Summer 1971

Estimating the Readability of Primary Books: A Comparative Study of the Formula Method and Subjective Appraisals

Mary E. Ball
Central Washington University

ESTIMATING THE READABILITY OF PRIMARY BOOKS: A COMPARATIVE STUDY OF THE FORMULA METHOD AND SUBJECTIVE APPRAISALS

A Thesis

Presented to

the Graduate Faculty

Central Washington State College

In Partial Fulfillment

of the Requirements for the Degree

Master of Education

---

by

Mary E. Ball

August, 1971

APPROVED FOR THE GRADUATE FACULTY

_____

Azella Taylor, COMMITTEE CHAIRMAN

_____

Doris Jakubek

_____

Michael S. Brunner

ESTIMATING THE READABILITY OF PRIMARY BOOKS : A COMPARATIVE STUDY OF THE FORMULA METHOD AND SUBJECTIVE APPRAISALS

by

Mary E. Ball

August, 1971

This paper surveys research on the factors that influence the readability of primary books, the efforts that have been made to measure it, and the comparative reliability of subjective judgment and objective analysis of the difficulty of certain primary books. An appraisal by the Fry Readability Graph is recommended for a first approximation of the difficulty of a book, to be followed by judgment decisions regarding concepts, interest and other considerations, to make a suitable match of book to child.

ACKNOWLEDGMENTS

I feel especially privileged to have had Dr. Azella Taylor as chairman of my committee. I am indebted to Dr. Taylor and to Mrs. Doris Jakubek for their expert guidance and help in "routine" matters, but, even more, I am grateful to them for their inspiration, their patience, and their availability in time of need. I also wish to thank Mr. Michael S. Brunner for agreeing to serve on my committee.

LIST OF TABLES

TABLE OF CONTENTS

Chapter Page

Chapter Page

CHAPTER 1

INTRODUCTION

THE PROBLEM

The task of selecting a suitable library book for a given child often entails considerable research. A casual inspection of a book's format, illustrations and content may indicate its possible usefulness, but as a basis for recommending the book, such information is too meager and unstable. The essential question that a teacher needs to ask is how readable the book is for a particular child, as to the ease or difficulty with which it can be read and understood.

The problem in the search for an authoritative statement as to the difficulty level of a book becomes complicated as various sources reveal different designations for the same book. The teacher is confronted for example, with the following estimations: (1) the publisher's label of 0508 on the book jacket; (2) the Children's Catalog classification as "Easy," with the further breakdown, "k-4;" (3) a trade journal's estimate: "Ages 5-9;" (4) the pronouncement by one panel of experts that "it can be read with ease in second grade;" and finally, (5) a formula evaluation, 2.8. All sources agree that this is a primary book!

De-coding the above labels will probably bring the teacher no closer to the exact readability level, for the following considerations must be weighed:

1. The publisher's estimate may be optimistic.

2. The term "easy" is purely relative and likely to be to be broadly inclusive.

3. The "k-4" designation gives no indication of the level at which a child can read the book independently, if indeed this is not a read-aloud at all these levels with top interest at fourth grade.

4. "Ages 5-9" may span as many gradations of ability as are found from kindergarten through fifth grade.

5. That children in one grade can read a book with ease does not in itself establish a difficulty level.

6. If 2.8 indicates an average grade level, it follows that half the book is more difficult; how much more is not disclosed.

7. It would be helpful to know whether the 2.8 represents an instructional or an independent reading level.

8. It would not be sound to assume an arbitrary designation within the generalized limits, since each description appears to have derived from a different cornerstone and to imply the use of different measurements.

With such reservations the picture is still more confused. It is difficult to see what rationale was employed in these appraisals.

Statement of the Problem

Broad generalizations as to the age or grade levels for which a book is suited do not give any real indication of the book's degree of

difficulty. What the generalities do indicate is that different standards of judgment have been used, as well as different modes of measurement; and the meaning of grade level designation has been left open to conjecture.

Is there a way in which the aspects of readability can be brought into sharper focus, adequately measured, and rated in explicit terms? The problem must be clarified through research.

The purpose of this study, therefore, was to discover what research reveals about the factors that influence the readability of primary matter, what efforts have been made to measure it and with what success, and the comparative reliability of subjective judgment and objective analysis of the difficulty of primary books. An important aspect of the investigation was the attempt to search out empirical studies that would lead to a consensus of opinion regarding the term "grade level."

Importance of This Study

Betts has repeatedly emphasized to teachers the necessity of distinguishing between the child's instructional reading level and his independent level. "Being aware of these differences is the starting point," he declares. Then he cautions not to set the bar higher than the child can jump. (3:143).

But over and over again, the bar has been set too high. Harris states that one of the reasons why so many children place reading down

on the list of their leisure-time activities is that, for the most part, the books which they have been given to read have been too difficult to allow for easy and enjoyable reading. Poor readers are further discouraged to find that among the stories which are easy enough for them, many have interest appeal only for younger boys or girls. (29:474).

Experienced teachers and librarians agree with Kottmeyer's observation that "most of the attractive trade books are too difficult for remedial readers, for little attention is paid by these writers to the problem of reading difficulty." (32:186). This leaves it up to the remedial reading teacher to determine the difficulty of the books available for his use.

Knowledge of the difficulty level of a book has a particular significance for a remedial reading teacher. Following diagnosis of a child's skills in reading, correction of the reading disability requires that the child's reading be brought step by step from one level of competence to another, and he needs reading matter along the way that will build his confidence. If a teacher urges him to "try this one," the book needs to be within his capability or he will become frustrated at the difficulty and quickly lose interest. While it is true that high motivation can carry an able child through reading matter normally above his range, there comes a point where interest will flag in the face of formidable vocabulary, sentence structure, context or concepts. With a disabled reader, this point comes "early in the game." It is important that he not be frustrated by material too difficult for him, such as

"k-4" might very well be, for this undermines his confidence and creates an unfavorable attitude toward the next offering. A teacher cannot afford much of this trial-and-error matching, for he is, in effect, sabotaging his own efforts to foster the child's pleasure in reading.

The need for information on the precise difficulty level of a book, expressed in meaningful terms, is, then, doubly important to the remedial teacher if he is to make a satisfying match of book to child.

Insofar as possible, the true meaning of the term "grade level" as a designation of the degree of difficulty, needs also to be explored and its connotations made clear.

It is reasonable to assume that no matter how unique the book, it shares, with others, in definite characteristics of content and structure that are amenable to classification and some kind of difficulty rating. This implies the existence of a criterion of comparison and of some mode of measurement and interpretation. The teacher must survey the field to discover an acceptable criterion and valid tools of analysis that he can use himself. He must acquire background information on readability, the factors that influence it, and the efforts to measure it, in order to reach a sound conclusion.

Limitations of the Study

The literature on the measurement of readability and its

applications is extensive. Reviews of research studies in the field will necessarily be limited to those having significance for the analysis of primary books.

A major part of the investigation will be concerned with the identification of factors that influence readability of primary matter, and their weighting, in formula fashion, into tools for measurement. This study will focus on the Spache Readability Formula and the Fry Readability Graph, the two most recently designed formulas that are applicable to primary books.

This investigator's interest in the subject was an outgrowth of a research paper on the availability of high interest-low vocabulary trade books for disabled readers in grade three. Only those titles were considered that had been adjudged by experts to have an interest level at least two grades above vocabulary level. A list was compiled from various sources, of approximately seventy-five high interest books assumed to be for independent reading at third-grade level. Comparisons of various appraisals will be restricted to books from this list.

Data tabulated will be limited to estimates of the reading difficulty of each book according to the publisher (if available), the Children's Catalog, certain reading experts, and this investigator's application of the Spache and Fry formulas.

Comments on publishers' appraisals of their trade books for primary grades will be with reference to the seven replies sent

in answer to the writer's questionnaire.

DEFINITIONS OF TERMS

Readability

Although there is no precise definition of the term, the delineation of its aspects by Chall has been widely accepted by researchers:

> In the broadest sense, readability is the sum total (including the interactions) of all those elements within a given piece of printed material that affects the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimum speed, and find it interesting. (11:7).

Readability in the more limited sense, referring specifically to reading ease or comprehension difficulty, is the concept pursued in this investigation.

Independent Reading Level

This is the highest level at which the child can read "on his own," with full understanding and freedom from frustration. At this level the child should be able to recognize ninety-nine per cent of the words he encounters and to comprehend at least ninety-five per cent of what he reads. (Betts 2:122).

Instructional Reading Level

This is the highest level at which the child can read satisfactorily under teacher supervision. The child may require help on the recognition of words, but never more than five per cent. His comprehension will be at least seventy-five per cent. (Betts 2:122).

CHAPTER 2

REVIEW OF THE LITERATURE

Chapter Two is devoted to a review of selected studies relating to the readability of primary books and is divided into the following sections:

1. Factors Affecting Readability
2. Various Methods of Determining Readability
   a. Professional Judgment
   b. Grading by Publishers
   c. Booklists
   d. Objective Measurement
3. Readability Formulas
   a. Internal Factors Studied
   b. Readability Formulas Applicable to Primary Books
   c. Effectiveness of Formulas in Determining Difficulty of Books
4. Review of Recent Advances in Readability Research
5. Needed Research
6. Summary

FACTORS AFFECTING READABILITY

Webster's Unabridged Dictionary defines "readable" as "legible . . ., easy to read, because interesting or pleasing. . ., that

permits or admits of reading." When a book reviewer states that a book is readable, he may be referring to its legibility, its ease, its power to interest, or some combination of these qualities. Gray emphasized the "individual characteristics of particular readers" as a factor of readability, pointing out that the term "readable" expresses a certain relationship between these characteristics and the qualities inherent in what is read. (26:492).

Educators are aware that interest, for example, does not depend entirely upon the subject matter used; it depends also upon such mechanical factors as size and style of type, length of the book, and clarity and color of the illustrations. It depends, too, upon the expressional elements in the book, one aspect of which is difficulty.

Ease of reading or understanding depends not only on such expressional elements as vocabulary and sentence structure but also on the reader's interest in the subject matter. Factors of format may also affect ease of understanding.

The literature abounds in studies relating to one or more of these factors of readability, but the area in which there is the most research deals with the ease-of-reading factor.

VARIOUS METHODS OF DETERMINING READABILITY

The problem of matching the readability of the material and the reading ability of the child has been approached in numerous ways by both objective and subjective means.

Professional Judgment

The earliest and most common procedure of determining readability was based solely on judgment, as reported by Chall. (11:9). Before any scientific prediction of difficulty was devised, teachers or librarians made recommendations based on their own experience. They were probably influenced by grade or age designations assigned to books by editors and publishers, who in turn used their judgment to appraise the difficulty of the book. In the opinion of Spache, such experts "are undoubtedly correct more often than they are grossly wrong;" but, lacking "a known point of reference against which intuition and judgment can be compared, their estimates must be [only] approximate." (41:26).

Grading by Publishers

Experienced teachers have learned that grade-level designation indicated by publishers may often be too low when compared to actual difficulty. Examination of the books themselves reveals, however, that some publishers have scrupulously rated their primary books according to a word-frequency count using the Thorndike (46), Gates (24:625-642), or Dolch (18:123-129) wordlist as criterion. Others rely entirely on judgment in their appraisals of readability. (See Appendix, page 85). Still others state in their sales brochures that the Spache Readability Formula (41:142) is used to rate books of third-grade difficulty and below.

It is interesting to note in turn, that Spache in preparing his own formula adopted the publishers' grade-level designations for the one hundred fifty-two books used in his study.

Book Lists

There are many lists extant of recommended books for readers of various ages and interests. Most of the recommendations are based on carefully considered opinions of one or several trained experts. Spache observed, however:

> As might be expected, these experts differ among themselves and with other groups of trained observers . . . .The various book lists and indexes now available represent a pooling of opinions of varying degrees of inaccuracy. (41:27).

Washburne and Vogel were the first to determine empirically the reading difficulty of juvenile books. (Chall 11:133). They expressed the level of each book by the average reading ability of children who read and enjoyed it. This information was published in 1926 as the Winnetka Graded Book List and was later revised to include books of third-grade difficulty.

There is a notable scarcity of book lists prepared for teachers of children retarded in reading. Chall (11:135) observed that "if the books listed were suitable in difficulty, they were usually too immature in content and format; whereas, if they were suitable in content and format, they were usually too hard." Investigation for this study confirmed Chall's statement, with the added finding that the scarcity was

most pronounced at the primary level.

One approach to the problem was the special production of books with more mature content and low reading difficulty. One of the first attempts to produce such books was made by Thorndike in the late nineteen thirties. It is Thorndike's specification for the selection of appropriate books that must be noted in this connection, for he was the first to enunciate the principle that a book is suitable if the child meets approximately one unknown word in one hundred, or better still, one in two hundred known words. (47:127). Betts was later to characterize a similar degree of competency as the child's "independent reading level" and to point out that books for such purpose might be as much as two grades lower in difficulty than those at his instructional level. (2:122).

Special book lists conforming to these specifications are occasionally found in textbooks on the teaching of reading, professional periodicals, or separately available from a college of education, or selected from compilations by professional reviewing agencies.

Objective Measurement

Over the past four decades researchers have gathered data enabling them to devise more objective procedures for evaluating printed matter. On the premise that the key to reading ease or difficulty lies primarily in the language components of the material itself, individual researchers have (1) chosen a criterion of known comprehension

difficulty, (2) analyzed the criterion to determine the rank order of its comprehension variables, and (3) weighted its most prominent or most measurable variables into a formula, or multiple regression equation, that can be applied to other material to rate its difficulty.

Studies of the development and applications of formulas comprise a large part of the literature on readability.

READABILITY FORMULAS

Internal Factors Studied

Of the several types of factors identified by investigators as contributing in some degree to the difficulty of a passage, only two have been embodied in formulas applicable to children's literature: (1) vocabulary load, and (2) sentence structure.

Vocabulary load. Lorge (34:405) noted that "one or more measures of vocabulary load is used as a predictor in every study of readability." The two principal measures are word frequency and word difficulty.

Most of the early studies reported that materials with fewer different words were easier than materials with a higher percentage of different words. (Gray 26:492-493). The first formula of note to be applicable to primary books, the Washburne and Morphett 1935 revision of the 1928 Winnetka formula (49:355-364), indicated positively that diversity appeared to be the best predictor of difficulty in children's materials,

especially in the lower grade levels.

Vocabulary difficulty has been measured in different ways. A common way was the use of some basic word list like Thorndike's, (46) the Stone revision of the Dale Easy Word List of 769 words (41:146-148) or Dolch's list of the First 1000 Words for Children's Reading. (18:123-129). The words within a particular list were considered easy; those not contained in the list were considered hard. Readability formulas using this approach to word difficulty were the Washburne-Morphett, Spache, and Dolch.

Another way of measuring vocabulary difficulty was by word length: the number of monosyllables or polysyllables, or the number of syllables per hundred words.

As early as 1930, Johnson (30:284) presented evidence that the percentage of polysyllabic words in a passage is a measure of the difficulty that children encounter in reading it. His formula took cognizance of no other loading factor. In 1952 Gunning (27:35) found that the portion of words of three syllables or more is the best key to word load. Although Gunning's formula was not applicable to primary books, his findings relating to word length were interesting and relevant:

> Word length is closely related to both familiarity and to abstraction. Among the 1,000 words E. L. Thorndike . . . found to be used most often, only 36 are of more than two syllables. In Dale's list of 3,000 most familiar words, only one out of 25 is of more than three syllables. On the other hand, among words beyond the 20,000 most often used, two out of every three are of three syllables or more. (27:36).

Wheeler (51:397-399) in 1954 added new data to earlier findings of Johnson in recognition of the then-current vocabulary norms, and together with a sentence factor, evolved a formula for evaluating books of from primer through fourth-grade difficulty.

Fry in 1968 (23:535) gave weight to the polysyllable factor in his Readability Graph.

Sentence structure. Almost every investigator who studied factors other than vocabulary found sentence structure significant in predicting difficulty; the degree of significance, however, was variously estimated.

Wheeler alone gave equal weight to "unit length," similar to sentence length, stating in explanation that "in every case throughout the nine series of books examined, there was a clearcut increase in sentence length from grade to grade." (51:398).

In 1959 Stolurow and Newman reported a factor analytic study of language elements indicating that a sentence factor was of "lesser but almost equal" importance with a word factor in accounting for variance in readability. Thus computer-extracted results affirmed an assumption which several investigators have used in their formulas without dependable data to support it. Of greater significance to an investigator of high-interest low-vocabulary books for primary children is another finding of Stolurow and Newman that "the relative predictive value of these factors changes with variations in the ability of readers." (43:250).

Both Fry and Spache apparently agreed with this hypothesis. The curve on Fry's Readability Graph was so drawn that in the lower levels sentence length played a major role in readability, although at the upper levels, word length accounted for most of the variability. (23:535). Similarly, Spache attached more importance to the sentence-length element than to the vocabulary element, in designing his readability formula for primary grades. (42:411).

Sentence structure has been measured in other ways than by length, and these are interrelated. Chall noted the use of the number or percentage of simple sentences as compared with complex sentences, and of clauses and prepositional phrases as indicators of sentence complexity. In general, easy materials are characterized by short, simple sentences with few prepositional phrases. (11:46).

Table I summarizes six primary-level formulas resulting from the foregoing studies and lists for each one the material used, the internal factors studied, the formula itself, and the criteria employed.

The Spache and Fry formulas will be expanded in more detail in Chapter 3.

Effectiveness of Formulas in Determining Levels

The idea that a piece of literature could be effectively "graded" by objective means has been the subject of continued controversy since its inception. Investigators have questioned the validity of underlying assumptions in the formula rationale, or the variables used in formulas, or the criteria employed, or the relative dependability of the ratings

TABLE I

SUMMARY OF READABILITY FORMULAS
APPLICABLE TO PRIMARY BOOKS

Author (1)	Material Used (2)	Internal Factors Studied (3)	Formula (4)	Criteria (5)
Johnson 1930 (30)	Textbooks, primer to Grade VIII; standardized tests, Thorndike word list	Percentage of polysyllables	Percentage of polysyllables. Also gives norms of polysyllables from primer to Grade VIII	Publishers' grade designations for texts; Thorndike frequency for words
Washburne-Morphett 1935 (49)	Basic Primary List of children's literature	(a) Number different words in 1000, (b) different uncommon words in 1000, (c) number of simple sentences in 75	Grade placement = (a) x .00255 + (b) x .0458 + 1.294 - (c) x .0307	Grade I and II, teacher judgment and children's reading; Grade III, Stanford Achievement scores
Dolch 1948 (18)	10 recently published basal reading series, Grades I to VI	Median sentence length, sentence length at 90th percentile, hard words (outside Dolch 1000)	Each of three factors converted to grade levels from tables; averaging the three yields grade levels of books	Inspection, comparing successive averages

TABLE I (continued)

Author (1)	Material Used (2)	Internal Factors Studied (3)	Formula (4)	Criteria (5)
Spache 1953 (41)	Basal reading books (129), social studies, health, and science (23) for Grades I to VIII	Percentage of hard words (outside Dale-Stone list), average sentence length	Grade level of books = .141 average sentence length per 100 word + .086 words outside Dale-Stone 769 + .839	"According to level of classroom use. . . pre-primers 1.2, primers 1.5, first readers 1.8. . ." --publishers' designations
Wheeler 1954 (51)	9 basic reading series	Unit length and percentage of polysyllabic words	Average unit length in random sample of ten to twenty pages x percentage of polysyllabic words x 10, graded by reference to table	Combined means of percentage of polysyllabic words and of unit length, for each level
Fry 1968 (21)	Miscellaneous readers for various grades	Word and sentence length	Plot of mean syllable and sentence length in three 100-word samples yields approximate difficulty level	Publishers' designations

obtained. Developers of formulas have been among the first to recognize the limitations of objective measurement and to caution against misuse of the formulas.

Limitations. 1. Each of the formulas is applicable only to material similar to the criterion on which it was based. Chall states that formulas in general were often criticized when the fault actually lay in their misapplication to a type of material for which they were not designed. (11:35). The limitation applied to range of difficulty as well as the type of material. Spache emphasized this restriction in writing of his own formula: "Although estimates of reading difficulty greater than 3.9 can be found by the formula, it is doubtful that these have any accuracy or even any real meaning." (41:150).

2. Readability formulas measure only a limited number of factors in reading difficulty. (Chall 11:56). This limitation is in part a built-in design, it having been determined that the inclusion of certain concrete variables other than vocabulary and sentence elements would only increase the complexity of the formulas without adding much to their predictive values. (Klare 31:12). However, the very objectivity of formulas has imposed another, more serious, limitation with regard to the lack of measurement of concepts, interest, content, style, format, and possibly other characteristics of printed matter relevant to readability; efforts of researchers to weigh and measure these attributes in statistical terms have met with little success. (Chall 11:156-157). The absence of

these kinds of factors in readability formulas has been the most vulnerable target for critics to date.

3. Formulas give only an approximate estimate of the difficulty of material," which should then be tempered with experienced judgment." (Chall 11:56).

4. Formulas are tools, not rules for writing. Gunning (27:39) advised would-be writers to use a formula "as a guide after you have written, but not as a pattern before you write." With regard to simplifying another author's work, Kress cautioned that shortening sentences or changing vocabulary to conform to a word list would not necessarily make material easier to read, nor could "the same concepts" be presented using different words. (33:98).

Formulas as practical tools. Supporters of the formula principle maintain that within the limitations noted, these instruments actually do reflect readability as well as researchers are currently able to measure it. Various agencies have relied on formulas as aids in the evaluation of reading materials or in the preparation of written materials. Publishers and editors have made wide use of formulas. Editors have adapted or simplifed materials for reluctant readers, using formula word lists as guides. Classroom teachers have found formulas helpful in determining the average reading ability needed for comprehension of certain books or for arranging books in rank order of difficulty. Spache has declared:

> There is no question that, applied intelligently, formulas have accomplished their aim . . . . The formulas are not intended to supplant any use of judgment, experience or knowledge of reading interests and habits but rather to complement these more ancient methods of estimating reading difficulty. (41:34).

Criticisms. Linguists in particular reject utterly the formula design, not for its structural weaknesses but for its presumption of adequacy in measuring any language elements. They take the position that a book is more than the sum of its words, phrases and clauses; that the elements of style, kind and quality of concepts expressed, and appeal to reader interest are properties unique to individual authorship, immeasurable by any statistical device. They deplore the application of readability measures to adapt or simplify literary works. As one critic expressed it, "The publishers are flooding the market with books written to order, watered down, doctored, squeezed, pounded and arranged to meet a prescribed grade level." Rheay pleaded for "freedom from this grade level strait jacket," particularly with respect to recreational reading. (39:479).

Ham granted the usefulness of formulas in making more or less accurate estimates of probable comprehension but regarded mere comprehension as a small part of the whole concept of readability. "The fact that a book can be read is no final measure of its readableness. . . . The reader determines readability, not the book." (28:167-170). Such factors as experience background, interest, and purpose of the reader were viewed by a number of critics as composing this final measure. (Manzo 35:962).

Reviews both favorable and unfavorable to formulas have noted that the procedure does not measure concepts. The concepts, nevertheless, are there. Kress pointed out that short, easy words, and even short sentences, may represent very high-level concepts. Conversely, simple concepts can be expressed in multisyllabic words and complex sentences (33:98). Dawkins assailed the validity of formula-derived ratings on this basis, charging, in effect, that such scores are not only inadequate representations of the difficulty of a book, but may be distortions as well. (17:515-521).

In further support of this thesis, Dawkins criticized a particular formula for assuming that words have only one assignable meaning and that words can be understood in isolation from contextual, syntactical, and grammatical elements. With respect to word lists, Dawkins noted the lack of consistency in designating certain hyphenated words, compound words, and idioms as "familiar" or "not familiar." In summation he declared, "The method of word counting used by the readability formulas shows no awareness of the nature of language."

Other critics were of the same persuasion. Blair deprecated the use of a word frequency index to indicate difficulty, in that abstractness of a term, or its morphological complexity (number and kind of prefixes, suffixes, etc.), were often overlooked. Further, he urged that the purpose and use of a word list be considered; a list based on the speech of children in one region of the United States must be used with greater care in other regions. (4:442-443).

Two discerning analyses by Bormuth relate to the structure of formulas. The first has to do with the form itself:

> The old readability formulas were presented in the form of what is called a multiple variable, linear equation. These equations have a characteristic that makes them unsuitable for use as readability prediction formulas. To use them we must assume that any correlation observed between two variables, say sentence length and word length, must always exist and that it must be of the magnitude observed in the original research. This is simply not true of the language features used in most formulas. . . . The result is that the old formulas yield misleading results whenever the correlation is anything other than the correlation the formulas assume. (10:844).

The second analysis concerns the measurement of the difficulty of a sentence:

> Two major objections can be raised to considering sentence length as the sole factor affecting grammatical complexity. First, it forces us to accept the dubious proposition that all sentences containing the same number of words possess the same degree of complexity . . . . Second, the number of words in a sentence does not measure a natural unit of language. We cannot simply add or chop off a few words to make a sentence more or less complex. Making a sentence more complex may or may not increase the number of words it contains; and increasing the number of words it contains may or may not increase the complexity of a sentence.
>
> The grammatical complexity of a sentence actually results from the grammatical structure of the sentence. (10:842).

Veatch called into evidence one of the primary linguistic principles that "sentences make words, rather than words make sentences," reasoning that since this is true, the basic principle of controlling the difficulty level of graded sequential readers by means of word count cannot be justified. (48:231-233; 243).

In the same vein, Schiller characterized the notion that words are the basic grammatical units and that they in turn combine to make

sentences, as "grammatically naive." Schiller reiterated the anomaly, subscribed to by linguists, that it is sentences, indeed groups of sentences, that give meaning to words. (40:17-32).

Summing up the position of linguists on grammatical structures, McCullough said in essence: Just as the sound of a letter depends upon the situation of the letter in the word, and the meaning of a word is altered by its relationships with other words, so the meaning of a sentence, the contribution of a sentence, depends upon its surroundings. Her epitome: "The elements of language are not islands. They create a fabric whose very open spaces are significant." (36:360).

Blair was more blunt in what amounted to a summation of the whole body of criticism against mechanical measurement of language:

> What don't these mechanical formulas measure? Among other things, contextual difficulty--what a word means in the context of surrounding words, what a sentence means in the context of surrounding sentences, etc. Abstractness of ideas isn't measured. Neither is density of ideas--how close together they are. Then there's reader interest in the subject. And style appeal--what you might call the literary quality of the writing. And how the material is organized. And whether it's interesting to look at--pictures, etc. And size of type, length of line, spacing, kind of ink and paper, lighting, etc. None of these is measured by the formulas. (4:442).

Manzo (35:962-965) called for a moratorium on readability research, for "there is probably nothing that can be done with formulas that cannot be done equally well without them."

RECENT ADVANCES IN READABILITY RESEARCH

Linguistics and psychology together, over the past decade, have contributed new insights into the nature of language and how language is processed in the mind of the reader. Research specialists, with this fresh knowledge, set out to develop more reliable methods for measuring the difficulty children have with reading materials. First they needed to identify and describe the linguistic features of materials that are really important in affecting comprehension. Computer techniques then enabled them to analyze their data in far more detail than was possible with traditional formulas.

Bormuth and Coleman's Analyses

Bormuth (7:4-54) and Coleman (15:166-178;16:316-324) have conducted extremely penetrating and comprehensive investigations into the "hundreds" (7:5) of variables which are likely to be involved in comprehension. As a result of these studies, Bormuth claimed in 1967 that whereas the older formulas, at best, could predict only 25 to 50 percent of the variation observed in the difficulties of materials, "today we have . . . several prototype formulas which are able to predict 85 to 95 percent of the variation." (10:840).

Bormuth and his team researched the more complex grammatical aspects of prose, and it might be assumed, therefore, that their studies relate only to sophisticated materials, beyond the primary level.

Yet even the simplest sentences, the shortest words, in fact, all but the most shallow types of writing, are seen in the light of these studies to be simple only on the surface. The findings outlined below are quite relevant to this investigation of readability.

> Word and sentence length are not independent linguistic variables; they are themselves dependent upon certain transformations which can be performed on the language. (7:10).

> A structure must be traced back to its underlying form before it can be understood. This may involve several transformations. The more structures, the more transformations; therefore the more likelihood of errors in comprehension. (7:19).

> All the vocabulary variables studied proved to be dependent variables. Word frequency is of minimal value in explaining why words vary in frequency and difficulty. Word length is a dependent variable; the number of letters or syllables in a word is dependent upon the number of affixes out of which the word is formed. The structures within the word probably cause words to vary in length as well as difficulty. (7:51).

> Word length may affect look-and-say learning in one of three ways: to the extent that words are processed visually, length in letters should have an effect; to the extent that they are processed vocally, length by syllables should affect readability; to the extent that they are processed mentally, length in morphemes should affect readability. (15:170).

> Structures such as prepositional phrases probably produce different effects upon the difficulty of sentences depending upon where those phrases are embedded in the sentence structure. (7:5).

> Abstractness also plays an important role in comprehension, but the variables have not yet been defined. (7:52).

> Structures which decrease the length of a clause were associated with the more difficult passages. (7:53).

Two conclusions from these factor analyses have particular significance for the investigator. Bormuth stated, "These studies make

it seem virtually certain that the previous practice of attributing grammatical difficulty to sentence length is not only illogical but contrary to fact." (7:53).

His second conclusion, however, presented no hopeful alternative: "The results . . . cast grave doubts on whether it is presently possible for a readability formula to exhibit economic practicality, face validity, and predictive accuracy." (7:54).

Cloze Method for Measuring Readability

Bormuth (9:1-11), Potter (38:1) and others have used and recommended cloze tests to measure a reader's understanding of a passage. The procedure involves the deletion of every fifth word of the selection, the blanks to be filled in with the exact words deleted. The cloze theory is explained by Potter:

> Just as there is an apparent tendency to "see" a not quite complete circle as a whole circle by "mentally closing the gap" and making the image conform to a familiar shape, a mutilated sentence is filled in by completing those words that make the finished language pattern conform to the intended or apparently intended meaning. (Merry Christmas, _____ New Year.)

One of the main advantages of this procedure, according to Bormuth, is that the measure of passage difficulty is not confused by injecting an extraneous reading task into the process; the instructor variable is thus reduced to a minimum. (10:841).

Bloomer, however, raised the question as to the efficacy of the cloze procedure for the less mature student, pointing out that apparently

he must be able to bring at minimum the ability to read well and clearly all the words presented to him in the cloze tests before he can make reasonable estimates of omitted words. If the child has difficulty mastering word recognition or phonics concepts, or has little knowledge of English language patterns and sequences, the cloze technique would not be of value. It was surmised that the minimum level for the use of cloze procedure would be a student with at least junior high reading skills in terms of word recognition, but who was having some difficulty with comprehension. (5:173-181).

There is some evidence, on the other hand, that the cloze technique is more versatile. Anderson (1:402-403), in an article acclaiming the advantages of this method, cited a study by Brual[1] who found it could discriminate among reading books "in the lower primary level."

Fry reported using an oral cloze technique with thirty second- and third-grade pupils as one means of ranking the difficulty of a passage. He found that this method was the most accurate and made the finest distinctions. Except that it requires an enormous amount of time to make the cloze tests, Fry stated, this procedure would be an excellent way to determine readability. "As a research tool, the method is

---

[1]T. G. Brual, "Readability" (unpublished thesis, University of Queensland, 1962).

excellent; but for practical purposes it is all but impossible to use." (23:534-536).

McLeod's Technique

In the same article Anderson discussed briefly another procedure departing from the traditional methods of estimating readability, a technique presented by McCloud[2]. McCloud conceived readability as "a threshold of difficulty." Tests were given to determine the students' reading level, and a criterion test was applied to the book. Then a graph was plotted of the proportion of children successfully reading a given book according to reading age. The threshold level expressed as a reading age gave the readability level of each book. Anderson praised the McCloud technique as a powerful measuring instrument, extremely reliable, with the particular virtue that it involves the reader. (1:402-403).

Further reference to the McCloud study could not be located by this investigator.

New Guidelines for Preparing Children's Books

Control of vocabulary, syntax and content. In another direction, research proceeded to investigate such matters as the control of vocabulary,

---

[2]J. McLeod, "The Estimation of Readability of Books of Low Difficulty," (The British Journal of Educational Psychology, 32:112-118, June, 1962).

syntax and content in reading materials prepared for children. In a review of recent research, Weinbraub reported these conclusions: (1) The words most frequently used in print, which comprise most lists of common words, should occur without planned control; (2) strict control of sentence structure in primary-grade materials is probably not necessary; children can and do use complex and compound sentences in their speech, and the sentence patterns they meet in reading should reflect similar structures; (3) the relevance, to primary-level children, of the content of their reading matter, should be given more attention. (50:195).

Coleman's studies. Coleman (15:166-178; 16:316-324) has recently conducted a series of experimental studies leading to the design of more readable books for beginners and suggesting guidelines for the construction of advanced materials. Viewing reading as "a complex hierarchy of sub-skills," Coleman points out that materials may be quite readable according to one sub-skill but unreadable according to another. (15:174). This principle holds for adult material as well as children's; it is commonly observed that the skills brought to the reading of light fiction, for example, are entirely different from those skills required for careful study of a piece of prose. Books devised for first-graders on the basis of phonics regularity will present little difficulty to children having facility in sounding out words, but may be unreadable to those relying on whole-word memorization. (16:316).

Coleman undertook to construct books for primary children that "optimally facilitate a combination of all sub-skills." (16:316). Using a list of regularly spelled words, supplemented by a very small number of function words misspelled in a transitional alphabet (thee, iz, wuz, etc.), and restricting the number of regularly spelled words to those that occur most frequently in English, he put the whole together with a series of little cartoon-like characters that suggest plot and sequence, to produce a twenty-word primer that was "extremely readable by any of the sub-skills." (16:317).

For readable writing above the first-grade level, Coleman stated his prescriptions succinctly: "Prefer grammatical transformations (1) that give short clauses and use active verbs; . . . (2) that do not use abstract nouns nominalized from verbs." (15:176).

Klare's Studies of the Role of Word Frequency in Readability

In 1968 Klare (31:12-22) reported findings from his own studies and from others, confirming the importance of word frequency as an element of the vocabulary factor in readability. He noted that (1) "familiarity" is determined almost uniquely by frequency, as shown in one study, to the extent of a .998 relationship; (2) there is a close correlation between word length and frequency of usage, words tending to be shortened as they are used more; (3) the percentage of monosyllabic words in a selection provides a fair index of reading difficulty;

(4) high frequency words have a low recognition threshold; (5) the most common words have the largest variety of meanings. In short,

> . . . not only do humans tend to use some words much more often than others, they recognize more frequent words more rapidly than less frequent, prefer them, and understand . . . them more readily. (31:12-22).

Measurement of Reading Ability

Unquestionably the success a reader has with a book is determined in large part by the reading ability he brings to it. It is relevant to the purpose of this investigation, therefore, to examine what recent research has to say about the usefulness of standardized tests in assessing reading ability.

Data and conclusions noted below are taken from the comprehensive review by Farr (20:38-71), of measurement problems and issues relating to reading.

Reading vocabulary tests. The wide array of procedures used to measure vocabulary casts doubt as to whether it is a specific sub-skill in reading. For example, in the tests surveyed, twenty-six different approaches were counted for measuring word meanings. Further, when time limits are imposed on the tests, what is being measured is some unknown combination of speed and vocabulary, not just speed or just vocabulary. As to the task of selecting the "best" synonym from a number of alternatives, linguists question the validity of defining any word apart from the semantic and syntactic context

clues of a passage.

Most vocabulary tests are quite similar to one another regardless of their intended grade-level use, yet studies show that younger children tend to use description-type definitions, while older children more often use synonym-type definitions.

Another point of controversy concerns the size of children's vocabulary. Tests and instructional materials have been based on word lists devised some years ago, and they vastly under-estimate the number of words known by children today at any grade level. (20:40).

There is a very high degree of over-lap in tests of various language skills. "Any attempts to diagnose reading vocabulary as distinct from reading comprehension or other areas should proceed on very cautious grounds." (20:43).

Reading comprehension tests. The division of comprehension into distinct sub-skill areas has not been based on any validity studies. Attempts to measure specific sub-skills have not been consistent. Sub-scores from a number of tests should not be used independently as a measure of the reader's skill but should be combined with others in an over-all measure. (20:52-71).

If research determines that reading comprehension is a composite of sub-skills, it probably will be discovered that the skills in turn are dependent upon a particular set of conditions, one of which is the reading difficulty of the selection in comparison to the reading skill of the examinee. (20:64).

In summary, Farr's analyses indicate that many tests fail to measure validly what they purport to measure, that no one seems to know whether sub-skills of reading can be measured, and that there is a lack of measures for assessing more complex reading behaviors. (20:71)

NEEDED RESEARCH

Some of the findings emerging from readability research are of immediate value in suggesting new approaches to instructional problems, new designs for instructional materials, new viewpoints of professional interest. Other findings may be significant in their implications, but not definitive; they require substantiation through cross-validation, through replication perhaps with other populations, or through other empirical evidence. Reference to the latter type of findings is included in the following list of aspects of readability that invite further study:

Continued study of the quantitative approach to evaluation, with emphasis on objectivity and efficiency in application. (Chall 11:158; Bormuth 9:4-9).

Study of those qualitative aspects of readability which so far have eluded objective measurement. (Chall 11:158; Bormuth 9:4-9).

Empirical determination of the instructional reading level. There is no reason to think that the traditional 75 percent criterion score will necessarily be the same for students of all ages, on passages

at all levels of difficulty, and on passages on all topics. (Bormuth 8:1-5).

Operational definition of reading comprehension in terms of specific reading tasks. (Farr 20:52).

Development of criterion tests for measuring comprehension. Usually standardized reading tests have been developed to compare one student's reading performance with that of another rather than with some specific goal. This constitutes one of the major shortcomings of all such tests. (Farr 20:64-65).

Development of differentiated procedures for measuring reading vocabulary at different age levels. (Farr 20:39).

Development of tests based on sound empirical evidence concerning the components of reading ability. (Farr 20:48).

Investigation of the widespread notion that vocabulary and comprehension are separate, measurable sub-skills of reading. (Farr 20:43, 65).

Determination of relative levels of difficulty of different syntactic structures. (Weintraub 50:195).

Updating of the estimates of size of students' vocabularies. (Farr 20:44).

Refinement of the vocabulary factor in formulas, to answer such questions as the following:

Should graphological length or phonological length be used as the measure of word difficulty?

What is the significance of repetition of hard words?

How should inflectional forms of a word be treated? Different meanings of a word? Names of persons and places? Slang and dialect? Numbers? (Chall 11:159-160).

SUMMARY

Critics assailed traditional formulas as crude, outmoded, ineffective, or useless, some declaring that experienced judgment is a more reliable method of determining the readability of a book, inasmuch as it takes into account all pertinent factors. Proponents countered that formulas were designed as an aid, not a comprehensive measure, in evaluating readability, and that within the limitations of objectivity, the value of formulas to writers, editors, publishers and educators continues to be demonstrated. It is pointed out with some justice that research so far has been unable to produce a better tool for the purpose.

It is claimed that researchers have designed new formulas that can account for as much as 95 percent of the difficulty variation in materials. It is admitted, however, that the expense of linguistic analyses and mathematical computations involved makes them impractical for common use.

Investigations relating to the style of children's books have led to an improved design for primary materials and have pointed the way to improvement at the intermediate level.

Recent studies on what can be measured in reading raise a number of questions about the nature of comprehension and the validity

of present standardized tests of reading. Statistical evidence tends to discredit the accuracy of test scores on "vocabulary" or "comprehension."

The current thrust of research, as evinced in the studies reviewed, is into the area of linguistics, as investigators probe into the nature of language and its patterns of expression to discover what actually makes a book readable.

Much of the on-going research and many of the projections for future study will be of only academic interest to the classroom teacher, until computer techniques and facilities are made available to the schools. What the primary teacher needs now is a simple, time-saving tool to which he can turn with confidence when he needs a measure of reading difficulty. Basically, the teacher wants to know these things about a book for a given child: Are the words too hard? Are the sentences too long or complicated?

One of the techniques described in Chapter Three can give him that first approximation of difficulty in fifteen minutes.

CHAPTER 3

DESCRIPTION OF THIS STUDY

The previous chapter covered reviews of research studies on readability, especially those aspects which pertained to primary-grade materials. Chapter Three will be devoted to a description of two particular formulas, the Spache Readability Formula (41:141-148) and the Fry Readability Graph (21:513-516); their application to each of the listed primary books; a tabulation of comparative data on readability levels from several sources; and an analysis of the data.

Preliminary information concerning the simplicity, ease of application and time-saving feature of the new Fry formula for estimating difficulty of primary books, commended it for further investigation. It was decided that a valid test of the reliability of the Fry instrument would be a comparison with the Spache formula, which is used by many publishers, editors and educators for estimating the readability of books at this level. The purpose of this study, originally aimed at determining the relative merits of objective and subjective evaluations, was consequently expanded to include an investigation into the relative merits of the two formulas.

SPACHE READABILITY FORMULA

Spache and his co-workers agreed with earlier research findings that the best predictors of difficulty are a measure of

vocabulary and a measure of sentence length. Their data differed, however, in indicating that sentence length is slightly more closely related to reading difficulty for primary reading material than is the percentage of hard words. The Spache formula, consequently, is weighted more heavily on this factor. (42:44).

Electing to use a word list as a basis for measuring vocabulary difficulty, Spache chose the Dale list of 769 words,[1] judging all words outside the list as hard. This list was composed of words found both in the International Kindergarten Union list[2] and the first thousand words of Thorndike's Teacher's Word Book (46). The first formula, in 1953, relied on the original Dale list, but the word list now used is a revision based on the suggestions of Stone in 1956. (Spache 41:145). Stone pointed out that a considerable number of words on the Dale list were really difficult, rather than easy, words for primary reading, and that the list omitted many words commonly used in pre-primers, primers, and first readers. Stone proposed the deletion of 173 of the words he considered unsuitable and the insertion of 173 words he judged to be

---

[1]Edgar Dale, The Dale "Easy Word List" of 769 Words. (Columbus: Bureau of Educational Research, Ohio State University, 1948).

[2]International Kindergarten Union, Child Study Committee, A Study of the Vocabulary of Children before Entering the First Grade. (Washington: distributed by the International Kindergarten Union, 1928).

easier on the basis of grade rating, familiarity, and length of word. (44:36-41). Spache adopted this revised list, commenting, however, that differences in the estimates of reading difficulty using the new list, averaged less than two months and in no case were greater than four months. (41:145).

Schoolbooks in common use in this country were analyzed in terms of the vocabulary and sentence factors. Spache drew several 100-word samples from each of 152 basal readers, social science, health and science books. "Each book was assigned the grade-level designation of the publisher: pre-primers 1.2, primers 1.5, first readers 1.8 and 1.9, second readers 2.1 and 2.7, and third readers 3.3 and 3.7" (41:141).

The regression equation constructed from these data was a simple one: Grade level of book = .141 average sentence length per 100 words + .086 words outside the Dale-Stone list + a constant of .839.

On the following pages are detailed the directions for using the Spache formula and rules and suggestions for applying the formula; in addition are reproductions of the form of worksheet recommended and of the Dale-Stone word list in its entirety.

Procedure in Computing Grade Placement (41:144)

(1) Prepare a Worksheet like that given on page 43.

(2) Count off approximately 100 words in the early part of the book. Begin at the beginning of a sentence and end the count with the last word of the sentence containing the hundredth word.

(3) Write the number of words in the Worksheet on line 1.

(4) Count the number of sentences in the sample. Write the number of sentences in the Worksheet on line 2.

(5) Check the separate words in the sample against the Stone Revised Word List. Make a count of the number of words not found in this list.

(6) Write the number of hard words in the Worksheet on line 3.

(7) Divide the number of words in the sample by the number of sentences to find the average sentence length (line 4).

(8) Divide the number of hard words by the number of words in the sample to find the percent of hard words. Drop the decimal point. (line 5).

(9) Multiply the average sentence length (line 4) by .141. Write the product on line 6.

(10) Multiply percent of hard words (line 5) by .086. Write the product on line 7.

(11) Add the figures on lines 6, 7 and the constant, .839.

(12) The sum is an estimate of the grade level of difficulty of the selection.

(13) Repeat steps 1-11, with samples from the middle and rear of the book. Use at least 5-10 samples, depending upon the length of the book.

(14) Determine the average grade placement of the book by adding the estimates and dividing by the number of samples. This is the final estimate of the grade level of difficulty of the entire book. Drop the last figure or round it off, as 2.367 = 2.4.

Rules and Suggestions for Applying the Formula (41:144-145)

(1) Count all letters and numbers in figures as familiar.

(2) Proper nouns, or names of persons, places are counted as familiar.

(3) Count regular verb forms as familiar. This includes ing, es, ed, and changes involving doubling of the final consonant, dropping the final e, changing y to i.

(4) Count regular plurals and possessive endings of nouns as familiar. Plurals in s, es, ies are familiar; those as in ox-oxen, goose-geese, are unfamiliar unless on the list.

(5) Count adjectival or adverbial endings, as ily, er, est, ly as unfamiliar unless on the list.

(6) Count a word as unfamiliar only once even though it appears again or with variable endings later in the sample.

(7) A group of words, consisting of the repetition of a single word or exclamation . . . is counted as a single sentence regardless of punctuation.

(8) Count hyphenated words as unfamiliar unless both parts appear in the word list.

(9) Count contractions, as didn't, unfamiliar unless on the list.

(10) Count hyphenated words, compound words and numbers in figures as one word.

(11) Analyze each sample independently, i.e., words counted as unfamiliar in any sample are again unfamiliar in subsequent samples.

(12) Count single or two-word sentences as such in determining average sentence length, as in directions and some pre-primers.

(13) Avoid sampling material that is not typical of continuous matter, e.g., avoid dialogue, headings, titles.

(14) Avoid sampling consistently at the beginning or end of chapters, since the Clymer study . . . indicates these are not typical.

Worksheet for Application of the Spache Readability Formula for Grades I-III

Article or Book_____ Date_____

Author_____ Publisher_____

	Page____ From____ To____	Page____ From____ To____	Page____ From____ To____	Page____ From____ To____
1. Number words				
2. Number sentences				
3. Number words not on Stone Revised Word List				
4. Ave. Sentence Length (Divide 1 by 2)				
5. Per cent hard words (Divide 3 by 1, multiply by 100)				
6. Multiply (4) by .141				
7. Multiply (5) by .086				
8. Constant	.839	.839	.839	.839
9. Estimated grade placement (Add, 6, 7, and 8)				

Average grade placement of_____ samples_____

Analyzed by_____

Date_____

Dale-Stone Word List

CLARENCE R. STONE'S REVISION OF THE DALE LIST OF 769 EASY WORDS

a
about
across
afraid
after
afternoon
again
air
*airplane
all
almost
alone
along
already
also
always
am
an
and
animal
another
answer
any
*anyone
anything
apple
are
arm
around
*arrow
as
ask
*asleep
at
*ate
*automobile
away

*baa
baby
back
bad
bag
*bake
*baker
ball
*balloon
band
*bang
*bark
*barn
*barnyard
basket
*bath
be
bear
beautiful
*became
because
bed
*bedroom
bee
been
before
began
begin
behind
being
believe
bell
belong
beside
best
better
between
big
*bigger
bill
bird
*birthday
bit
black
*blew
blow
blue
board
boat
book
both
bottom
bow
*bowl
*bow-wow
box
boy
branch
bread
break
breakfast
bright
bring
brother
brought
brown
*bug
build
building
*bump
*bunny
*bus
busy
but
butter
buy
*buzz
by

*cabbage
*cage
cake
*calf
call
came
can
*candy
cap
car
care
careful
carry
*cat
catch
*caught
cent
chair
*chick
*chicken
child
children
Christmas
*circus
city
*clap
clean
*climb
close
clothes
*clown
*cluck
coat
*cock-a-doodle-doo
cold
color
come
coming
cook
*cooky
corn
corner
could
count
country
cover
cow
cried
cross
*crumb
cry
cup
cut

dance
dark
day
dear
deep
*deer
did
*dig
dinner
*dish
do
does
dog
*doll
done
don't
door
down
draw
dress
drink
drive
drop
dry
*duck

each
ear
early
east
eat
egg
*elephant
else
end
*engine
enough
even
ever
every
everything
eye

face
fall
family
far
farm
farmer
fast
fat
father
*feather
feed
feel
feet
fell
felt
fence
few
field
fill
find
fine
finish
fire
first
fish
fit
five
*flag
*flew
floor
flower
fly
follow
food
foot
for
found
four
*fox
fresh
friend
*frog
from
front
fruit
full
*fun
*funny

game
garden
gate
gave
get
girl
give
glad
go
*goat
God
going
gold
gone
good
*good-by
got
*grandfather
*grandmother
grass
gray
great
green
grew
ground
grow
guess

* Added to Dale's list by Clarence R. Stone.

Dale-Stone Word List (continued)

had
hair
hall
hand
*happen
happy
hard
has
hat
have
*hay
he
head
hear
heard
heavy
*held
*hello
help
*hen
her
here
herself
*hid
hide
high
hill
him
himself
his
*hit
hold
hole
home
*honey
*hop
*horn
horse
hot
house
how
*hungry
hunt
hurry
hurt

I
ice
if
*I'll
in
Indian
*inside
into
is
it
its

*jar
*joke
jump
just

keep
kept
kill
kind
*kitchen
*kitten
knew
*knock
know

lady
laid
*lamb
land
large
last
late
laugh
lay
learn
*leaves
left
leg
let
*let's
letter
lie
light
like
line
lion
listen
little
live
*log
long
look
lost
lot
loud
love
*lunch

made
mail
make
man
many
march
matter
may
me
meat
meet
men
*meow
met
*mew
*mice
might
mile
milk
*milkman
mill
minute
miss
*Miss
money
*monkey
*moo
more
morning
most
mother
*mouse
mouth
move
Mr.
Mrs.
much
*mud
music
must
my

*nail
name
near
neck
need
nest
never
new
next
nice
night
no
noise
north
nose
not
note
nothing
now
*nut

of
off
often
*oh
old
on
once
one
only
open
or
*orange
other
our
out
outside
over
own

paint
*pan
paper
*park
part
party
*pat
*paw
pay
*peanut
*peep
*pennies
people
*pet
pick
*picnic
picture
*pie
piece
*pig
*pink
place
*plan
plant
play
please
*pocket
point
*policeman
*pond
*pony
poor
*pop
post
present
press
pretty
*puff
pull
*puppy
*push
put

quick
quiet
quite

*rabbit
race
rain
*rake
ran
read
ready
real
red
rest
ride
right
ring
river
road
*roar
*robin
rock
*rode
roll
roof
room
*rooster
*root
*rope
round
row
*rub
run

said
same
sand
*sang
sat
save
saw
say
*says
school
sea
seat
see
seed
seem
seen
sell
send
sent
set
seven
shake
shall
she
sheep
*shell
shine
shoe
shop
short
should
show
shut
sick
side

* Added to Dale's list by Clarence R. Stone.

Dale-Stone Word List (continued)

sign
sing
sister
sit
six
*skate
skin
*skip
sky
*sled
sleep
*sleepy
*slide
slow
small
*smell
smile
smoke
*sniff
snow
so
soft
sold
some
something
sometime
song
soon
sound
*soup
*splash
spot
spring
*squirrel
stand
star
start
station
stay
step
stick
still
stone
stood
stop
store
story
straight
street
*string
strong
such
suit
summer
sun
*sunshine
*supper
sure
surprise
*swam
sweet
*swim
*swing

table
tail
take
talk
tall
*tap
teach
teacher
*teeth
tell
ten
*tent
than
thank
that
the
their
them
then
there
these
they
thin
thing
think
this
those
though
thought
three
*threw
throw
*ticket
tie
*tiger
time
*tired
to
today
*toe
together
told
tomorrow
too
took
top
town
*toy
train
tree
*trick
*tried
*truck
*trunk
try
*turkey
turn
*turtle
two

*umbrella
uncle
under
until
up
upon
us
use

*vegetable
very
visit
*voice

*wagon
wait
*wake
walk
want
war
warm
was
wash
watch
water
wave
way
we
wear
*wee
*weed
week
well
went
were
west
*wet
what
wheat
wheel
when
where
which
while
white
who
why
wide
wild
will
win
wind
window
wing
winter
wish
with
without
*woke
woman
wonder
wood
*wolf
word
work
world
*worm
would
write

yard
year
yellow
yes
you
your

*zoo

"EASY" WORDS OMITTED FROM DALE'S LIST

above
act
against
ago
American

bank
beat
bless
blind
blood
body
bone
born
brave
bridge
broken
built
burn

captain
case
cause
center
chance
change
chief
choose
church
circle
class
clear
clock
cloth
cloud
coal
company
cool
cost
course
crowd
crown

dead
die
different
doctor
double
dream
dust

earth
easy
edge
eight
either
England
English
evening
except
expect

fair
fancy
fellow
fight
finger
fix
forget
forth

gift
glass
golden
grain

half
hang
heart

Study of the Spache Sampling Technique

Clymer (14:245-250) undertook to determine for the Spache formula the number of samples and the technique of sampling which would give the best estimate of readability.

Spache originally recommended that three samples of approximately 100 words each should be used. These samples were to be taken one each from near the beginning, the middle, and the end of the book. (42:412). In a later presentation, Spache directed: "Use at least five to ten samples, depending on the length of the book." (41:144). Clymer noted that no data had been published to support these recommendations, or to give the reason for the change.

To determine the "true" readability of a number of second- and third-grade texts, Clymer applied the formula to the entire content of each, in successive 100-word samples, then found the average of all the scores. This sampling procedure, Clymer claimed, demonstrates more clearly the true readability value than does the treatment of a book as one single sample. This is because a word is counted "hard" only once in a sample. Contrary, then, to what some previous research has shown, the Spache sampling technique does not over-estimate "true" readability but is the only procedure to be followed when the formula has been developed through that process. (14:248-249).

The system of "stratified" sampling recommended by Spache, Clymer found to be a reasonable practice, on the assumption that primary

texts probably increase in difficulty from beginning to end.

With regard to the number of samples to be drawn, Clymer stated:

> This will depend upon the precision with which the book must be judged. For a quick check of a book to be taken from the library for use with a child of known reading ability, three samples are surely enough. . . . If the results of this experiment are typical, by using three samples a book will be rated within a month of its "true" readability about thirteen out of eighteen times. (14:249).

In applying the Spache formula to books for the present study, three samples were used unless the length of the book warranted more, or unless the effect of using different samplings was being experimented with.

Criteria for the Spache Formula

An observation by Chall on this subject is pertinent:

> Spache's 1953 formula for grading primary materials is based . . . on standards derived from current textbooks. Hence, the grade-level estimates derived from the formula can be used . . . only for estimating difficulty in relation to current primary materials. (11:121).

Word List Component of the Spache Formula

Along with the implication that the Spache criteria might become out of date, there should be some comment about the age of the word list used. Although it was revised in 1956, over 76 percent of the words remained unchanged, deriving, through the Dale list of 1948, from the International Kindergarten Union list of 1928 (see page 39),

and the Thorndike list of the nineteen thirties. This lack of up-dating gives rise to some speculation as to how valid the list actually is for application to present-day materials.

Accuracy of the Formula

According to Spache, the accuracy of this formula compares favorably with that obtained from other formulas. He stated that "in half the predictions the error in estimating grade level will be less than 3. 3 months, and in the remaining predictions the error will probably be greater than three months." (41:143).

Fry commented, however, that scores obtained from the Spache formula give an appearance of accuracy that could be misleading. He interpreted Spache's report of "a probable error of estimate in predicting grade level, of 3. 3 months," in this way:

> This means that half the time the true score of a book lies within a 6. 6 months' band centered around the score obtained by working the formula, and that half the time the real grade level lies outside the 6. 6 months' band. Therefore, accurately judging the grade level of a passage to within one tenth of a grade level by using the Spache formula is not possible. (23:537).

Perhaps this appearance of accuracy is one of the "pitfalls for the unwary" that formula users are cautioned to avoid. In any event, this apparent power to discriminate so precisely between tenths of a point of difficulty has undoubtedly added much to the popularity of the Spache formula. It was already unique in its field; there was none to challenge it until the Fry Readability Graph was published.

THE FRY READABILITY GRAPH

In 1968 "a readability formula that saves time," by Dr. Edward Fry of Rutgers University, was announced by editors of the Journal of Reading, offering a "handy graph for pinpointing readability levels with reasonable accuracy and uncommon simplicity." (21:513). To this recommendation it could have been added that the Graph is unique in its range, covering all grade levels from pre-primer through college.

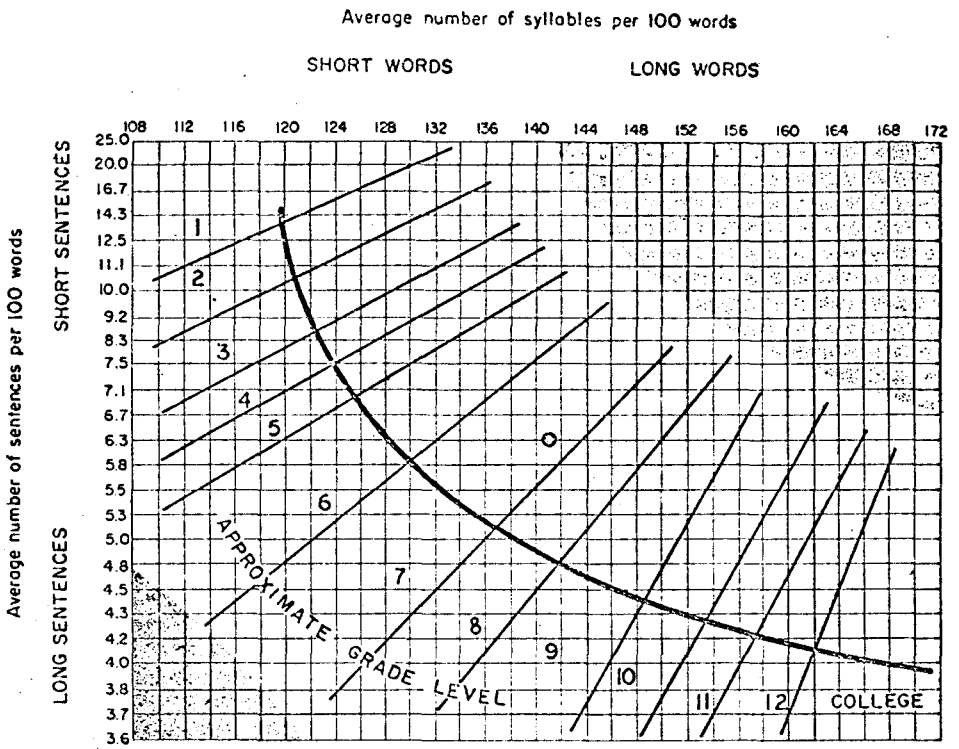
"The rationale for 'one more' readability formula is its simplicity," stated Fry, noting that it is often lack of simplicity that has kept other readability formulas from more widespread use. His formula considers two factors: word length expressed in syllables, and sentence length. There is no word list to be referred to, there is a minimum of computation involved, and anyone who can count to one hundred can work out the formula. A copy of the Readability Graph, reproduced on the next page, is the only accouterment needed.

Directions for Use

Both the directions and the process are simple: (21:514).

(1) Select three one-hundred-word passages from near the beginning, middle and end of the book. Skip all proper nouns.

(2) Count the total number of sentences in each hundred-word passage (estimating to nearest tenth of a sentence). Average these three numbers.



DIRECTIONS: Randomly select 3 one hundred word passages from a book or an article. Plot average number of syllables and average number of sentences per 100 words on graph to determine the grade level of the material. Choose more passages per book if great variability is observed and conclude that the book has uneven readability. Few books will fall in gray area but when they do grade level scores are invalid.

EXAMPLE:

	SYLLABLES	SENTENCES
1st Hundred Words	124	6.6
2nd Hundred Words	141	5.5
3rd Hundred Words	158	6.8
AVERAGE	141	6.3

READABILITY 7th GRADE (see dot plotted on graph)

Figure: Fry Readability Graph

(3) Count the total number of syllables in each hundred-word sample . . . . It is convenient to count every syllable over one in each word and add 100. Average the total number of syllables for the three samples.

(4) Plot on the graph the average number of sentences per hundred words and the average number of syllables per hundred words. Most plot points fall near the heavy curved line. Perpendicular lines mark off approximate grade level areas.

If great variability is encountered either in sentence length or in the syllable count for the three selections, then randomly select several more passages and average them in before plotting.

As to accuracy of the score, Fry stated that it is "probably within a grade level" (21:514).

Criteria Selected

In common with the formula designers who preceded him. Fry accepted publishers' designations of grade level as a point of reference. He plotted "lots of books which publishers said were third-grade readers, fifth-grade readers, etc.,. . . looked for clusters and 'smoothed the curve.' " (21:515).

Validity of Internal Factors Employed

Early studies had indicated the validity of using the word length approach as a measure of vocabulary, as opposed to the word list approach. Johnson (30:283-298) in 1930 and Wheeler (51:397-399) in 1954 had used a polysyllabic factor in their formulas. Then in 1959 Stolerow and Newman in their factor analysis of forty-four variables had noted a high correlation between reading ease and monosyllables. (43:243-251). Fry's selection of word length as the vocabulary factor

in his formula, thus was firmly supported by research findings.

Similarly, the importance of a sentence factor in predicting reading difficulty had long been recognized, but as a secondary element. The analyses of Stolerow and Newman indicated, however, that in the lower levels of difficulty the relative predictive values of the word and sentence factors apparently shifted, to give the sentence factor a slight dominance. (43:250).

Spache, too, had found earlier that sentence length is a better predictor of difficulty in primary materials than is a measure of vocabulary load. He wrote:

> Apparently, editors of material for primary reading materials exercise more control over sentence length than over the introduction of hard words. Above the primary grades, sentence length is less controlled and perhaps less significant in reading difficulty, since the child has acquired a modicum of reading skill. (41:411).

Perhaps the variation is better accounted for by the "curvilinear relationship," noted by Bormuth, which apparently exists between semantic and syntactic structures. His illustration clarifies the term somewhat:

> For example, adding another syllable to a one-syllable word increases its difficulty far more than adding another syllable to a seven-syllable word. The same is true of many other features. Interestingly, it was the variables most frequently used in the old formulas that showed the greatest amount of curvature. (10:844).

The shape of the curve on the Readability Graph may possibly be due to this variation, in Fry's opinion. (23:535). Inspection of the graph shows that sentence length plays a major role at the lower reading

levels, while word length accounts for most of the variation at the upper levels.

Validation of the Graph at Primary Levels

In one study to test the validity of his formula, Fry compared the mean readability grade-level scores on the Spache and Fry formulas and cloze error and oral reading scores of thirty students on six primary-level books. He reported that the cloze method seemed to be the most accurate and made the finest distinctions, although this technique, and the oral reading method as well, are too time-consuming for practical purposes. Both the Spache formula and the Readability Graph showed very high correlations with the cloze method and satisfactory correlations with each other. (23:538).

Usefulness of the Graph at Primary Levels

Although it is new and relatively untried, the Readability Graph has attracted notice in several professional publications. Gaver, in an article appearing in School Libraries in 1969, hailed the new technique as "a real boon to school librarians," especially for its ease of application--"less than five minutes per book,"--and its high reliability. By using the Fry system, Gaver stated, it is possible to determine in minutes the difference between the interest level and the reading level of any book, and to determine the levels of books for independent reading, especially those with potential for the primary grades. (25:23-25).

Pauk (37:207-210), in an article comparing the Fry formula with others at upper levels, also noted the simplicity and time-saving advantages of the new formula, commenting that it takes but fifteen minutes to apply.

FRAMEWORK OF THE BOOKLIST COMPARISONS

The booklist which follows has undergone certain changes in size and content due to the non-availability of some of the titles at the time of this evaluation. All titles listed, however, are presumed from at least one source to be suitable for independent reading at third-grade level.

The Spache and Fry grade-level designations were computed by this investigator according to directions prescribed by the formulators. Following Clymer's recommendation (see page 47), only three samples were used for each scoring by the Spache formula, except in instances where a wide variation in difficulty made further sampling advisable. It is not claimed that the same sample lengths were used in both the Spache and Fry evaluations, since the Fry formula requires exactly one hundred words and the Spache formula may extend a number of words farther, to the end of the sentence. Insofar as the samples are comprised of the same hundred words, it can be said that the same material was used for both applications.

Criteria Used by Evaluators

Publishers. Criteria vary, as indicated by letters from publishers. (See Appendix, pages 85-92). Brochures from some publishers state that their primary books are carefully controlled in vocabulary through the use of various word lists and/or the Spache formula. A number of the earlier-published books show a break-down of words employed at each level of difficulty. Many publishers apparently rate suitability by judgment only.

Children's Catalog.

> Demonstrated usefulness of the books is vouched for by a representative group of experienced librarians and specialists in children's literature. (13:5).
>
> The figures (4-6) indicate that this book is useful for children in the 4th to 6th grades. It is difficult to make generalizations as to the reading ability of children and for this reason the grading given is rather flexible . . . . (13:9).

Harris.

> The grade placement for each book is this writer's judgment of the minimum level of reading ability necessary if a boy or girl is to enjoy reading the book independently or with very little help . . . . To be included [in the list], each book had to pass the test to having interest appeal to children at least two grades above the difficulty rating. (29:594).

Kottmeyer. In his "Bibliography for Retarded Readers," books were listed by grade level of difficulty, "included because they have been used with remedial pupils for some time." (32:189). It is

assumed that Kottmeyer has followed his own prescription for choosing material to be read independently: one grade level lower than the pupil's silent reading test score. (32:185).

Sullivan.

> After more than twenty years of supplying books for children who need more mature content than their recognition vocabulary will allow them to comprehend, we have concluded that no readability formula is suitable. Use and judgment decide where a book fits as to grade in vocabulary, sentence structure, and degree of abstractness. Such must be determined in relation to each book . . . . Books . . . have been indicated of grade one or two level on this list only if the child can read them himself. (45:3-5).

Eakin. At the easy book level, all books graded third grade and below were tested on the Spache formula for primary materials. (19:xii).

TABLE II

DATA ON READABILITY LEVELS OF SELECTED BOOKS

Author and Title of Book	Publisher[1]	Professional Judgment: Children's Catalog	Harris	Kottmeyer	Sullivan	Eakin	Formula: Spache	Spache[2]	Fry[3]
1. Anderson: BILLY AND BLAZE	T	k-2	3	3			3	3.8 *(3.1-4.7)	5
2. Beim, J.: ANDY AND THE SCHOOL BUS	U	k-2			1	2		2.2	2[L]
3. Beim, J.: BOY ON LINCOLN'S LAP	U				1			3.2 (2.8-3.6)	3
4. Beim, J.: SMALLEST BOY IN CLASS	U	1-3	3			3		3.0 (2.6-3.3)	3
5. Beim, J.: THIN ICE	U	1-3	1		1			2.7	1[H]
6. Beim, J.: TIM AND THE TOOL CHEST	U	1-3		2	1	2		3.0 (2.7-3.5)	2[L]
7. Beim, L. &J.: TWO IS A TEAM	M	1-3		2		3		2.9 (2.7-3.5)	3
8. Benchley: RED FOX AND HIS CANOE	N 2							2.5 (1.6-3.3)	1

[1]Code letters refer to publishers listed in Appendix, page 93

[2, 3]Levels determined in the present study

*Indicates range

TABLE II (continued)

DATA ON READABILITY LEVELS OF SELECTED BOOKS

Author and Title of Book	Publisher[1]	Children's Catalog	Professional Judgment: Harris	Kottmeyer	Sullivan	Eakin	Formula: Spache	Spache[2]	Fry[3]
9. Berres: THE SEA HUNT	G 1.8		2				2	2.6 (2.2-3.1)	1
10. Brown: SNEAKERS	L		1		1	3		4.1 (3.8-4.5)	6^{L}
11. Brown: WHEEL ON THE CHIMNEY	Q	k-3	1		1	3		3.7	5^{L}
12. Bulla: SURPRISE FOR A COWBOY	L			4	1			2.4	1
13. Burton: MIKE MULLIGAN	P 3	1-3	3	3	2	RA	2-3	4.4 (3.6-5.4)	6^{L}
14. Cerf: ANIMAL RIDDLES	W 2	k-2					2	2.3	1
15. Cerf: RIDDLES	W 2	2-5	1					2.2	1^{L}
16. Cerf: MORE RIDDLES	W 2						2	2.5	2^{M}
17. Chandler: COWBOY SAM	B Pr-1		1	Pr-1	1		1	2.1	Pr
18. Chandler: ... & FREDDY	B 1^{H}		2	Pr-1			2	1.8 (1.1-2.0)	1^{L}
19. Darby: LEONARD VISITS OCEAN FLOOR	G Pr							1.8 (1.6-2.0)	1
20. DeRegniers: LITTLE HOUSE OF YOUR OWN	M	k-2			1			3.0	4

TABLE II (continued)

DATA ON READABILITY LEVELS OF SELECTED BOOKS

Author and Title of Book	Professional Judgment							Formula	
	Publisher[1]	Children's Catalog	Harris	Kottmeyer	Sullivan	Eakin	Spache	Spache[2]	Fry[3]
21. Dolch: ONCE THERE WAS A CAT	I 1[2]			1				2.2 (1.7-2.5)	1
22. Dupre: TOO MANY DOGS	H 2-4							2-4 (1.9-2.8)	1[L]
23. Flack: ANGUS AND THE DUCKS	F	k-2	1	Pr-1	1	RA k-2		3.6 (2.9-4.3)	3[H]
24. Flack: BOATS ON THE RIVER	Y	k-3	2					4.0 (3.4-5.0)	7[L]
25. Flora: MY FRIEND CHARLIE	M	1-4						3.4	3[L]
26. Freeman: YOU WILL GO TO THE MOON	W 2	1-3					1.8	1.9 (1.2-2.2)	PPr
27. Geisel: CAT IN THE HAT COMES BACK	W 1[H]		1			1	2	2.1	PPr
28. Geisel: GREEN EGGS AND HAM	W		1					1.8	PPr
29. Georgiady: GERTIE THE DUCK	H 1						2	2.6 (2.1-3.6)	1

TABLE II (continued)

DATA ON READABILITY LEVELS OF SELECTED BOOKS

Author and Title of Book	Publisher[1]	Professional Judgment: Children's Catalog	Professional Judgment: Harris	Professional Judgment: Kottmeyer	Professional Judgment: Sullivan	Professional Judgment: Eakin	Formula: Spache	Formula: Spache[2]	Formula: Fry[3]
30. Goldin: SPIDER SILK	D 2	1-3						3.3 (3.0-3.7)	2^{M}
31. Guilfoyle: NOBODY LISTENS TO ANDREW	H 1	1-2	1				2.1	2.3 (1.9-2.6)	1
32. Hader: THE BIG SNOW	L	k-2			1	RA k-1		3.8	4
33. Hoff: THE FOUR FRIENDS	H 2							2.2	1^{H}
34. Hoff: CHESTER	N 2					2	1	2.1	Pr
35. Hoff: GRIZZWOLD	N 2							2.7 (2.0-3.6)	1
36. Holt: LANCE AND HIS FIRST HORSE	L				1			3.6	4^{L}
37. Hurd: COME AND HAVE FUN	N	k-3						1.9	PPr
38. Hurley: DAN FRONTIER GOES HUNTING	C Pr		1				2	1.8	Pr
39. Ipcar: WORLD FULL OF HORSES	F		2		1			3.1	4^{H}

TABLE II (continued)

DATA ON READABILITY LEVELS OF SELECTED BOOKS

Author and Title of Book	Publisher[1]	Children's Catalog	Professional Judgment: Harris	Kottmeyer	Sullivan	Eakin	Formula: Spache	Spache[2]	Fry[3]
40. Kay: SNOW BIRTHDAY	O	k-3			2	RA 3		3.3 (3.0-3.7)	3^{M}
41. Kessler: DUCK ON THE TRUCK	K 2^{M}	k-3						2.1	Pr
42. King: MABEL THE WHALE	H 2	1-3					2	2.5	1
43. Lenski: LITTLE AIRPLANE	V	k-2	1				3	3.4	3^{H}
44. Lenski: PAPA SMALL	V	k-2	1		1			2.3	1
45. Lent: JOHN TABOR'S RIDE	R							4.1 (3.5-4.9)	2
46. Lexau: BENJIE	E	k-3						3.3 (2.8-4.1)	2^{H}
47. MacDonald: RED LIGHT, GREEN LIGHT	F	k-2	1		1	2		2.6 (2.1-3.1)	Pr
48. McCall: BUTTONS TAKE A BOAT RIDE	C Pr		1	2		1	1	1.8	PPr
49. McCloskey: BLUEBERRIES FOR SAL	Y	k-2			2			3.4 (3.0-3.8)	5^{H}

TABEL II (continued)

DATA ON READABILITY LEVELS OF SELECTED BOOKS

Author and Title of Book	Publisher[1]	Children's Catalog	Professional Judgment: Harris	Professional Judgment: Kottmeyer	Professional Judgment: Sullivan	Professional Judgment: Eakin	Formula: Spache	Formula: Spache[2]	Formula: Fry[3]
50. McCloskey: MAKE WAY FOR DUCKLINGS	Y	k-2				2		4.0 (3.5-4.5)	4
51. Minarik: FATHER BEAR	N	k-2						2.2	2^L
52. Minarik: LITTLE BEAR VISIT	N 2	k-2						2.2	1^H
53. Moore: TOO MANY BOZOS	J 2							2.3	1^L
54. Newell: HURRY UP SLOW POKE	K 1^H							2.1	1^L
55. Perkins: DON AND DONNA GO TO BAT	W 2							2.9 (2.6-3.2)	2^L
56. Rambeau: JIM FOREST AND THE BANDITS	G 1.9						2	2.6	1^H
57. Rambeau: ... THE TRAPPER	G 1.7							1.8	Pr
58. Rey: CURIOUS GEORGE	P Pr-1	k-2	3	3		2	3	2.9 (2.4-3.7)	2^L
59. Rey: ... FLIES A KITE	P 1,2	k-2	3			2		2.7 (2.3-3.1)	2^L

TABLE II (continued)

DATA ON READABILITY LEVELS OF SELECTED BOOKS

Author and Title of Book	Professional Judgment					Formula			
	Publisher[1]	Children's Catalog	Harris	Kottmeyer	Sullivan	Eakin	Spache	Spache[2]	Fry[3]
60. Robinson: PICTURE BOOK OF ANIMAL BABIES	L	E			1			2.5 (2.1-2.7)	2^{H}
61. Rowand: GEORGE	R					RA k-3		4.5 (3.2-5.5)	6^{L}
62. Selsam: GREG'S MICROSCOPE	N 2	1-3				2		2.4	1^{M}
63. Selsam: LET'S GET TURTLES	N 2	1-3						2.5 (2.0-3.0)	1
64. Seuss: CAT IN THE HAT	W 1	k-2	1				2	2.3 (2.0-2.9)	PPr
65. Sharp: DAFFY	X	Pr, 1	1				1	1.4	Pre-Pr
66. Stolz: EMMETT'S PIG	N 2	k-3						2.5 (2.0-3.4)	2^{M}
67. Tresselt: RAIN DROP SPLASH	S	k-1			1	RA k-1		3.9 (2.8-4.8)	5^{M}
68. Tworkov: CAMEL WHO TOOK A WALK	A	k-3			1			4.4 (2.7-6.2)	7^{L}
69. Udry: A TREE IS NICE	N	k-1			1	RA k-2		2.5	1^{H}

ANALYSIS OF THE DATA IN TABLE II

Comparisons of Ratings

As the data indicate, the thirty-five publisher ratings were similar to professional judgment ratings in two-thirds of the cases; they were lower or similar to the Spache ratings in approximately the same number of cases; and they were higher or similar to the Fry ratings in approximately the same number of cases.

The forty-four professional judgment ratings were lower than the Spache ratings in three-fourths of the cases, and they were higher in none of the cases; fewer than half of the cases were lower than Fry, and about a third were higher than Fry.

The sixty-nine Fry and Spache ratings break down as follows:

Fry ratings higher than Spache ratings:

3	from two and one-half to three grades higher
5	one and one-half to two grades higher
4	one grade higher

Fry ratings lower than Spache ratings:

3	two grades lower
13	one and one-half grades lower
17	one grade lower

Fry and Spache ratings approximately the same: 24.

In almost half the cases it is apparent that the Fry ratings are significantly lower than the Spache ratings; in more than a third of the cases the ratings are similar; and in fewer than one-sixth of the cases

are the Fry ratings higher than Spache ratings.

Accuracy of the Professional Judgment Scores

Inspection of the data listed for Harris, Kottmeyer, and Sullivan leads to the conclusion that either the children they have had in their remedial clinics were exceptionally good readers or that the books have been rated primarily on the interest factor.

One of the more extreme illustrations is the case of Little House of Your Own (item 20), which according to the Spache formula is rated 3.0, and by the Fry Graph, 4.0. Sullivan's rating of the book at first-grade independent reading level implies that a child of matching reading ability has a facility with non-phonetic word elements as in built, course, and doesn't; that he has a phonics sense that will enable him to sound out umbrella, remember, and secret; that he has an understanding of the compound words everyone, grownups, nobody, and tablecloth; that the sentences, from ten to fourteen words long in these samplings, are not too complex for a first-grader; and that the degree of abstractness is not beyond his grasp. This is a charming little book, actually an appealing little essay on human rights and needs, primary style; it is an excellent read-aloud for kindergarten through about third- or fourth-grade difficulty; but it is very doubtful if a first-grade child could read it himself, without either skipping words or having some word attack difficulties, considering the word attack skills presented at this level.

Harris and Sullivan both place Angus and the Ducks (item 23) at first-grade independent reading level, and Kottmeyer at kindergarten through second grade. The sentence length in this little book is formidable for a child reading at any one of those levels, one sample containing nineteen words; but this factor may be offset by the publisher's device of placing single four-to-five word phrases by themselves on a page. The story contains many words judged "hard" by either a word list or polysyllabic count. However, the subject-matter is interesting, the illustrations are attractive, the writing style is brisk and vivid, and the plot moves along quickly to a satisfying conclusion. The consensus of opinion among the three who judge suitability by professional expertise is that the book can be read independently at kindergarten through first grade, yet the upper level of difficulty as determined by the Spache formula is 4.3. Query: Does the average first-grader possess reading skills that are equal to this kind of challenge?

Accuracy of the Spache and Fry Scores

In thirty-six cases, the appearance of accuracy in the Spache scores is misleading, as it covers more or less a wide range of difficulty. For example, the unwary teacher or librarian might infer from the designation "3.8" for Billy and Blaze (item 1), that a child with reading achievement score of 3.8 could read the book; but actually, the book as a whole is too difficult for a child with less than

a reading ability level of 4.7. Red Fox and His Canoe (item 8) requires a matching reading level of 3.3--the upper level of difficulty--rather than the 2.5 average indicated.

Similarly, the 4.4 Spache rating for Mike Mulligan (item 13) is a whole grade lower than the upper level of difficulty of one of the samples; Boats on the River (item 24) extends to 5.0, rendering the 4.0 average score meaningless; the same underestimate of maximum difficulty is noted in the ratings for Gertie the Duck (item 29), Grizzwold (item 35), George (item 61), and--the most misleading of all--The Camel Who Took a Walk (item 68), with a designated average of 4.4 but with an upper level of 6.2.

It would seem more advisable to adopt as an index of reading difficulty the highest score yielded by the samples instead of the average score indicated by the Spache formula.

The same qualification applies to the Fry sampling scores; however, a Fry designation of "3" is more inclusive than a Spache designation of "3.2" (item 3), and therefore is not so misleading.

Looking at the Fry ratings on the cases just noted, it can be seen than on item 1 the Graph score of fifth grade more nearly represents the ability level called for than does the Spache formula average of 3.8, and on item 13 the Graph rating is closer to the upper difficulty level of the book than is the Spache average.

Explanation of the Variance in Spache and Fry Ratings

This investigator suggests that variance in the estimates obtained by the two formulas, where it does occur, is caused by one or both of two factors, one having to do with the nature of vocabulary measurement and the other having to do with the sampling pattern.

(1) Whenever there is a very marked discrepancy between scores obtained by application of the two formulas to identical samples (as in this study), this variance can be accounted for by the vocabulary factor only, since the sentence factor is fairly constant between them. Thus, a Fry score of 7-low and a Spache score of 4.0 (item 24, Boats on the River) indicate, in all likelihood, that the material contains a relatively large number of multisyllabic words. Since each word, by the Fry formula, carries the weight of its syllables ("usually," for example, has a count of four), the vocabulary count mounts faster than with the Spache formula, where each word carries a weight of one point, regardless of word length.

The element of repetition of certain multisyllabic words or phrases within a sample, as in this story, will also enlarge the Fry score.

The principle operates in reverse in cases where the Fry estimate is much lower than the Spache estimate; for example: on Gertie the Duck (item 29), Fry's rating is 1 and Spache's is 2.6; on Grizzwold (item 35), Fry's rating is 1 and Spache's is 2.7. Words in

these stories tend to be short, thus keeping Fry's scores down; but the occurrence of even short words outside the Spache "easy" list--lake, posts, safe, cheered, themselves, engine; caves, fur, can't, mountain, ow, pole--puts the Spache vocabulary average up to 2.6 and 2.7.

Another feature of vocabulary measurement that may cause variance in scores obtained by the two formulas is the inflexibility of the Spache word list. One aspect of this feature has been noted previously: its increasing tendency to become less timely. A periodic up-dating of the word list would enable formula users to turn to it with more confidence. You Will Go to the Moon (item 26), instead of carrying a Spache designation of 1.8 (41:47) might be more in line with the Fry pre-primer designation, were moon, rocket, space, station, earth, ship, and TV not considered "hard" words.

The word-list approach may cause variance in still another way, by its failure to provide for specialized vocabulary. A case in point is The Sea Hunt (item 9), with these words counted "hard" by Spache: tanks, eels, sharks, nets, hooks; barrel, diving, helmet; porpoises, octopus. Of these, only the two- and three-syllable words were counted "hard" by the Fry formula.

(2) The scores are manipulable by varying the sampling pattern. An infrequent exception to this truism occurs with a book such as Dolch's Once There Was a Cat (item 21), which is carefully

controlled for uniformity of vocabulary and sentence-length difficulty). To shift the pattern even slightly may change the word, syllables, or sentence count and so change the score. Since it would be by the merest chance than any two analysts would draw the identical passages from a given book, the scores derived by any two analysts can be expected to differ, depending upon the difficulty of the passages analyzed. Examples of this variance are shown in the following cases, where the Spache published ratings on various books differ sharply from the scores computed for this study:

> Mike Mulligan (item 13) published rating: level 2-3; computed scores: 5.4, 4.0, 4.4, and 3.6.
>
> Nobody Listens to Andrew (item 31) published rating: level 2.1; computed scores: 1.9, 2.5, and 2.6.
>
> Chester (item 34) published score: level 1; computed scores: 2.1, 2.1, 2.0.

It must be assumed that Eakin, who used the Spache formula for the books she evaluated, has also selected samples of generally lower difficulty than those upon which the published Spache scores were based. In addition, the Eakin scores differ from those computed for this study: only one (Chester, item 34), was similar, and the other three were one grade level below.

Eakin very appropriately indicated six titles as read-alouds, suitable for use in kindergarten through first, second or third grade.

The question arises: What is the value of these grade-level designations if the scores can be so manipulated? The teacher or

librarian must make several judgment decisions before recommending a given book to a particular child, and one of the first considerations must be: How typical of the over-all difficulty of the book are these samplings? It is recommended that enough samplings should be taken to assure a fair picture of the range of difficulty within the book. According to this investigator, it is not the average score on these samples that is significant; rather, it is the highest score, representing the most difficult passage among the samples, that should be noted. Only through an awareness of this full range of difficulty within the book can the teacher insure a successful reading experience for a given child.

SUMMARY

Chapter Three was devoted largely to an examination of the Spache Readability Formula, the Fry Readability Graph, and a comparison of the two through the medium of a selected list of sixty-nine primary-level books.

Limited data on appraisals by publishers and experienced judgment were compared with each other and with each formula.

Ratings by publishers and judges tended to agree with each other, tended to be lower than the Spache ratings, and higher than the Fry ratings. The Fry ratings were lower than those of Spache in a ratio of three to four.

No statement can be made as to which system of appraisal is the more accurate, since each type of data was derived from different

criteria. These were: useful within an approximate age range, useful within a certain grade range, based on sentence length and an "easy"-word list, based on sentence length and word length in syllables, judged on interest and/or other features. It is more important to speak of the practicality of a method, in terms of its efficiency in measuring a limited number of factors that help to determine readability.

The Spache Readability Formula has the weight of tradition behind it in its use of the word list for discriminating between easy and hard words. It also has the distinction of being the only one available to publishers, editors, and educators for evaluating primary books, until recently, and so is a more familiar tool. The word list, however, is in need of up-dating; a reading level expressed in precise to-a-tenth-of-a-month terms can be misleading; and the formula is not valid for books beyond mid-third level.

The Fry Readability Graph is distinguished by its simplicity, its ease of application, and its time-saving feature. The use of poly-syllables as a measure of word difficulty is supported by research, and appears to be particularly valid at the primary level. (11:31; 31:31; 43:52). The Fry formula has the added advantage of being applicable to all ranges of difficulty, from primer to college level.

Whether a given book is evaluated by the Spache formula or the Fry formula, it is important to note the range of difficulty represented by the samples. If the range is wide, and the difficulty of the book is

represented by the "average" score, half of the book will be a frustrating reading experience to the child whose standardized test rating is below that "average."

It should be kept in mind at all times, as Chall (11:56) cautions, first, that formulas measure only a limited number of factors important in reading difficulty; second, that they give only a first approximation of difficulty, which should then "be tempered with experienced judgment."

CHAPTER 4

SUMMARY AND CONCLUSIONS

Research has much to say about the relationships between structural elements of a passage and how these affect readability. There is general agreement, for example, on the identity of the two main predictors of difficulty, a vocabulary and a sentence factor, although points of view differ as to how these factors should be expressed in a formula. But little has been said about research efforts to search out or establish a reliable, external criterion of difficulty to which passages of unknown difficulty can be compared. Formulators have used either "standard" test passages or publishers' designations of grade level for their criterion. Thus it appears to this investigator that criteria for the formula were selected by publisher prestige rather than by an empirical judgment.

A stated purpose of this investigation was to research the term"grade level." In all the literature surveyed for this study, only two particular references were made to the term as such. Fry (21:515) noted that there are no rigorous standards of just what is one grade level as opposed to the next higher grade level.

> There seems to be some loose sort of agreement between publishers and educators which is based on experience and perhaps a little on test data as to what grade level designations mean. However, even standardized test data are not exact. . . . Hence the problem . . . is complicated by trying to determine grade

> level when grade level won't stand still and when subjective "judgments" are about as good a standard as can be found.

Bormuth's generalization is perhaps as close to a definition of the term as can be encountered:

> When a grade placement number is attached to a book, it can be interpreted as the average level of reading achievement attained by children who are able to read the book at the minimum level of comprehension. (6:435).

It is assumed that "the minimum level of comprehension" refers to the instructional level, but of what population of children? Does the term always refer to an instructional level?

Wheeler (51:397-399) has stated that most formulas tend to give instructional levels. Every teacher and librarian should be aware of this general practice before recommending a rated book to a child, particularly if the book is intended for leisure reading.

Bormuth (see page 27) proposed a cloze test procedure that "seems to provide an accurate measure of the difficulty of a passage, almost regardless of how difficulty is measured." (6:429-436). Perhaps research will look to an adaptation of this new method for use at primary levels.

In the meantime, educators, publishers, editors and book reviewers look to the tools they have--the readability formulas--which, with all their limitations, are yet the best instruments for objective measurement that research has provided. Critics do not hesitate to point out the limitations of formula evaluations:

(1) They do not measure concepts.
(2) They do not measure interest.
(3) They do not measure experience background.
(4) They do not measure content.
(5) They do not measure purpose of the reader.
(6) They do nothing that expert judgment cannot do better.

Proponents of the formula procedure counter that no objective tool should be criticized for not doing a job it was never designed to do; that a formula is designed to give only a first approximation of the reading level of a book according to its word and sentence difficulty; that the expressed level is probably within a year's span of the true readability level of the book; and that a formula was never proposed to supplant professional judgment, but rather to supplement this older means of evaluating a book.

The contention of this investigator is that a formula rating is a "must" for any book being considered for use with children of limited reading ability, particularly for those in remedial classes. The teacher seldom has access to any published resource that will give a reliable estimate of the actual reading level of a given book, and it therefore devolves upon him to apply the formula himself. By means of the Fry Readability Graph he can estimate the difficulty of a book in a few minutes.

But this is not the total assessment. It is at this point that subjective evaluation becomes necessary, for the book may be eminently readable by a particular child, yet be entirely unsuitable from the standpoint of interest. Teacher judgment, or that of other professional

persons, based on an understanding of the child's needs, capabilities and interests as well as on a knowledge of the book's difficulty, is the final ingredient in matching the book to the child, to assure a satisfactory reading experience.

SUGGESTIONS FOR FURTHER RESEARCH

These suggestions are offered for further readability research: (1) Up-dating of the Dale-Stone word list; (2) abandonment of the "average" designation of reading level in favor of a stated range of difficulty over all the samples drawn for computation; (3) publication of the fact that a formula evaluation equates with an instructional reading level, not an independent level; (4) empirical determination of the instructional reading level; (5) continued research on both quantitative and qualitative aspects of readability, with emphasis upon objectivity and efficiency in application.

CONCLUSION

The research of Bormuth and his associates is revealing some startling new hypotheses regarding the syntactic and semantic structure of language. For example, according to the conventional wisdom of our time, the word and sentence variables are independent units, causative of comprehension difficulty and measurable as such. Bormuth finds, however, that these variables themselves are dependent upon certain transformations that must be traced back to their underlying forms

to be understood. It is the complexity of each succeeding transformation, Bormuth postulates, that should be used as a measure of difficulty of the passage.

With research "taking off" in this direction, it is perhaps not too visionary to hope that formulas of the future will exhibit an entirely new pattern, one that takes into account many of the variables contributing to readability, and is as practical and available to the classroom teacher as the dictionary on his bookshelf.

REFERENCES CITED

1. Anderson, Jonathan. "Research in Readability for the Classroom Teacher," Journal of Reading, 8:402-403, May, 1965.

2. Betts, Emmett A. "Informal Inventories," pp. 121-125 in Readings on Reading Instruction, ed. Albert J. Harris. New York: David McKay Co., 1963.

3. _____. "Success Levels for Retarded Readers," pp. 442-447 in Readings on Reading Instruction, ed. Albert J. Harris. New York: David McKay Co., 1963.

4. Blair, Allen M. "Everything You Always Wanted to Know about Readability but Were Afraid to Ask," Elementary English, 48:442-443, May, 1971.

5. Bloomer, Richard R. "The Cloze Procedure as a Remedial Reading Exercise," Journal of Developmental Reading, 5:173-181, Spring, 1962.

6. Bormuth, John R. "The Cloze Readability Procedure," Elementary English, 45:429-436, April, 1968.

7. _____. Development of Readability Analyses. (Microfiche, ERIC No. ED 029 166, March, 1969). 171 p.

8. _____. Empirical Determination of the Instructional Reading Level, paper presented at the International Reading Association Conference, Boston, April, 1968. (Microfiche, ERIC No. ED 020 084). 6 p.

9. _____. New Data on Readability, paper presented at a meeting co-sponsored by the International Reading Association and the American Educational Research Association, Seattle, May, 1967. (Microfiche, ERIC No. ED 016 586). 11 p.

10. _____. "New Developments in Readability Research." Elementary English, 44:840-845, December, 1967.

11. Chall, Jeanne S. Readability: An Appraisal of Research and Application; Bureau of Educational Research Monograph No. 34. Columbus: Ohio State University Press, 1958. 202 p.

12. Children's Catalog: a Catalog of 3310 Selected Books for Public and School Libraries. Ed. by Dorothy Herbert West and Rachel Shor. Tenth ed. New York: Wilson Co., 1961.

13. Children's Catalog. Ed. by Rachel Shor and Estelle A. Fidell. Eleventh ed. New York: Wilson Co., 1966.

14. Clymer, Theodore. "A Study of the Sampling Reliability of the Spache Readability Formula," pp. 245-250 in Reading in Changing Society; International Reading Association Conference Proceedings, vol. 4, ed. J. Allen Figurel. New York: Scholastic Magazines, 1959.

15. Coleman, E. B. "Experimental Studies of Readability: Part I," Elementary English, 45:166-178, February, 1968.

16. _____. "Experimental Studies of Readability, Part II," Elementary English, 45:316-324, 333, March, 1968.

17. Dawkins, John. "A Reconsideration of the Dale-Chall Formula," Elementary English, 33:515-519, December, 1956.

18. Dolch, Edward W. Problems in Reading. Champaign: Garrard Press, 1948.

19. Eakin, Mary K. (comp.). Good Books for Children: A Selection of Books Published 1950-1965. Chicago: University of Chicago Press, 1966.

20. Farr, Roger. "Problems in Measuring Reading Sub-Skills," pp. 33-71 in Reading: What Can Be Measured? Newark, Delaware: International Reading Association Research Fund, 1969.

21. Fry, Edward B. "A Readability Formula That Saves Time," Journal of Reading, 11:513-516, April, 1968.

22. _____. "A Readability Graph for Librarians, Part I," School Libraries, 1:13-16, Fall, 1969.

23. _____. "The Readability Graph Validated at Primary Levels," Reading Teacher, 22(6):534-538, March, 1969.

24. Gates, Arthur I. "Construction of a Reading Vocabulary for the Primary Grades," Teachers College Record, 27:625-642, March, 1926.

25. Gaver, Mary V. "A Readability Graph for Librarians, Part II," School Libraries, 1:23-25, Fall, 1969.

26. Gray, William S. "Progress in the Study of Readability," Elementary School Journal, 47:491-498, May, 1947.

27. Gunning, Robert. The Technique of Clear Writing. Rev. ed. New York: McGraw-Hill Book Co., 1968.

28. Ham, Catherine. "Aspects of Books That Affect Readability and Use: In Literature," pp. 167-170 in Matherials for Reading, Supplementary Educational Monograph No. 86, ed. Helen M. Robinson. Chicago: University of Chicago Press, 1958.

29. Harris, Albert J. How to Increase Reading Ability: A Guide to Developmental and Remedial Methods. Fourth ed. rev. New York: David McKay Co., 1961.

30. Johnson, George R. "An Objective Method of Determining Reading Difficulty," Journal of Educational Research, 21:283-287. April, 1930.

31. Klare, George R. "The Role of Word Frequency in Readability," Elementary English, 45:12-22, January, 1968.

32. Kottmeyer, William. Teacher's Guide for Remedial Reading. St. Louis: Webster Publishing Co., 1959.

33. Kress, Roy A. "Finding Readable Materials for Remedial Reading," pp. 94-112 in Readability: Finding Readable Material for Children: A Report of the Tenth Annual Conference on Reading, ed. Gerald A. Yoakum. Pittsburgh: University of Pittsburgh Press, 1954.

34. Lorge, Irving. "Predicting Readability," Teachers College Record, 45:404-415, March, 1944.

35. Manzo, Anthony V. "Readability: A Postscript," Elementary English, 47:962-965, November, 1970.

36. McCullough, Constance M. "Linguistics, Psychology, and the Teaching of Reading," Elementary English, 44:353-362, April, 1967.

37. Pauk, Walter. "A Practical Note on Readability Formulas," Journal of Reading, 13(3):207-210, December, 1970.

38. Potter, Thomas C. A Taxonomy of Cloze Research, Part I: Readability and Reading Comprehension. Inglewood: Southwest Regional Laboratory for Educational Research and Development, June, 1968. (Microfiche, ERIC No. ED 035 514). 51 p.

39. Rheay, Mary L. "Grading Children's Books," Elementary English, 36:478-479, November, 1959.

40. Schiller, Andrew. "Teach Children to Read Their Father Tongue," pp. 17-32 in Co-ordinating Reading Instruction, ed. Helen M. Robinson. Glenview: Scott, Foresman and Co., 1971.

41. Spache, George D. Good Reading for Poor Readers. Rev. ed. Champaign: Garrard Publishing Co., 1966.

42. _____. "A New Readability Formula for Primary-Grade Reading Materials," Elementary School Journal, 53:410-413, March, 1953.

43. Stolerow, Lawrence M. and J. Robert Newman. "A Factional Analysis of Objective Features of Printed Language Presumably Related to Reading Difficulty," Journal of Educational Research, 52(7):243-251, March, 1959.

44. Stone, Clarence R. "Measuring Difficulty of Primary Reading Material: A Constructive Criticism of Spache's Measure," Elementary School Journal, 57:36-41, October, 1956.

45. Sullivan, Helen B. and Lorraine E. Tolman. "High Interest-Low Vocabulary Reading Materials: A Selected Book List," Journal of Education, 139(2):9-59, December, 1956.

46. Thorndike, Edward L. A Teacher's Word Book of Twenty Thousand Words. Rev. ed. New York: Bureau of Publications, Teachers College, Columbia University, 1932.

47. _____. "Improving the Ability to Read," Teachers College Record, 38:123-144, November, 1936.

48. Veatch, Jeannette. " 'Linguistic' Instruction in the Teaching of Reading: Kill or Cure?" Elementary English, 39:231-233, 243, March, 1962.

49. Washburne, Carleton and Mabel V. Morphett. "Grade Placement of Children's Books," Elementary School Journal, 38:355-363, January, 1938.

50. Weintraub, Samuel. "What Research Says about Learning to Read," pp. 180-201 in Co-ordinating Reading Instruction, ed. Helen M. Robinson. Glenview: Scott, Foresman and Co., 1971.

51. Wheeler, Lester R. and Edwin H. Smith. "A Practical Readability Formula for the Classroom Teacher in the Primary Grades," Elementary English, 31:397-399, March, 1954.

APPENDIX A

PUBLISHERS' REPLIES TO QUESTIONNAIRES

Field Educational Publications, Inc.
A SUBSIDIARY OF FIELD ENTERPRISES, INC.

Robert K. Hendin
Assistant Vice President
Managing Editor/Language Arts

July 14, 1970

Mrs. Kermitt C. Ball

Dear Mrs. Ball:

Thank you very much for your letter of June 19, 1970. The formula we use for assessing readability levels for grades 1-3 is the Spache Readability Formula and for higher grades the Dale-Chall Formula.

Additional Deep Sea titles are currently being readied for publication in the spring of 1971; and did you know that there are now eight titles in the Checkered Flag series and twelve titles of the Jim Forest Readers?

We do certainly appreciate your enthusiasm about the Harr Wagner books for remedial readers and are, of course, very gratified to hear that they have been the mainstay of your program for years. Thank you again for your letter, and I hope the above information will be of value to you in your master's thesis.

Sincerely,

Robert K. Hendin

RKH:mg

LETT PUBLISHING COMPANY

July 10, 1970

Mrs. Kermitt C. Ball

Dear Mrs. Ball:

In the past we have used the Dale-Schall Test and the Spache Readability test as well as the Botel Readability formula. But I must say what we rely on principally is long experience with the reading ability of children in the primary grades. Editors with this kind of experience can judge the reading level of a book quite accurately by looking at the text.

Sincerely yours,

(Mrs) Esther K. Meeks
Editorial Director
Children's Book Department

EKMrw

Please Note: Personally Identifiable Information was redacted due to privacy concerns and signature was redacted due to security concerns.

June 19, 1970

Juvenile Books Editor
Golden Press, Inc., Division of
Western Publishing Co., Inc.

Dear Editor:

This is a request for information, if you please, as to how you assess readability levels of your trade books for primary grades. As a remedial reading teacher I am of course concerned about the readability aspect of a given book, and such information straight from the publishers would be particularly valuable to me now in preparation of a master's thesis.

If you will take a moment to check or reply to the items below and return this sheet to me in the enclosed self-addressed envelope, I shall appreciate it very much.

Very truly yours.

Do you use a formula? No

If so, which one? _____

If not, by what method do you assign grade level? Inspection by people who have had experience in the field of reading instruction.

Please Note: Personally Identifiable Information was redacted due to privacy concerns and signature was redacted due to security concerns.

June 19, 1970

Miss Margaret K. McElderry, Juvenile Books Editor
Harcourt. Brace & World. Inc.

Dear Miss McElderry:

This is a request for information, if you please, as to how you assess readability levels of your trade books for primary grades. As a remedial reading teacher I am of course concerned about the readability aspect of a given book, and such information straight from the publishers would be particularly valuable to me now in preparation of a master's thesis.

If you will take a moment to check or reply to the items below and return this sheet to me in the enclosed self-addressed envelope, I shall appreciate it very much.

Very truly yours.

Do you use a formula? No

If so, which one?_____

If not, by what method do you assign grade level?_____

Please Note: Personally Identifiable Information was redacted due to privacy concerns and signature was redacted due to security concerns.

1817

Harper & Row, Publishers

New York Evanston London

June 29, 1970

Mrs. Kermitt C. Ball

Dear Mrs. Ball,

Please forgive the delay in answering your letter of the 19th. We do not use a readability formula. Grade levels are assigned to books according to the story and vocabulary content. (We do not use vocabulary lists, by the way.) Since reading ability varies from child to child, we invariably suggest an age-span covering three or four years. In this way we are reasonably certain that the story will appeal to the child whether he reads it himself or it is read to him.

Yours sincerely,

Barbara A. Dicks
Office Manager
Harper Junior Books

BAD/jvr

Please Note: Personally Identifiable Information was redacted due to privacy concerns and signature was removed due to security concerns.

June 19, 1970

Mrs. Phyllis Cerf, Juvenile Books Editor
Random House, Inc.

Dear Mrs. Cerf:

This is a request for information, if you please, as to how you assess readability levels of your trade books for primary grades. As a remedial reading teacher I am of course concerned about the readability aspect of a given book, and such information, straight from the publishers, would be particularly valuable to me now in preparation of a master's thesis.

If you will take a moment to check or reply to the items below and return this sheet to me in the enclosed self-addressed envelope, I shall appreciate it very much.

Very truly yours.

Do you use a formula? Yes.

If so, which one? It depends on book being graded

If not, by what method do you assign grade level? _____

* Below 3rd. grade reading level: Spache
* above " " " " : Dale-Chall

These formulas are very helpful, but in the hands of amateurs who can't bring sufficient judgement to the property being graded, they often produce poor results resulting for children who want want to read.

June 19, 1970

Miss Velma V. Varner, Editor, Junior Books
Viking Press, Inc.

Dear Miss Varner:

This is a request for information, if you please, as to how you assess readability levels of your trade books for primary grades. As a remedial reading teacher I am of course concerned about the readability aspect of a given book, and such information straight from the publishers would be particularly valuable to me now in preparation of a master's thesis.

If you will take a moment to check or reply to the items below and return this sheet to me in the enclosed self-addressed envelope, I shall appreciate it very much.

Very truly yours,

Do you use a formula? No

If so, which one? _____

If not, by what method do you assign grade level? Interest and appeal

_____

Please Note: Personally Identifiable Information was redacted due to privacy concerns and signature was removed due to security concerns.

APPENDIX B

PUBLISHERS OF LISTED BOOKS

A Aladdin Books
Ariel Books (see
Holt, Rinehart and Winston)
B Beckley-Cardy Company
C Benefic Press
D Thomas Y. Crowell Company
E Dial Press
F Doubleday and Company
G Field Educational Publications, Inc.
H Follett Publishing Company
I Garrard Publishing Company
J Golden Press
K Grosset and Dunlap
L E. M. Hale and Company
M Harcourt, Brace and World
N Harper and Row
Harr Wagner Company (see
Field Educational Publications)
O Holt, Rinehart and Winston
P Houghton, Mifflin Company
Q J. B. Lippincott Company
R Little, Brown and Company
S Lothrop, Lee and Shepard
T Macmillan Company
U William Morrow and Company
V Oxford University Press
W Random House
X Steck Company
Y Viking Press