</td>
<td>1 - 3</td>
<td>4.92</td>
<td>5 - 6</td>
<td>5</td>
</tr>
<tr>
<td>Heidi</td>
<td>4 - 6</td>
<td>6.21</td>
<td>5 - 6</td>
<td>6</td>
</tr>
<tr>
<td>Mary Poppins</td>
<td>4 - 7</td>
<td>5.30</td>
<td>5 - 6</td>
<td>5</td>
</tr>
<tr>
<td>Huckleberry Finn</td>
<td>6 - 9</td>
<td>5.31</td>
<td>7 - 8</td>
<td>7</td>
</tr>
<tr>
<td>Tom Sawyer</td>
<td>6 - 9</td>
<td>6.18</td>
<td>5 - 6</td>
<td>6</td>
</tr>
<tr>
<td>Henner's Lydia</td>
<td>3 - 5</td>
<td>5.24</td>
<td>5 - 6</td>
<td>6</td>
</tr>
<tr>
<td>Ringtail</td>
<td>3 - 5</td>
<td>4.37</td>
<td>5 - 6</td>
<td>5</td>
</tr>
</tbody>
</table>
the heading "Survey of Related Literature." Contradictory formula scores as related to other formula scores can be contributed to formula elements and will be discussed later in the study.

Grade placement scores which are contradictory in comparisons with formula results with library selection aids may be attributed to different factors. One factor which might contribute to a discrepancy in score between formulas and selection aid tools was of interest. Most formulas are devoid of an interest count, whereas interest can be taken under consideration when grade and/or age level are listed in library selection aids by "humans" who consider the book's content, format, typography, illustration, and the like. Formulas do not take the foregoing factors under consideration.

Another factor which formulas cannot or do not measure but which is taken under consideration by library selection tools is the reader. The reader, in relation to the reading material, is thought of in the sense of his maturity, experience, motivation and interest or lack of it in the material itself.

Both measures of readability have their merits. It must rest on the user to choose the method which he thinks will assess the material with the greatest validity. In the selection of a formula for reading assessment the following advantages can be found: (1) almost one hundred per cent objective judgment, (2) set rules and guidelines to follow, and (3) approximate grade level is given. Disadvantages of formulas are: (1) some subjective judgments must be made in applying formulas, (2) most formulas do not take into account the reader, (3) some rules for formula application are not stated explicitly, and (4) certain formula elements
make it impossible to receive a readability score no lower than third grade, and (5) one might find the application of the formulas too time consuming.

The main advantage of using library selection aid tools is apparent. This is the most economical method time wise because the reading level is given. Another advantage of selection tools is that the reader and the appearance of the book are taken under consideration during the evaluation. However, this method is most subjective and gives wide reading range scores for books (ex. Peter Pan 4-7, meaning the reading level is 4th through the 7th grade).

The user of the two readability methods should first consider his purposes for wanting the reading material assessed and then decide which method would be feasible. The user should always keep the reader foremost in mind before making final judgments on the readability method to be used.

Element Analysis of the Lorge Formula

The Lorge formula was one of the first easy-to-apply readability formulas and is still among those widely used. By counting the relative number of prepositional phrases index of readability in terms of scores can be obtained. Although this formula is one of the easier to use, the writer found it rather time consuming.

To those familiar with readability formulas, it is obvious that in using a readability yardstick certain "factor counts" must be made. In the Dale-Chall formula there are two factor counts and in the Flesch formula there are three. In using the Lorge formula four counts are required. The continuous checking of the Dale List of 796 Easy Words with
the content material was quite timely. A smaller word count would render
the formula easier and quicker to use; however, by using a smaller word
list a question of the validity of the list would then come into view.

The hardword factor count.--The Lorge formula is not one which is
entirely objective. While applying this formula to material the writer
had to make certain subjective decisions. An example of the foregoing
statement would be in the area of the hardword count. The rule for this
area is, "... whenever an unfamiliar word (a word included on the Dale
List of 796 words) appears in a passage (regardless of the number of
times it appears) it is counted only once." The words "daisies" and
"daisy" were found in a passage taken from Alice in Wonderland. The
question arises of whether to count both words as one unfamiliar word or
two unfamiliar words. The problem here was that both are the same word
but with different endings and in a sense making them different words.
Lorge might have done well to include in this formula procedures to take
in counting unfamiliar words in the same passage which have been changed
by the use of affixes. The user, coming in contact with the unfamiliar
word affix problem, will probably use either of three methods. Method
one and method two are interrelated. The user will count consistently
the unfamiliar words with affixes or, two, he will consistently not in-
clude the unfamiliar words with affixes in his count. By the use of one or
the other method the user will be one hundred per cent accurate or one
hundred per cent inaccurate in his count. The third method that is used
is an inconsistent method which involves counting more affixes and leaving
some uncounted, hoping for at least a fifty per cent accuracy in this area
of the factor count.
Counting compound words in relation with the Dale List of 796 words also presented a problem. The rule Irving Lorge writes for this is, "Compound names of persons or places like New York, United States, St. Louis, Santa Claus, and Van Dyke count as single words." The above rule really applies to compound words which are not found on the list and also to make it emphatic to the formula user not to count compound words as two words. Nowhere does he mention how to count two unfamiliar words coming together forming a compound word. Words such as "daytime," "yourself," "nobody" and "handwriting" were found in passages taken from Huckleberry Finn and subjective judgment had to be used in the counting of these words. The two words when separated were found on the list of 796 words but were not included "joined together" or as compound words. Therefore, in a sense the words should be counted as unfamiliar if they are not found on the list as one word. Then again the compound words should be considered as familiar words because both words can be found on the list—a very confusing situation.

The next question that arises while applying this formula is how to count familiar words with affixes such as: "uneasy," "goodness," and "unhappy," which are found in Cinderella and Heidi. Here again subjective decisions are made.

The sentence count.—The sentence count is also a factor in the Lorge formula. The only reference given in following this procedure is, "Begin at the beginning of the sample and count the number of complete sentences." The Lorge formula made no provisions for conversation which was especially difficult to interpret. For example, the following from Peter Pan may be interpreted as one or two sentences: "There are such a lot of them," he said. "I expect she is no more."
Prepositional factor count.—The Lorge list of prepositions, when compared with preposition lists that are found in grammar textbooks, was found to be incomplete. Because of this, there is considerable possibility of error in the prepositional phrase count.

**Element Analysis of the Flesch Formula**

In this formula, Dr. Flesch used three factors: average sentence length, relative number of affixed morphemes (prefixes, suffixes, inflectional endings), and relative number of personal references.

The count of affixes.—The most serious shortcoming of the Flesch Formula was the count of affixes, which was found to be rather arbitrary, in the sense that two people making a count on the same samples would usually come out with different number of affixes. If one was extremely careful and consulted a dictionary to be certain that all affixes were included, he would find the work too time consuming.

Personal reference count.—The second shortcoming of the Flesch Formula was the count of personal references. From the writer's formula analyses it was found that the personal reference count may not be considered a reliable index of difficulty. For example, when one speaks of "John" and "Mary" and "he" and "she" referring to John and Mary, there is a justification for subtracting from difficulty. This is because in writing about John and Mary we usually say things that are not abstract or general. However, subtracting from difficulty of personal-reference such as "R. J. Thomas" of the automobile industry, or "Senator Austin," when one is writing about atomic energy or the United Nations, does seem a bit inaccurate. If the reader does not know these persons, the difficulty of the written material is not decreased. In fact, these individuals are
no longer personal, they are abstractions. Flooding printed materials with personal references to those "abstract" persons will add little to "human interest" and ease of comprehension. Another example of the abstractness of the personal reference count is taken from Winnie-the-Pooh, page 157. Here the personal reference "it" is being used but it does not refer to a person.

"And what did happen?" asked Christopher Robin.

"When?"

"Next morning."

"I don't know."

"Could you think and tell me and Pooh some time?"

"If you wanted it very much."

To continue discussion on the information of personal references and how unreliable this factor can be in certain instances the writer cites S. S. Stevents and Geradine Stone whose article appeared in the American Psychologist. They reported that after readability formula had been applied to K. Koffka's Principles of Gestalt Psychology, a Flesch score much lower than had been expected was received. In fact, it came out only a little higher than elementary textbooks in psychology. This was quite startling news for them. They wrote:

The Harvard graduate students don't believe it, because they read Koffka and sweat.

... Now how can Koffka, the students' choice for unreadability, score so low? ... A few things appear evident, however, for one thing, Koffka helps his score by peppering his passages with personal pronouns: 5.8 per hundred words. But his "I," "we," and "you," are rhetorical devices—he is actually rarely talking about us or about himself. He is talking about abstracts and complicated relations and he and we get into it as mere guinea pigs in an experiment.1

1Stone and Stevents, American Psychologist, II (July, 1947), 233.
Here is an example taken from Koffka:

In the first cases, real moving objects present in the field, the shift of the retinal pattern leads to behavioral motion of objects, whether I fixate a nonmoving object or follow a moving one with my regard; in the second class, when my eyes roam over stationary objects, such a shift will not have this result. Although the two facts belong closely together, the second one will be fully discussed in Chapter IX, after we have introduced the ego. Here we concentrate mainly on the first, even if we cannot entirely avoid referring to the second. Thus we turn to the theory of perceived motion.¹

This passage has seven personal references per hundred words. According to Flesch's Quick Reference Chart, as presented in Table 2, similar number of personal reference characterizes materials that are standard in difficulty and are comparable to digest magazines.

From the above statements one can readily see that the personal reference factor count needs further research.

Flesch's Quick Reference Chart.--The Flesch formula does not give a grade point average or a readability index score at the end of the formula. Flesch uses a difficulty score and this is converted into the school grade level and also gives a description of the style and type of material being read. The Quick Reference Chart is presented in Table 2.

Element Analysis of the Dale-Chall Formula

The use of the Dale-Chall wordlist in the hardword count.--In analyzing the Dale-Chall formula the writer concluded that a larger wordlist would make a better prediction but at the same time become involved because of its length. One might find a user hesitant in applying this

### TABLE 2

**QUICK REFERENCE CHART**

<table>
<thead>
<tr>
<th>Description of Difficulty</th>
<th>Words in Average Sentence</th>
<th>Affixes per 100 Words</th>
<th>Personal References per 100 Words</th>
<th>Typical Magazine</th>
<th>Potential Audience (Typical Audience 1-Step Below)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Easy</td>
<td>Up to 1</td>
<td>8 or less</td>
<td>22 or less</td>
<td>19 or less</td>
<td>Comics</td>
</tr>
<tr>
<td>Easy</td>
<td>1 to 2</td>
<td>11</td>
<td>26</td>
<td>14</td>
<td>Pulp Fiction</td>
</tr>
<tr>
<td>Fairly Easy</td>
<td>2 to 3</td>
<td>14</td>
<td>31</td>
<td>10</td>
<td>Slick Fiction</td>
</tr>
<tr>
<td>Standard</td>
<td>3 to 4</td>
<td>17</td>
<td>37</td>
<td>6</td>
<td>Digest</td>
</tr>
<tr>
<td>Fairly Difficult</td>
<td>4 to 5</td>
<td>21</td>
<td>42</td>
<td>4</td>
<td>Quality</td>
</tr>
<tr>
<td>Difficult</td>
<td>5 to 6</td>
<td>25</td>
<td>46</td>
<td>3</td>
<td>Academic</td>
</tr>
<tr>
<td>Very Difficult</td>
<td>6 and Up</td>
<td>29 or more</td>
<td>54 or more</td>
<td>2 or less</td>
<td>Scientific</td>
</tr>
</tbody>
</table>

*Reproduction of Flesch Quick Reference Charts for School-Grade-Conversion Score.*
formula because of the number of words used in this list.

Sentence factor count.—In using the Dale-Chall formula one would also encounter a sentence factor problem in trying to judge conversation. For this factor the Dale-Chall formula only states, "Count the number of complete sentences in the sample." Another example of this dilemma can be seen in Winnie-the-Pooh, "What about me?" said Pooh sadly. "I suppose I shan't be useful?"

Use of the Dale-Chall's Correction Table.—Because this formula uses only a word sentence factor it is very economical time wise and easy to apply. The Dale-Chall formula does not give a grade score. One receives a raw score with this formula and refers to a Table to receive the corrected grade level score. Dale-Chall's Correction Table can be seen in Table 3, page 34.

Because only two factor counts are used in the Dale-Chall formula (word and sentence factor) one would be justified in wondering if such factor counts as prepositional phrases, affixes, and personal references in other formulas are necessary.

Comparison of Formula Elements in the Lorge and Dale-Chall Formula

Each formula studied furnished information which would establish the relative difficulty of a number of books. Since readability is important in any book, whether it be an arithmetic, science, social-studies, or language-arts book, the use of these formulas gives one good basis for comparison.

The Dale-Chall Formula and Lorge Formula proved to be similar when compared with the Flesch Formula. Both (Dale-Chall and Lorge Formulas)
TABLE 3

DALE-CHALL CORRECTION TABLE

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

*aReproduction of Dale-Chall's Correction Table for School-Grade-Conversion Score.*

formulas are in agreement as to the type of passage that should be selected in a sample. Dale-Chall and Lorge also agree upon procedures to follow in counting certain words (hyphenated words, contractions, numbers which are written in symbol forms, and compound names) which appear in the selected passages. However, Dale-Chall add one other procedure under this count, a guide for counting initials.

For sentence counting, both formulas stated only to "count the number of complete sentences in the sample." Here, as stated previously, there should have been further explanation as to what should be done in counting conversational sentences.

Dale-Chall and Lorge agree that there should be a count made of the number of hard words in a sample. However, there is some variation on how
each makes the hard word count.

The Lorge formula instructs the user to "Use the word list and cross out in the sample every word on the list, regardless of its meaning (ex. spring—meaning season, jump, water, or steel coil)." Lorge further states that, "The count is of the number of different hard words, so that each hard word is counted only once."

In return, Dale and Chall give this rule, "Words which do not appear on the Dale List are considered unfamiliar. Underline all unfamiliar words, even if they appear more than once." Here a contrast in formulas can be seen. Whereas hard words are counted only once, regardless of the number of times they appear in a sample using the Lorge formula, Dale-Chall's instructions state that they should be counted the number of times they appear in the same sample.

Specific procedures were listed to guide the user while making the word count. Word count procedures which were similar were those governing "common nouns."

Dale-Chall, in their formula, went a step further and listed rules to be followed in counting proper nouns and abbreviations.

Lorge gives a clear set of procedures to follow when counting hyphenated words, "In case of uncommon hyphenated words follow Webster's Unabridged Dictionary (2nd edition). Any hyphenated word is considered as one word if it is listed in the dictionary as a hyphenated word; otherwise it is counted as two words." Here Lorge makes it emphatic that unfamiliar hyphenated words should be counted as two words.

Dale and Chall give these procedures to follow, "Count hyphenated words as unfamiliar if either word in the compound does not appear on the list. When both appear on the list, the word is familiar." Dale-Chall
inform us that we do not have to even bother to look in a dictionary for
words--if they are not on the list then they are unfamiliar. Therefore,
one must count them as hard words. What Dale and Chall do not tell is
if we should count the unfamiliar hyphenated words as one or two unknown
words. And to go a step further--how must the user count the hyphenated
words if one is familiar and the other unfamiliar?

Under the miscellaneous special cases procedures both formulas
agreed that when two or more endings are added to a word, the word should
be counted as unfamiliar. Dale-Chall add two more procedures pertaining
to suffixes and numbers written in symbols. Lorge's additional rules
apply to words of different spelling and words formed by adding "y" to
them.

Lorge adds one additional count to his readability formula, the
prepositional phrase count.

The word count lists given by Dale-Chall and Lorge prove to be in-
teresting and produced information as to the reason why a readability
score of no lower than fourth could be reached.

The Lorge formula uses the Dale List of 769 Easy Words (an earlier
word list by Edgar Dale). The Dale-Chall formulas used the Dale List of
3,000 Familiar Words (developed later than the first list).

Words which are a part of the Dale-Chall and Lorge Formula word
list are those which are considered as our 220 basic sight vocabulary
words and are usually mastered by the end of the first grade. Other words
listed are those words which ought to be a part of the vocabulary taught
on the different elementary grade levels.

To test the theory of the word list being the reason why a score of
no lower than fourth grade could be obtained using the Dale-Chall and Lorge formulas, the writer checked both formulas' word lists with vocabulary lists given for grades one through six. The formulas' word lists were then compared with vocabulary list taken from Morton Botel's *How to Teach Reading*, the Bucks County 1185 Word List (a study made in 1955 of all words, according to grade level, taught in five major basal reading series), and Dr. Durrell's *Improving Reading Instruction* which contained 90% of the words ordinarily used in the written composition of intermediate children and Edward Dolch's Basic Sight Vocabulary List of 220 Words.

On completion of the wordcheck lists the writer found that the Dale-Chall formula, with its word list totaling 3,000, could be broken down thusly:

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>Number of Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer</td>
<td>192</td>
</tr>
<tr>
<td>1 - 2</td>
<td>498</td>
</tr>
<tr>
<td>3 - 4</td>
<td>2,305</td>
</tr>
<tr>
<td>Unclassified</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3,000</strong></td>
</tr>
</tbody>
</table>

The Lorge formula with a total of 769 words yields the following results:
TABLE 5
LORGE WORDCHECK LIST

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>Number of Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer</td>
<td>117</td>
</tr>
<tr>
<td>1 - 2</td>
<td>223</td>
</tr>
<tr>
<td>3 - 4</td>
<td>429</td>
</tr>
<tr>
<td>Total</td>
<td>769</td>
</tr>
</tbody>
</table>

The writer also found that the higher the hard word count the higher the readability index will be and the lower the hard word count the lower the readability index. This was especially true for the Dale-Chall formula because the crux of this formula is the hard word count.

Therefore, if the user applies the Dale-Chall formula, which considers familiar all words up to the fourth level, to a second grade book and adds the constant (3.6365) in the end computation the readability index produced could be no less than 4.00. The constant of 3.6365 in itself contributes three and a half years or a score of third grade six months to the readability score initially.

The Lorge formula adds a constant of 1.9892 in the end computation to the values. This constant really adds a score of approximately two years to the readability index received. Here again an initial grade score (the constant) is added to the readability index score before it is received. Although it is very doubtful that a readability index less than 3.5 can be received with this formula it is, however, possible that the Lorge formula might produce a lower readability index than the Dale-Chall
formula because of its lower constant. The reader at this point is referred to Table 1 for comparison of the more significant scores of the two formulas dealing with this issue. Notice should be given especially to Peter Pan, Cinderella, Huckleberry Finn, and Ringtail.

Lorge's shorter word list, which he incorporates into his formula, cuts down on the number of familiar words, thus increasing the hard word count and also the readability index score.

Lorge, as stated earlier, adds a prepositional phrase count to his formula in addition to the hard word count. The effect that the prepositional phrase count will have on the readability index will of course vary according to the difficulty of the passage. It was found that the easier the passage the lower the prepositional phrase count; the more difficult a passage, the higher the prepositional phrase count.

Comparison of Formula Elements and Findings of the Flesch Formula with those of the Lorge and Dale-Chall Formulas

The Rudolf Flesch formula is similar to Lorge and Dale-Chall formula to the extent that he also directs the user to select samples by a numerical scheme.

Flesch too, includes a word and sentence count in his formula. However, his guidelines to follow in applying the two counts differ from Lorge and Dale-Chall. Lorge and Dale-Chall instruct the user to use the entire page of each sample selected. Flesch instructs the user to use a certain number of words from each page from which the sample is taken: "Take each sample and count each word in it up to 100. Count contractions and hyphenated words as one word. Count as words numbers or letters separated by space."
Flesch's rule for the sentence count helps the user decide how conversation should be counted, whereas Lorge and Dale-Chall in their formulas neglected this fine point. Flesch, on sentence count, states, "Find the sentence in each sample that ends nearest to the 100-word mark—that might be at the 94th word or the 109th word—count the sentences up to that point and divide the number of words in those sentences by the number of sentences. Do this for each sample. In counting sentences, follow the units of thought rather than the punctuation; sometimes sentences are marked off by colons or semicolons instead of periods—like these. But don't break up sentences that are joined by conjunctions like and or but."

The Flesch formula does not include a hard word count; instead, a count of affixes and personal references is included. On comparing the count of affixes to hard word count (included in the Dale-Chall and Lorge formulas) in relation to time the writer found the count of affixes to be more time consuming. Dale-Chall and Lorge furnish a list of words the user can readily refer to. Although Flesch furnishes a list of the most common affixes, it is not exclusive and the user must continuously refer to the dictionary. The count of affixes, as previously stated, in most instances turned out to be rather arbitrary. This is so because in many instances certain affixes may be overlooked by the user and considered as the "whole" word itself. For example: real— the "al" should be counted as an affix; ignorant— the "i" is counted; in said— the "d" is counted.

If a thorough count of affixes is made, one could not help but receive a high index score. This is so because the higher the count of affixes the higher the index score, and the user will certainly receive a high affixes count if every "d" added to "said" is counted in juvenile
fiction because words such as these make up the vocabulary of the story book.

The count of affixes when compared with the prepositional count still proved to be more time consuming.

The Flesch formula uses a personal reference count. The personal reference count is also referred to as the level of interest count. This count is taken under the assumption that the more personal references the higher the reading interest and the easier the material will be. His personal reference count consists of counting all names, nicknames, pet names, personal pronouns and words that deal with human beings or their relationships; an example of this: man, woman, girl, cousin, husband, etc. Flesch does not consider other words, like "teacher" or "doctor" as personal reference words.

The personal reference count affects the numerical computation of the formula. In computing the reading index for the Lorge and Dale-Chall formulas the user must add all values and constants. The Flesch user adds only the values given for the average number of personal references from the foregoing value. After this score is received the user has to subtract the constant from that score. Then by using the Quick Reference Chart the readability index is received.

A summary table comparing formula elements to enable readers to see formula differences at a glance can be found on page 42. Elements which formulas agree on or incorporate as formula factors are checked.

**Evaluation of the Dale-Chall, Flesch and Lorge Readability Formulas**

The Lorge, Flesch, and Dale-Chall formulas tend to assign similar
TABLE 6
SUMMARY TABLE OF FORMULA ELEMENTS

<table>
<thead>
<tr>
<th>Formula Elements</th>
<th>Dale-Chall Formula</th>
<th>Lorge Formula</th>
<th>Flesch Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type and Length of Passage Used</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Procedure for Word Count</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Procedure for Sentence Count</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Prepositional Phrase Count</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Count of Affixes</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Personal Reference Count</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Hard Word Count</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

grade scores when assessing readability levels of juvenile fiction.

Neither formula proved to be outstanding in grading primary books on or below the fourth grade level. If one would note in Table 1 there is no grade score below fourth grade. Although a score 1-4 was listed for Adventures of Pinocchio and Winnie-the-Pooh the overall grade score on these books is fourth grade. The reason for this is that the raw scores given by the formulas were right on the borderline—they were 4.9 and 4.6 respectively. The raw score grades can be seen in Table 1.

For general use in evaluating books for elementary grades the Dale-Chall and Lorge formulas are easier for the average person to use than is the Flesch formula. In selection of a formula out of the two previously mentioned, the Dale-Chall formula would be recommended first. There
might be fewer errors made on the user's part because of the two simple factor counts which are to be applied. If a large number of books is to be measured, finding the time to run any of these formulas would still be a difficult problem. The Lorge formula should include a longer prepositional list with its formula because this count proves to be rather arbitrary at times, though not as much as the count of affixes. Lorge and Dale-Chall formulas need further study in their use of constants. Their constant factor seems to be one of the main reasons for such high readability scores.

Flesch count of affixes seems to be the reason for a readability index of no lower than fourth grade being received. Flesch uses a constant of only .75.

The Dale-Chall, Flesch and Lorge formula proved to be valid measures of juvenile fiction for grades four, five, six and above. Because of the great chance of errors which can occur while applying these formulas (factor count errors) the writer finds it mandatory to add that these formulas are valid to the extent of giving the user an approximate grade score.

The formulas should not be applied to books lower than the third level because a readability index for these books will not be received. There was close agreement between the library selection aid tools grade placement and the formulas grade placement scores.

The user of readability formulas should keep in mind that no readability score is the complete answer to the proper evaluation of reading materials. However, if carefully and conscientiously used, any of the three formulas which have been studied can provide helpful information regarding the relative difficulty of the books which are being considered.
CHAPTER III

SUMMARY AND RECAPITULATION

The idea behind readability measurement is the matching of reader and printed material. It assumes that readers differ in their ability to read and that the printed material in turn varies in the amount and kind of ability needed to read and understand it.\(^1\)

Presently, formulas are the best means devised for measuring certain aspects of reading difficulty and for comparing various pieces of writing.

Readability has become a problem of prime importance for a number of reasons: (1) the trend to emphasize reading as the chief aid to learning appears to be on the increase, (2) a better professional understanding of the relationship between the readability of instructional materials and frustrations in reading abilities at any grade level, (3) discrepancies between grade scores achieved on standardized tests and the ability to read instructional materials designated for that grade level, (4) discrepancies between the readability of books with the same grade level designations, (5) the trend to reduce the vocabulary load of basal textbooks, especially in grades one and two, (6) the slow extension of practices in the direction of the experience approach to reading, and (7) the commercial values of both tradebooks and texts.

\(^1\)Chall, op. cit., p. 44.
No one maintains that any readability formula is perfect. Research is still being conducted and doubtless will continue for years. In the meantime, these formulas can be extremely helpful in measuring reading difficulty on an objective basis. With the measure given by a formula, we have the means by comparing the difficulty of one book with another.

A study of the science of measuring readability can give us many clues to the success and failure of certain books for children. The effort of the classroom teacher and librarian could also be facilitated by the products of these studies in the form of readability index of books or a National Bureau of Readability Standards which would provide comparable ratings on all school instructional materials.

Although the study of the science of measuring readability involves many areas, the main area of this study centered on determining the validity of three formulas as measures of readability: Rudolf Flesch's Yardstick for Measuring Readability, The Lorge Formula for Estimating Difficulty of Reading Materials, Dale-Chall's Formula for Predicting Readability.

Purposes for the study were to: (1) determine which formula of those used in this study is more consistent with the readability levels assigned by specific library selection aids, (2) determine which formula(s) can be used in measuring readability with the greatest efficiency, (3) analyze elements which are counted in the computation of each of the formulas to be used, and (4) compare formula results and elements based on an analysis of the formula elements.

In order for the study to evolve, a list of thirty books randomly selected from various book award lists (John Newberry, Spring Book
Festival, Boy's Club of America, Best Books of the Year, Horn Book Magazine, and Children's Catalog), had to be compiled. Grade placements were assigned to each book by library selection aid tools. Books in which library selection aid tools were in agreement on grade placement were selected for this study.

The number of books selected for the study was twelve. Library selection aid tools used for the study were: Best Books for Children, Good Books Too Good to Miss, Good Books for Children, and Children's Catalog.

Each book selected for the study was subjected to three measures of readability: the Dale-Chall, Flesch, and Lorge Readability Formulas. Reading grade levels suggested by library aids for each book were compared with each of three computed readability measures charted for the purposes of comparison and evaluation.

A systematic and thorough study was made of each formula for the purposes of analyzing, comparing and evaluating the three readability formulas. After collection and organization of research data a summary of findings, conclusions, recommendations and implications were presented.

**Summary of the Survey of Related Literature**

Research of literature on readability produced the following data:

1. The idea of gradation in a series of texts for reading instruction was first introduced in the McGuffey readers. The selections were assigned to lower or higher levels largely on the basis of author's judgment. A check against pupils and books proves that grade for grade, the McGuffey books are relatively difficult.

2. In 1923, Lively and Pressey conducted one of the first experiments to find the elements that made for difficulty
in reading materials. Lively and Pressey measured the vocabulary difficulty of books by assigning the Thorndike frequency number to each different word and taking the average of these numbers. This measurement did not directly relate the books to reading comprehension. Using only one kind of measurement, it merely arranged books in relation to one another in terms of vocabulary difficulty of the material.

3. It was not until Vogel and Washburne produced their Winnetka Formula in 1928 that a formula was created which related elements of difficulty in written materials to specific reading levels. The term "readability" was coined some years later.

4. Since the Vogel and Washburne formula, many workers have devised formulas based on the statistics of regression equation, which are variations of the Washburne-Vogel technique.

5. Klare in 1950 estimated that 34 formulas or methods for determining the reading difficulty of printed material had been devised.

6. Many studies prior to 1950 were concerned with testing of hypotheses rather than with critical evaluation. They reflected an interest in exploration rather than validation of readability measures.

7. The evaluative studies that have been done in readability have proven to be contradictory to each other in certain instances.

8. Some of the more recent studies of readability formulas found the following results:

   (a) Readability formulas are not universally applicable. They are especially ill adapted to poetry and plays.

   (b) Flesch, Russell and Fea, and Chall agree that readability formulas currently available cannot measure readability with a high degree of effectiveness.

   (c) Existing readability ratings produce approximate assessments of readership levels on factual material. However, on creative material, they are relatively ineffective.

---

1Vogel and Washburne, op. cit., p. 329.
2Hildreth, op. cit., p. 329.
3Klare, op. cit.
(d) No readability formula is effective at all levels of difficulty. In addition, each formula is most reliable when used with material of the same level as that on which it was standardized. This further limits the use of each formula.

(e) Analyzing difficulty levels in literature is an aid to teachers, publishers, librarians and parents. It enables them to match reading material to individual level of students.

Findings of the Study

The following findings were drawn from the study:

1. All of the methods required analyzing of samples from the reading matter for certain features of language construction, the two factors most commonly employed being (a) vocabulary load, and (b) sentence structure.

2. The difficulty level of reading material increases in proportion to number of different words in the text, the use of long sentences and long, complicated paragraphs, and the frequent use of less common polysyllabic words. Simple material was found to be characterized by concreteness, easy vocabulary, and simplified sentence structure with few dependent clauses and a small proportion of prepositional phrases. This finding can be seen readily in comparing the language used in Cinderella, which has a library selection aid grade placement of 1-3 with that of Huckleberry Finn, 6-9 library grade placement.

3. Dale-Chall, Lorge and Flesch formulas are capable of producing a readability index for only fourth grade and above. Dale-Chall and Lorge above fourth index scores were contributed to their constants of 3.6365 and 1.9892, respectively. The second factor contributing to a high readability index score is the wordlist which the Dale-Chall and Lorge formulas use. Dale-Chall uses the Dale List composed of 3,000 words and Lorge uses the Dale List composed of 769 words. These lists were made up of words ranging from primer up to fourth. Flesch used a very low constant of .75; however, the count of affixes increases the readability index for this particular formula.

4. The Lorge Formula results produced slightly lower scores than the Dale-Chall and Flesch Formulas, although these scores were not lower than the selection aid tools scores. However, the Lorge formula did produce only one score lower than the selection aid scores, this score being computed for Huckleberry Finn, selection aid score 6-9 and Lorge formula score 5.31.
5. The Dale-Chall's Formula for Predicting Readability and the Lorge Formula for Estimating Difficulty of Reading Materials proved to be similar in application when compared with the Rudolf Flesch's Yardstick for Measuring Readability.

6. In the Dale-Chall and Lorge formulas no provisions are made for partial sentences and conversation is especially difficult to interpret. An example of this was taken from Winnie-the-Pooh, "What about me?" said Pooh sadly. "I suppose I shan't be useful?"

7. Two findings drawn particularly from the Lorge formula concerning the word and prepositional phrase count are:

   a. Shorter word list (Dale List of 369 Easy Words) incorporated into the Lorge formula decreases the number of familiar words, thus bringing up the hard word count and also increasing the readability score.

   b. The effect the prepositional phrase count asserts on the readability index will vary according to the difficulty of the passage. In other words, the easier the passage the lower the phrase count and the harder the passage the higher the phrase count. A very good example of this can be found in comparing Adventures of Pinocchio with a library placement of 3-6 and Tom Sawyer with a 6-9 library grade placement.

8. Rudolf Flesch's guidelines in the Flesch formula application are more explicit in giving directions for factor counts than the Lorge and Dale-Chall Formulas. Although Flesch formula presents explicit guidelines, there are still certain shortcomings of this formula:

   a. The most serious shortcoming of this formula was the count of affixes, which was rather arbitrary. The count of affixes was also found to be time consuming.

   b. A second shortcoming was the count of personal references. The writer could not consider this factor as being a reliable index to difficulty if applied to certain materials. The count of personal references affects the readability index to the extent of the higher the personal references count the lower the readability index, and the lower the count the higher the readability index.

9. Formula analyses and evaluation produced these findings drawn from the Dale-Chall, Flesch and Lorge readability formulas.

   The formulas do not:

   a. Give any measure of conceptual difficulty in textual material.
b. Take into consideration the way the material is organized.

c. Allow for variation in the meaning of multiple-meaning words.

d. Accept the fact that a fresh or unusual word may make a sentence or idea clearer than a commonplace word. This was quite prevalent in *Tom Sawyer* and *Huckleberry Finn*.

e. Vary their ratings in terms of different interests which persons may have at different developmental levels or in individual activities.

f. Take account of physical factors such as format and illustrations. This point is well taken because in primary books such as *Adventures of Pinocchio*, *Winnie-the-Pooh*, *Cinderella* and *Henner's Lydia* the physical factor will also affect interest in and reading of the book.

10. All formulas' index scores gave close approximations for juvenile fiction with grade placements above fourth grade. However, neither formula proved to be outstanding in scoring books below the fourth grade level.

**Conclusions**

1. The Lorge Formula for Estimating Difficulty of Reading Materials produced readability scores closer to the selection aid tools than the Flesch or Dale-Chall formula—though not to a significant degree.

2. The Dale-Chall's Formula for Predicting Readability proved to be more economical for use than the Flesch or Lorge formulas. This is attributed to the fact that this formula has only two factor counts.

3. For general use in evaluating juvenile fiction books for the elementary grades the Dale-Chall and Lorge formulas are easier for the average person to use than the Flesch formula.

4. The use of these formulas in measuring a large number of books would produce a time problem. No formula is economical, time wise.

5. The Dale-Chall, Flesch and Lorge Formulas are valid measures of juvenile fiction for grades four, five and six, and because of the great chance of error which can occur while applying these formulas, it is necessary to add that these formulas are valid to the extent of giving the user an approximate grade score.
6. Formulas should not be applied to books of the third level and lower because a readability index for these books will not be received. The result of this is from the effects of the affixes, hard word count and numerical computation of the constant.

Implications

The implications drawn from the findings of this study follow below.

1. Future research in readability might take three directions: (a) basic research on critical, underlying variables in the process of reading and learning from print, leading to some kind of a theoretical framework for understanding them; (b) attempts to identify and/or incorporate new factors important to accurate readability measurements into formulas; (c) refinement of factors and methods now used in readability formulas.

2. Because of the widespread use and observation that formulas are often accepted as being more accurate and valid than they are, a method for the exchange of information (concerned with readability formulas) and general distribution of summaries should be available to keep users informed of the state of affairs and correct misconceptions about what formulas "can" and "cannot" be expected to do.

3. The use of readability criteria offers new possibilities in better grading and preparation of materials for use at different levels of readership.

4. Trend for the future use of refined readability formulas in relation with juvenile reading materials:
   a. make it possible for all published materials for children to be identified by a reading-ease index determined by some standard technique.
   b. Therefore, the typical classroom teacher who is busy with a roomful of active youngsters can scarcely be expected to undertake the time-consuming task of rating children's reading material according to statistical formulas. This must be left to research workers, those who construct texts, and the librarians who are specialists in children's literature. However, the concept of readability can be utilized by teachers in guiding children's reading and in assigning texts. The teacher who has the basic principles in mind is in a better position to make informal judgments of the suitability of material for a particular grade or
an individual pupil, and to construct readable materials for them. The readability principle can serve as an aid in adapting instruction to individual differences.

c. Readability formulas can be used in choosing materials to be read aloud to children and in preparing talks for children.

d. Teachers will find readability formulas useful in modifying and rewriting reading materials for slower learners.

Recommendations

In accordance with the findings, conclusions, and implications, the following recommendations are made:

1. There needs to be continuous refinement of factors and methods now used in readability formulas. Refinement of existing readability factors and methods should be of at least three kinds: study of style factors other than words and sentences, examination of the exact relationships of words and sentences to difficulty and the use of new word lists.

2. The readability formulas studied should try to identify and/or incorporate, through research, new factors which are important to accurate readability measurement. New factors suggested for future investigation:

   a. Conceptual difficulty--taking into account the different shades of word meanings, comprehension of ideas presented in the material, and retention of these ideas.

   b. Material organization--the way details, facts, etc., are presented in logical order.

   c. Physical factors--legibility, illustration, format, etc.

   d. Reader's attitude--interest of the reader in the material, agreement of reader with the point of view expressed in the material.

   e. Style of writing--whether the writing is in good or poor taste.

3. Additional research is needed in readability to discover some of the critical, underlying variables in readability as they affect the process of reading and learning from print. Variables for future research, which are listed below, can be related to the following factors (these factors as mentioned earlier are also in need of extensive research) comprehension, learning, or retention of material:
a. **Word frequency (or familiarity) and length**—Increased frequency and shorter length of words mean the reader will have a larger number of words within his memory span. This increased span is related to comprehension, learning and retention measures; so also is the tendency for more frequent words to have more common (familiar) dictionary meanings.

b. **Sentence length and redundancy**—Short sentences and highly redundant material are important because of a reader's limited span. This, once again, affects comprehension, learning, and retention measures.

c. **Educational level and special reading experience**—As these increase, comprehension, learning, and retention measures tend to show a corresponding increase.
BIBLIOGRAPHY

Books


Articles and Periodicals


54


Lorge, I. D. "Reading and Readability," Teachers College Record, LXI (November, 1949), 90-97.


Publications

Unpublished Materials