

A REVIEW OF OBJECTIVE MEASURES FOR PREDICTING THE READABILITY OF SCHOOL BOOKS

Graham Wagner

Introduction

A good reading programme has informal and formal aspects, a balance between methods that motivate and develop a love for reading and good literature and methods that ensure that children have opportunities to read material which is at their level of reading ability. The best of children's literature, especially that presented in picture book form, is not produced or judged according to readability formulas but plays a vital role in the development of language and literacy (Elley and Mangubhai, 1981). However, readability level is a vital consideration when books are to be used for instructional purposes whether for improving reading ability or for reading in content areas.

Most English language teachers consider that they know best what their students' reading needs are. Yet even when carefully trained, research evidence indicates that teacher estimates of reading difficulty are largely subjective and open to considerable bias. What is required is a method which helps a teacher make a more objective assessment when matching books to her class. Objective methods are available. The teacher can either directly assess the books actual readability by testing it out on the class, using for example, the cloze procedure (Stamp, et al, 1979) or she can predict its likely readability using a readability formula. It is the last method that should prove of most use to the busy classroom teacher, given that most decisions about the suitability of books will be made out of class and in the absence of the children.

Readability Formulas

Readability formulas predict the likelihood that a selected extract, or extracts, from a book can be understood by a criterion group. In most cases the formulas produce a score, or difficulty index, that is translatable into Grade Level (usually USA) or Reading Age Level (more common in the UK, New Zealand and Australia).

Formulas in general, use word counts (vocabulary difficulty) and sentence variables (structural complexity) in a regression equation which is usually worked out by hand, although increasingly nowadays the calculations are undertaken by a computer which has been programmed accordingly.

In his review of the latest trends in readability formulas, Klare (1974-75) concludes that a simple two variable formula should be suitable for most

purposes, especially if one variable is a word or semantic variable (such as counting word length or counting the number of words not on a special list of familiar words) and the other is a sentence or syntactic variable (such as the average sentence length). Of the two the word variable is the more highly predictive when each is considered singly, for other languages as well as for English, but the sentence variable can also make an important contribution too. Klare says "It may seem surprising that counts of the 2 simple variables of word length and sentence length are sufficient to make relatively good predictions of readability. No argument that they *cause* ease or difficulty is intended; they are merely good *indices* of difficulty." (p. 97). As can be seen below not all formulas use, nor need, a sentence variable.

Two Long Established Readability Formulas

Most readability formulas can be a little "off-putting" unless one has ready access to a calculator. For example, two well-known formulas are expressed mathematically as follows:

- (1) Dale-Chall formula (1948)

$$Xc50 = .1579X_1 + .0496X_2 + 3.6365$$

Where $Xc50$ = reading grade score of a pupil who could answer one half the test questions on a passage correctly;

X_1 = Dale score, or percentage of words outside the Dale list of 3000;

X_2 = average sentence length in words.

- (2) Flesch formula (1948)

$$RE = 206.835 - .846wl - 1.015sl$$

Where RE = Reading Ease;

wl = number of syllables per 100 words;

sl = average number of words per sentence.

(To change RE scores to US school grades four algorithms are used).

Not all formulas are mathematical regression equations such as the Dale-Chall and the Flesch. Some like the Fry Readability Graph and Mugford Chart use graphs and charts to simplify calculations. In a few cases such as the Elley (1969) Noun Frequency Method there is no regression equation involved but like the Botel (1962) method of predicting readability it is used as if it were a formula. Furthermore it is included here because of its high validity as a measure of predicting readability in the South Pacific and because, more importantly, it is comparatively easy to use.

Limitations of Readability Formulas

The limitations of Readability Formulas have been reported elsewhere (Klare, 1963); (Blair, 1971); (Carozzi, 1972); Otto and Smith, 1970); Tremaine and Wagner, 1980).

Briefly the following points should be noted:

- (1) Formulas measure only one aspect of writing-style: they indirectly touch on the content, and say nothing about organisation, word order, format or imagery. They do not take into account the motivation, previous experience and intelligence of the readers.
- (2) Formulas measure only one aspect of style-difficulty: they do not measure a writer's dramatic effectiveness or ability to create a mood.
- (3) Formulas do not even measure difficulty perfectly: they are seldom accurate to (or even within) one grade or reading age level. Furthermore they are prone to sampling and application error.
- (4) Formulas are not measures of good style: poor writing may be quite readable according to a formula score but this information would not help the writer to improve his writing.

Even though style is only one characteristic of a piece of prose, it is an important one in terms of its relationship to reading speed, acceptability, understanding and learning. Klare (1963) puts the case for formulas when he says: "If formulas are thought of as efficient predictors of difficulty, more accurate in prediction than individual writers most of the time, that is all that should be expected" (p.25). Therefore a readability index is a useful starting point for a writer in editing his work, and for that matter, a teacher or parent in choosing books for children.

Selecting an Appropriate Readability Formula

With over 100 formulas to choose from it is a difficult job to sift through the mountain of data available and pick the right one. Fortunately, there have been extensive reviews of readability formulas in recent years which are designed to make the selection of the right formula, for the task in hand, a comparatively easy matter.

The following is an adaptation of an NZCER Set article (Wagner, 1981):
Which Readability Formula Should the Teacher Use?*

There are two major selection criteria:

1. The predictive accuracy of the formula.
2. The speed with which an answer can be obtained.

Secondary criteria for consideration are:

3. The range of applicability (whether it was to be used on children's and/or adult reading material).
4. Its ability to measure the level of abstraction and difficulty of ideas rather than mere style difficulty.
5. Its usefulness for assessing reading materials not in ordinary test (e.g., tests and questionnaires).

Because of the exclusiveness of the criteria it is impossible to designate one best formula. However, here are some general suggestions which apply in

* See Appendix I for References.

most cases to formulas intended primarily for adult materials but which will in most cases also reach down to low school reading age levels. These are as follows:

(1) *The Most Accurate Formula:*

Dale-Chall formula: See the recalculated versions by Powers, Summer and Kears (1958) and Holmquist (1968).

(2) *The Fastest Formulas:*

1. Farr-Jenkins-Paterson: Simplification of the Flesch Reading Ease Formula.
2. Gunning: Fog Index.
3. Fry: Readability Graph.
4. Mugford: Readability Chart.
5. McLaughlin: Smog Grading Formula.

(3) *The most popular formula:*

Flesch Reading Ease formula: This is a compromise choice between the Dale-Chall and the Farr-Jenkins-Paterson version of the Flesch).

(4) *A formula for non-text material:*

Forbes-Cottle formula: This formula is designed for use with psychological tests and inventories but relies on an out-of-print version of the Thorndike Junior Century dictionary and is therefore rather limited in its application.

While the Flesch Reading Ease formula extends down to 10-year-olds and the Dale-Chall formula to 9-year-olds there are advantages in using a formula designed specifically for children. A list of suggestions is as follows:

(5) *Popular general formulas for children's materials:*

Bormuth formulas — 35, 45, 55 per cent cloze criterion formulas (9-17 years).

Elley Noun Frequency Method: Relates to PAT levels 1-10 (from 8 to 15 + years).

Washburne-Morphett formula: 6-to 14-year-olds:

Lorge formula: (8- to 17 year olds).

The Washburne-Morphett is the most widely used overseas and is considered the most accurate. The Elley Noun Count has more potential in Australia, New Zealand and the South Pacific.

(6) *Formulas for the lowest children's levels:*

Spache formula: 6- to 8-years.

Wheeler and Smith formula: 5- to 9-years

Harris and Jacobson formulas: 6- to 9-years and 9- to 12-years.

(7) *Formulas for special tasks:*

Jacobson formula: Chemistry and physics textbooks from high school to college level.

Rogers formula: Oral comprehension 6- to 17-years.

Shaw formula: 9-year-old science.

Henshall formula: (shorthand dictation).

(8) *A formula for measuring abstractness:*

Bloomer formula: Approaches the problem of measuring the abstractness of writing by working out the adverbial load.

Selecting the Best Formula

Without a doubt, validation research on readability formulas in Europe and North America over the years has proved that the Dale-Chall formula is consistently the best. Be that as it may, the Dale-Chall is also one of the most tedious to use. Hence, teachers and authors have preferred easier, but slightly less valid measures as long as they can be assured that these have an acceptably high degree of age level accuracy (i.e. can predict true reading levels with no more error than about plus, or minus, 6 months).

In a recent comparative study of well-known formulas, or their equivalents, Harrison (1980) demonstrated that two formulas, the Dale-Chall and the Mugford had the highest levels of Validity and Age Level Accuracy. ~~although both missed out somewhat in an ease of application comparison.~~ however both missed out somewhat especially the Dale Chall, in an ease of application comparison.

Although mentioned, but not reviewed by Harrison, a readability measure used extensively in New Zealand, and to a certain extent in Australia and the South Pacific, is the Noun Frequency Method (Elley, 1969). This method has the high predictive validity and reliability of the Dale-Chall and the Mugford with the additional bonus of being easier to use. Because of the Noun Count's obvious utility in this part of the world, and because the Mugford is new and one of the few readability measures with the validity and reliability of the Dale-Chall developed outside North America, both measures are utilised below to illustrate how readability assessment of the non-mathematical kind works. It is left to the reader to decide which one to use, for both have all the requirements of a good readability formula.

It should be noted that neither the Mugford nor the 2000 Nouns used in the Noun Frequency method are included here because they are readily available elsewhere (See Elley, 1980, and Mugford, 1970).

Two Readability Formulas

Readability formulas usually require the use of a small pocket calculator to work out the calculations involved although the two that follow use tables or a word frequency list so that the degree of mathematical skill required by the teacher, or author, is nothing more than the ability to add and divide numbers.

1. *The Mugford Chart*

The Mugford Chart (1970) intended for use with children from 5½ years to 15 years, has been “intuitively” derived after years of research and classroom tests. It has now been extended for use with adults.

Mugford uses a letters-per-word variable that takes into account monosyllabic words like “queue” which many formulae would ignore. In addition, Mugford uses a sentence length in words variable and like the Noun Frequency Method takes repetition into account. To get a final score involves only the addition of four whole numbers. The score, given as a Difficulty Index, is interpreted as a UK Reading Level if it is less than 16. (To turn a UK Reading Level into a US Grade Level simply deduct a constant 5. To do the reverse add a constant of 5).

The method is as follows:

- (1) Count a 100 word sample treating hyphenated words (e.g. close-fitting) as separate words. Count contracted words such as couldn't as one word. Count each number expressed in figures as a word.
- (2) Make up four lists of words in the sample as follows:
List 1 Contains all *polysyllabic* words (i.e. words of three or more syllables).
List 2 Contains *non-polysyllabic* words (i.e. words of one or two syllables) seven or more letters long.
List 3 and 4 Contain respectively the six-letter and the five-letter words of one or two syllables.

Note: Do not list any word more than once and count derived words as separate from the base words (e.g. child, children, children's should be regarded as separate words). Proper nouns, representations of animal noises, nonsense words, etc., should be treated in the same way as ordinary words, but numbers expressed in figures should not be *listed*.

- (3) Count the number of words in each list. Find the word length score for each list from a Table designed for this purpose, then add these scores together to obtain the word length score for the passage.
- (4) Next, count the number of sentences noting that, “Where are you?” he cried., is classified as one sentence only.
- (3) Find the Difficulty Index by inserting the word length score from 3 above in the appropriate column of a Second Table. (If the Difficulty Index is 7.2 or less the user is referred to the article by Mugford in the references for the special procedure designed to determine reading age levels between 5.5 and 7.2 years).

2. *The Noun Frequency Method*

Although not a regression equation the Noun Frequency Method, devised by Elley (1969), is used as a formula. Elley found, after investigating various approaches to assessing readability, that the best single predictor of

readability was some measure of vocabulary load. Choosing nouns he discovered that the frequency of nouns in five series of passages correlated more highly than other measures (including the Dale-Chall) with judges' criteria of difficulty for the passages. This method consists of choosing at least 25 nouns and then assessing their frequency using a special list of 2000 nouns. (See Elley (1980) for details.)

Briefly, the method consists of:

- (1) Select three passages at random from a story or book, each long enough to contain at least 25 different nouns.
- (2) Identify the nouns in each passage and using a special list of 2000 nouns, record the frequency level of each noun. Nouns on the list are rated from 1 to 8. Nouns not on the list are rated 9.

Note the following:

- (a) Nouns in a passage are counted only once.
 - (b) Do not count people's names.
 - (c) Give plural nouns the same rating as single nouns.
 - (d) Hyphenated words follow the normal rules.
 - (e) Doubtful words are best omitted (e.g. gerunds, abbreviations, and recently covered, widely known words not on the list).
- (3) Add up the frequency level numbers and divide by the number of nouns.
 - (4) Refer to Table I in Appendix II to determine the approximate age range level for which the reading material is considered suitable for instructional purposes.

Two Examples to Demonstrate How the Mugford Chart and Noun Count Methods Are Used

Extract One Taken from "Stories for Us — Part 3", p.33, and used in Fiji in the third term of Class 5.

(Note: Nouns have been italicised for later use.)

"Long ago there were no coconut *trees* in *Samoa*. At that *time* there lived a *woman* called *Pai* on the *island* of *Savaii*.

One *day* *Pai* wanted some salt *water* for cooking and she went down to the *sea* to get some. In those days *people* did not have *buckets* because there were no *shops* then. But the *Samoans* knew how to make *pots* and *cups* from *clay*, and *Pai* carried a cup made of clay when she went down to the *sea*. She filled it up with *water* and returned *home*. While she was walking, she saw a tiny *eel* swimming in the cup. She decided to show it to her *daughter*, *Sina*.

Sina was excited when she saw the tiny *eel*. She decided to keep it. "It will make a nice *pet*," she said. "I'm going to feed it and make it grow bigger."

Sina put the *eel* in a cup which was filled with *sea water*, and every day she fed her *pet*. Soon the *eel* grew too big for the cup. *Sina* went to her *father*

and said, "Where can I keep my eel, Father? It's too big to live in the cup any longer."

"I'll dig a *pond* for it," her father said.

Sina's father dug a pond which was large enough for the eel to move about in the water. Sina was very pleased. Every day she went down to the pond to feed her eel. She thought it was a wonderful pet until one day the eel made her feel very frightened.

On that day Sina's *parents* were away in their *plantation*. When Sina went to feed the eel, she found it waiting at the *place* where she always gave it its *food*. The girl bent down with some food in her hand and in the next minute, she jumped back with a cry.

Mugford Chart Analysis of Extract One

Note that two sections of 100 words each have been used to assess the overall Difficulty Index. The first section starts with the word "Long ..." and the second with the word "Sina's ..."

Section I — Number of Sentences = 6.65 i.e. 7

List 1 (polysyllabic words)	List 2 (one or two syllable words - 7 letters or more)	List 3 (Six letter words)	List 4 (Five letter words)
coconut	cooking	called	there
Savaii	buckets	island	trees
Samoan	because	wanted	lived
Samoa	carried	people	woman
(4)	returned	filled	water
(16)	walking	(5)	those
	(6)	(8)	shops
	(15)		while
			(8)
			(8)

$$\begin{aligned}\text{Word length} &= 10 + 15 + 8 + 8 \\ &= 47\end{aligned}$$

$$\text{Difficulty Index} = 8.5$$

Section II — Number of Sentences = 6.86 i.e. 7.

List 1	List 2	List 3	List 4
wonderful	pleased	father	Sina's
plantation	thought	enough	which
(2)	frightened	always	large
(8)	parents	jumped	about
	waiting	(4)	water
	minute	(6)	Every
	(6)		until
	(15)		their
			found
			place
			where
			(11)
			(11)

$$\begin{aligned}\text{Word length} &= 8 + 15 + 6 + 11 \\ &= 40\end{aligned}$$

$$\text{Difficulty Index.} = 8.0$$

$$\begin{aligned}\text{Average Difficulty Index over both sections} &= (8.0 + 8.5 + 2) \\ &= 8.25\end{aligned}$$

$$\begin{aligned}\text{UK Reading Level} \\ &= 8.25 \text{ years}\end{aligned}$$

Noun Frequency Method Analysis of Extract 1

Note that the extract should contain at least 25 separate nouns. In the extract counting stops with the sentence ending with the word "food".

Noun Count

trees 1	Savaii 9	people 1	cups 2	pet 2
Samoa 9	day 1	buckets 5	clay 3	father 1
time 1	water 1	shops 2	home 1	pond 4
woman 2	plantation 6	Samoans 9	eel 9	parents 3
island 3	sea 1	pots 6	daughter 4	place 1
				food 2

$$\begin{aligned}\text{Average Frequency Level} &= \frac{89}{26} \\ &= 3.42\end{aligned}$$

By reference to the Chart in Appendix II it is possible to work out the Approximate Reading Age Level. In this case it is *8 to 9 years*.

Extract Two Taken from "Link 4 — Pupils Book", p. 141 used in Fourth Form English classes in Fiji.

"They knew they were in a desperate *situation*, no matter how much the *wind* might drop. The *ship* stayed firmly stuck on the *sandbar*, and they had to think of saving their *lives* as best they could. A *lifeboat* had been tied to the *stern* before the *storm*, but it must have broken free and had either sunk or to think of saving their *lives* as best they could. A *lifeboat* had been tied to the great *deal of doubt* as to how they could release and launch it safely, while being thrown about by the heavy seas. However, there was no *time* to debate the *problem*, for the ship was still taking a battering from the *waves* and could break into *pieces* at any *moment*.

The *mate* took hold of the lifeboat. With the help of the *rest* of the *men*, he fought against the *weather* to release and lower the lifeboat over the ship's *side* without being crushed by it as it swung wildly. The men then took their *lives* into their *hands* and jumped into the lifeboat as it tossed and dipped insanely below them. With the *salt* burning their *eyes*, they peered ahead through the *spray* to find the *shore*. Their only hope was to find a bay or the mouth of the river where they could shelter in the smoother water. But would their tiny craft reach the shore?"

Mugford Chart Analysis of Extract Two

Note that two sections have been used. The first starts with the word "They ..." and the second with the words "The mate ...".

Section 1: Number of sentences = 4.2 i.e. 4

List 1	List 2	List 3	List 4
desperate	sandbar	matter	might
situation	lifeboat	stayed	stuck
another	release	firmly	think
However	(3)	saving	their
(4)	(7)	before	lives
(16)		broken	could
		either	stern
		driven	storm
		launch	board
		safely	there
		thrown	great
		(11)	doubt
		(17)	while
			being
			about
			heavy
			could
			(17)
			(17)

Reference to a Mugford Chart gives the above word length scores and the Difficulty Index as follows:

Word length = 57
 Difficulty Index = 10.1
 UK Reading Level = 10.1 years

Section II: Number of Sentences = 5

List 1	List 2	List 3	List 4
insanely	lifeboat	fought	lower
(1)	against	wildly	being
(4)	weather	jumped	swung
	release	tossed	lives
	without	dipped	their
	crushed	peered	hands
	burning	(6)	below
	smoother	(9)	ahead
	through		spray
	shelter		shore
			mouth
	(10)		river
	(25)		where
			could
			water
			ships
			(16)
			(16)

Word Length = 54
 Difficulty Index = 9.4
 UK Reading Level = 9.4

Average difficulty index for Sections 1 and 2 = $9.75 (10.1 + 9.4) \div 2$

Average UK Reading Age = 9.75 years

Noun Frequency Method

situation 8	stern 9	problem 7	men 1
wind 2	storm 4	waves 6	weather 2
ship 2	sea 1	pieces 2	side 2
sandbar 9	board 3	moment 5	hands 2
lives 2	doubt 8	mate 4	salt 3
lifeboat 9	time 1	rest 2	eyes 2
	deal 5	shore 4	spray 8

Noun Count Total = 113
 Number of Nouns = 27

$$\begin{array}{r} \text{Average Frequency Level } \frac{113}{27} \\ - 4.2 \end{array}$$

Approximate Age Range = 9 to 10 years

Discussion

It is usual to take at least three extracts from a story or book to obtain an indication of its overall readability. In fact more than three extracts are required if the book is a text book containing different contributions by different authors. Naturally, there is a limit to what is practical so in the case of the Link 4 Pupils' Book (which contains 8 Units) 3 extracts from each Unit might be considered reasonable to give an overview. If three extracts are taken, even though the teacher or author would be advised to select extracts from each third of the story or book, the selection within each third should be randomly chosen.

The two extracts used above are therefore of limited use when it comes to determining whether the publications they come from are written at the same uniform standard throughout and are readable (i.e. understandable) by the criterion group they are designed for. That being so, some tentative interpretations can be made to show that readability indices are only the starting point in an evaluation process leading to the acceptance or rejection of a book or story.

The first thing to note is that there is little to choose from between the Mugford Chart and the Noun Frequency Method. Both approaches to the assessment of readability produced almost equivalent reading age levels on the demonstration extracts. Given that they are both highly valid and reliable as predictive readability measures, they are also both comparatively easy to use, although the Noun Frequency Method is the easier of the two. While both can be used for children between 8 and 15 years the Mugford Charts have the added advantage of extending down to 5½ years and extending past 15 years to cover adult levels of readability. Both would appear to have face validity in the South Pacific given this region's heritage of English language teaching emanating from the United Kingdom, Australia and New Zealand. Even so it would be a useful research exercise to norm both formulas on different South Pacific populations to investigate whether there are ESL cultural differences which might influence the interpretation of scores.

A close examination of the two demonstration extracts indicates that an analysis of extracts like these should provide authors with useful ideas about improving their writing to make it more readable. Although the often quoted maxim of short sentences and simple vocabulary is generally helpful in leading to readable writing this is not always the case, especially with

advanced text books. Even so an author would be advised when writing school books to use as few polysyllabic words as possible, in simple rather than complex sentences.

As mentioned above, it is unwise to draw any conclusions about the levels of readability of the books that provided the two demonstration extracts. Clearly a comparative study is potentially possible and would aid teachers in how they might use the text books in classroom work. Furthermore, it is comparisons of this kind, applied to the whole range of children's books used in schools, libraries and homes, that demonstrate the wide utility of readability formulas. Nevertheless, readability formulas should be used strictly according to the rules laid down by their designers, and the indices or scores interpreted carefully before any comparison is made. A lot more information on children's reading is required before teachers can confidently say that in an ESL context, Class 5 children should be reading at an 8 or 9 or 10 or 11, etc., reading age level.

Conclusion

This paper has looked at a wide range of readability formulas and provided guidelines on selecting a valid, reliable and useable formula for choosing children's books for classroom use and home reading. Some caveats were made about the use of readability formulas and the interpretation of results.

The Mugford Chart and Noun Frequency Method were used in this paper to illustrate how to undertake and interpret a readability analysis of two extracts from two different English Language books used in Fiji schools.

Both readability measures were tentatively judged suitable for use in the South Pacific.

Readability formulas are not a panacea for matching books to reading age levels. In the average school classroom there will be a wide range of reading abilities which the busy teacher will try to take into account when she plans her reading programme for the year. That being so, readability formulas are a very useful indicator of a book's readability, especially when it is not possible to directly test the books out on a target group of readers. In this respect, readability formulas, like the Mugford Chart and the Noun Frequency Method, have more efficacy in the classroom than they are often credited with.

Acknowledgement: Grateful thanks to Barbara Moore and Elizabeth Wagner for their helpful suggestions during the preparation of this paper.