



P-ISSN 2355-2794
E-ISSN 2461-0275

Readability Formulas: An Analysis into Reading Index of Prose Forms

Revathi Gopal^{*,1}
Mahendran Maniam¹
Noor Alhusna Madzlan¹
Siti Shuhaida binti Shukor¹
Kanmani Neelamegam²

¹Department of English Language and Literature, Faculty of Languages and Communication, Sultan Idris Education University, Tanjong Malim, Perak Darul Ridzuan 35900, MALAYSIA

²Department of Science and Humanities, Faculty of Engineering, University College of Engineering Kanchipuram, A Constituent College of Anna University, Bangalore, Tamil Nadu 631552, INDIA

Abstract

Text comprehension will suffer if the readability level is not accessible to the students. Readability formulas predict text complexity, assisting in appropriate text selection that complements students' reading abilities to improve their language development. Therefore, the study aims to find out the reading index of the prose forms in the literature component catered to lower secondary school students ages 13 and 14 years old in Form One (seventh grade) and Form Two (eighth grade) classrooms in Malaysia. The reading index is measured by using four readability formulas which are Dale-Chall, Fog, SMOG, and Flesch-Kincaid that focuses on the words, sentences, syllables, and polysyllable words. These formulas are used to predict the level of difficulty of the prose forms. The reading index calculated from these readability formulas reveals the grade level of the prose forms. The grade level indicates the best age for reading and understanding the prose forms. Two prose forms were chosen as samples in the study. A passage is chosen from each prose form to be uploaded using the online tool. The indices obtained from the readability formulas

* Corresponding author, email: revathi@fbk.upsi.edu.my

Citation in APA style: Gopal, R., Maniam, M., Madzlan, N. A., Shukor, S. S., & Neelamegam, K. (2021). Readability formulas: An analysis into reading index of prose forms. *Studies in English Language and Education*, 8(3), 972-985.

Received March 15, 2021; Revised June 28, 2021; Accepted August 15, 2021; Published Online September 16, 2021

<https://doi.org/10.24815/siele.v8i3.20373>

predicted that both of the prose forms were below students' reading age. The study implicates reading index must be taken into consideration in literary texts selection because it is an indicator of the years of education that an individual requires to comprehend the literary text clearly. Suitable reading material at students' age level can enhance literature learning and teaching in the ESL classroom.

Keywords: ESL classroom, readability formulas, readability index, text difficulty.

1. INTRODUCTION

Literature teaches readers about the unity of emotion and real-life across the world (Vučković, 2017). Studying these works can help students better understand intercultural and critical cultural issues (Byram, 2014) and foster national identity. In Singapore, literature acts as a unifying agent amongst its people (Dass et al., 2012). In Malaysia, the literature component is included in the teaching practices of English language syllables. The integration of the literature component has lifted students' awareness of language use, linguistics benefits, and the development of interpretative skills. Poems and prose forms (short stories) are two genres compulsory for Form One (seventh grade) and Two (eighth grade) students. Students' comprehension of prose forms can be difficult if the readability level is not within their reading ability. Readability takes into account factors such as text, readers, and the context. These three factors must be aligned to students' interests so that they will be motivated and engaged in reading can take place. When choosing reading materials for our students, these three factors play a significant role. However, there are readability formulas in judging how difficult or easy a reading text can be. The formulas give out a reading index; the number is meant to predict how easily students with different literacy levels would be able to read and understand the reading text.

Several studies on different aspects of readability have been conducted in Malaysia and abroad. In Malaysia, some of the local researchers are Nor (1997), Abdullah and Hashim (2007), and Janan (2011) with their studies on readability encompass not just education but also enterprise and industry. Abdullah and Hashim (2007) examined the readability of a variety of Malaysian English works. Janan (2011) developed a new readability model that emphasizes the psycholinguistics of readers and reading processes. Hamid et al. (2020) studied the readability of Covid-19 information by the Malaysian Ministry of health. Crossley et al. (2019) looked at international research and performed a report on going beyond traditional readability formulas. This research created new readability formulas for both text comprehension and reading speed based on advanced NLP tools.

Scholars from around the world have debated various aspects of readability. Brown et al. (2012), investigated the degree of correlation between each of the L1 readability estimates and the achievement of Russian EFL students on exact cloze passages and discovered that a variety of L1 readability indexes for the set of 50 passages are slightly correlated with the average performances of Russian university students on cloze versions of those same passages. Besides, Tabatabaei and Bagheri (2013) found that the readability formula produces good results when used to calculate

the readability of reading passages in various textbooks published by the Iranian Company as learning materials for pre-university students. [Rahmawati and Lestari \(2014\)](#) claimed that readability could be used to calculate the readability of reading passages in high school reading materials for both government officials and commercial tenth graders.

In this study, the reading index of the prose forms currently used for lower secondary students in Form One and Two classrooms in the literature component is determined. It is intended to determine whether the indices are compatible with the age of the students who would read and understand the prose forms prescribed for them. It is necessary to assist students out there in selecting appropriate reading materials based on their age, language proficiency, and interest in reading in order to participate in reading and promote language learning. Hence the gap of the study is addressed. The research determines the readability levels of two prose forms and addresses the research questions:

- What is the readability level of the prose form ‘Fairs Fair’ by Narinder Dhani?
- What is the readability level of the prose form ‘Cheat’ by Alan Baillie?

2. LITERATURE REVIEW

2.1 Readability

Scholars have proposed various definitions of readability. Each researcher defines the readability according to their understanding and definitions, which rely on the purpose and the need for their studies during this period. The earliest definition of the term readability was given by [Dale and Chall \(1949\)](#) and they stated that readability was “the total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers has with it” ([DuBay, 2004, p. 3](#)). The emphasis here is on the text and the reader, whereby the reader’s success in reading depends on the degree to which they understand the text and reading it at optimum speed. [Klare \(1963\)](#) defined readability as the ease with which a piece of writing may be understood or comprehended. This definition focuses solely on the writing style as a text element. Likewise, [Ismail et al. \(2016\)](#) posited readability determines whether a book is simple or complex to comprehend. Readability, according to the experts, is only concerned with the linguistic aspects of a document for example word count, word concreteness, syntax, and coherent devices. Another readability scholar, [McLaughlin \(1969\)](#), defined readability as “the extent to which a class of people finds certain texts understandable and comprehensible” ([DuBay, 2004, p. 3](#))

[Srisunakrua and Chumworatayee \(2019\)](#) state that readability focused on making sure that a reading material matches the target readers’ proficiency. Material with a high readability level is considered to be difficult to read, whereas a material with a low readability level is said to be simple to read. [Kassim \(2010\)](#) held a similar viewpoint, stating that readability is a measurement of the ease or difficulty of reading materials. The reader’s ability to read, understand, and analyze any reading material demonstrates that the reader has mastered readability and legibility aspects.

[Rosenblatt \(1978\)](#) claimed that the reader, the text, and the setting also contributed towards readability. The setting refers to cultural and social factors. They

are significant factors in structuring meaning (Chin, 1994). Pikulski (2002), another readability scholar, proposed that readability is determined not only by the interaction between the text and the reader but also by the purpose of reading. This viewpoint is similar to Rosenblatt (1978), but Pikulski (2002) included the term context in which he addressed it as the purpose of reading, and Janan (2011) shared the same view as the above scholars. She claims that readability is a dynamic interactive process involving text-reader factors and is bounded by certain contexts.

The definitions of readability have changed over the years. Firstly, Dale and Chall (1949) focused on understanding, fluency, and interest. Secondly, Klare (1963) focuses on the clarity of the text's written style that helps to understand words and phrases. During the years, scholars stressed the understanding of the text by which the difficulty of the text is placed within the text itself, not within the reader. (Janan, 2011).

The matching aspects between the text's characteristics and the reader's characteristics were prominent during the 1970s. From this period onwards, the definitions of readability were more of an interactive nature. During these years, the term 'context' was included in the definitions by readability scholars, for example, Goodman (1969), Rosenblatt (1978), Pikulski (2002), and Janan (2011). Nonetheless, what constitutes the context differed among the scholars.

It can be noted that within most definitions, three elements are prominent, which are the text, the reader, and the context in reading to understand a text. Readability is essential in the education field because it judges how readable a text is and the success of a reading process. A loophole between the reader and the text appears to mean the text is overlooked (Janan, 2011). To avoid the mismatch, educators depend on readability formulas to assess if the text is readable by the intended audience (Zamanian & Heydari, 2012).

2.2 Readability Measurement

Readability is seen as a measure of a text's accessibility that shows how well it reaches the target audience (Pikulski, 2002). The measure takes the form of a formula. Because they are objective, readability formulas are utilized for a wide range of literature. They do not take into account many outside influences, and they have been frequently employed by textbook publishers of the twentieth century (Oakland & Lane, 2004). The goal of readability evaluation is to ensure what grade level a reader needs to read and comprehend a text. Readability formulas rely on linguistic criteria for the prediction of text difficulty. The syntactic difficulty is based on the number of words per sentence, as well as the sentence's familiarity and syntactic structures. Sentences with more words and complicated structures are more confusing (McNamara et al., 2004). And thus, it is essential to investigate the readability level of the prose forms prescribed for the lower secondary school students, Form One and Two by analyzing the sentence structures and vocabulary difficulty, and finding out the level of difficulty of each prose so that it would not hinder text comprehension.

Readability formulas are created to predict text difficulty. The ability to predict text difficulty helps those in the teaching field to select appropriate texts for students. Following are details on the reading measurements that will be used in the present study, which are Fog (Gunning, 1952), Simple Measure of Gobbledygook or SMOG (McLaughlin, 1969), Flesch-Kincaid (Kincaid et al., 1975), and Dale-Chall (Dale &

Chall, 1995). The following are the readability formulas used in this study to identify the reading index for the prose forms.

2.2.1 Dale-Chall Readability Formula

Dale and Chall (1949) devised an early formula that included variables such as the percentage of unfamiliar terms in a sentence and the average number of words per sentence. In the United States, reading grade level is calculated with the Dale-formula Chall's formula (Harrison, 1984):

$$(0.1579 \times \text{Percent UFMWDS}) + (0.0496 \times \text{WDS/SEN}) + 3.6365$$

UFMWDS = unfamiliar words

WDS/SEN stands for "words per sentence average"

The New Dale-Chall readability formula (1995) will be used in this study to determine word difficulty by counting difficult words. The text's grade level is determined by the length of sentences and the number of difficult words. The New Dale-Chall formula (1995), which is based on the use of unfamiliar words, is an accurate readability formula.

2.2.2 Fog Readability Formula

The FOG formula of Gunning (1952) was the easiest to work out the reading index, and this fact shows its popularity. The formula consists of the following variables:

- 1) Average word count per sentence.
- 2) the polysyllabic word percentage.

The FOG formula (Harrison, 1984) is as follows:

$$\text{Grade level of Fog reading in the United States} = 0.4 \text{ times } (\text{WDS/SEN} + \text{percent PSW})$$

WDS/SEN stands for "words per sentence average"

PSW stands for polysyllabic word proportion

Robert Gunning came up with this formula in 1952. For every 100 words, it utilizes the average sentence length and the number of words with more than two syllables. The Fog index indicates how many years of schooling a reader will need to comprehend a paragraph or text.

2.2.3 SMOG Readability Formula

SMOG reading grade level is:

$$US = 3 + \sqrt{P} \text{ (3 plus the square root of P)}$$

P = the nearest perfect square to the number of polysyllabic words in thirty sentences.

The SMOG Readability Formula (McLaughlin, 1969) is a simple approach for determining a written material's reading level. A single variable and the amount of polysyllabic (three or more syllable) words in 30 sentences made up the SMOG formula. The following is the SMOG formula (Harrison, 1984):

$$US = 3 + P \text{ (3 plus the square root of P)}$$

P = the nearest perfect square to the number of polysyllabic words in thirty sentences
SMOG reading grade level

The SMOG readability formula (DuBay, 2004) is as follows:

1. At the start of your text, count ten sentences in a row. In the middle, count ten sentences. Near the conclusion, count ten sentences. There are 30 sentences total.
2. In each set of sentences, count every word with three or more syllables, even if the same word appears more than once.
3. Total the number of words you've counted. To determine the grade level of the reading content, use the SMOG formula as described above.

The Simple Measure of Gobbledygook SMOG Readability Formula was devised by G. Harry McLaughlin in 1969, and the calculations provide an approximation of the text's readability level (Doak & Doak, 2010). The abbreviation SMOG stands for Simple Measure of Gobbledygook. The length of the word and the length of the phrase are taken into consideration in the computation (DuBay, 2004). The SMOG formula's results approximate the level at which students have progressed if they can understand 90-100 percent of what is being read. In other words, most of the content will be understood if students read at or above grade level (Bailin & Grafstein, 2016).

2.2.4 Flesch-Kincaid Readability Formula

Three factors were incorporated in the Flesch-Kincaid formula (Kincaid et al., 1975, as stated in DuBay, 2004, p. 50):

1. The length of the words
2. The number of the unknown word
3. The length of a sentence

The Flesch-formula Kincaid's is:

$$=.39 wl + 11.8 sl - 5.59 \text{ (Kincaid et al., 1975, as cited in DuBay, 2004, p. 50).}$$

wl stands for word length.

sl stands for sentence length.

The Flesch-Kincaid readability formula determined the reader's grade level. It takes into account two factors: sentence length (words per sentence), which indicates syntactic complexity, and average syllables per word, which represents word difficulty (DuBay, 2004).

2.3 Formula's Ratings: Flesch's Reading Ease Scores and Dale-Chall's (1995) Scores, and Grade Levels

A readability formula is a mathematical calculation that predicts how easy a book will be to read. The number of years of schooling required to comprehend a text is estimated (Kondru, 2006). The semantic factor (word difficulty) and the syntactic factor are used to calculate readability formulas (difficulty of sentences). The readability formulas were used to determine the level of readability of school textbooks (Heydari & Riazi, 2012). It aided researchers in their tests by allowing them to alter the difficulty of passes (DuBay, 2004). The Flesch reading ease formula is commonly used in readability formulas. The Flesch reading ease formula is a straightforward method for determining the reader's grade level and the difficulty level of an English-language school text. It necessitates counting the number of words per phrase (in a sample of 100 words) and the number of syllables in one hundred words. It assigns a 100-point scale to texts, with the greater the score, the easier the material is to comprehend (DuBay, 2004). Table 1 shows Flesch's reading ease scores.

Table 1. Flesch's reading ease scores (adapted from [DuBay, 2004](#)).

Score	Ease of readability	Estimated reading grades
90-100	Very Easy	5 th grade
80-89	Easy	6 th grade
70-79	Fairly Easy	7 th grade
60-69	Standard	8 th and 9 th grade
50-59	Fairly Difficult	10 th to 12 th grade
30-49	Difficult	13 th to 16 th
0-29	Very Confusing	college graduate

Flesch's rating chart shows the scores from 0 to 100, the ease of readability from very confusing to very easy. The lesser the scores the more difficult is the reading text for the students. Next, Edgar Dale and Jeanne Chall were inspired by Rudolph Flesch's Reading Ease Formula to develop the Dale-Chall formula for adults and children over the age of four to address some of the flaws in the Flesch reading ease formula. The Dale-Chall formula was originally published in 1948 and updated in 1995 in *Readability Revisited: The New Dale-Chall readability formula* (Dale & Chall, 1995). This formula used word length to assess word difficulty and uses a count of hard words (DuBay, 2004). Table 2 shows Dale and Chall's (1995) formula's ratings according to scores and grade level. The Dale-Chall's (1995) formula shows that reading index scores from 10 above.

Table 2. Dale-Chall's (1995) scores and grade levels (adapted from [DuBay, 2004](#)).

Scores	Grade levels
4.9 and below	Grade 4 and below
5.0 to 5.9	Grade 5-6
6.0 to 6.9	Grade 7-8
7.0 to 7.9	Grade 9-10
8.0 to 8.9	Grade 11-12
9.0 to 9.9	Grade 13-15 (college)
10 and above	Grade 16 and above (college graduate)

3. METHODS

The study's design is based on the quantitative approach. Four readability formulas are used as instruments to find out the reading index for the literary texts 'Fair's Fair' by Narinder Dhami and 'The Cheat' by Allain Blaille. The readability formulas are Fog, SMOG, Flesch-Kincaid's readability formula, and Dale-Chall (1995).

3.1 Predictor Variables in Readability Formulas

Readability formulas were used to predict the readability index of the prose forms used in the lower secondary. Readability formulas used are Dale-Chall (1948), Fog, SMOG, and Flesch-Kincaid. Common variables used in these formulas are sentence length and difficult words. Table 3 below shows the predictor variables used in the formulas.

Table 3. Predictor variables used in the four readability formulas.

Formulas	Predictor variables		
	Words	Sentence length	Polysyllable words
SMOG (1969)			
Fog (1952)			
Flesch-Kincaid (1975)			
Dale-Chall (1948)			

The SMOG and Fog readability formulas use polysyllable words, Flesch-Kincaid uses words, sentence length, and difficult words, whereas Dale-Chall uses sentence length and difficult words in judging the reading index for reading materials.

3.2 Procedures in Obtaining Reading Index

The texts used to calculate the readability indices are selected randomly from the prose forms for Form One, a full text on chapter three (pp. 51-52) from the short story 'Fair's Fair' written by Narinder Dhami and Form Two, two pages (pp. 102-103) from the short story, 'Cheat' written by Alan Baillie. The paragraphs chosen were between 300 and 700 words long, which is the suggested length for the computerized reading assessments utilized in this study (Owu-Ewie, 2014). These texts were typed on Microsoft Word and the four formulas were applied at one go using a computer program by Alain Trottier called Words to count (<http://www.wordscount.info/>). The indices obtained from different readability formulas predicted whether the text is easy or difficult.

The following screen shows an example of counting the readability indices of one of the texts in this study. The procedure of setting the text is shown on the first screen. Figure 1 shows the words count website, uploading the text. Furthermore, the complexity level of the prose forms could be calculated manually, in addition to using the words count website, as described in the literature review section.

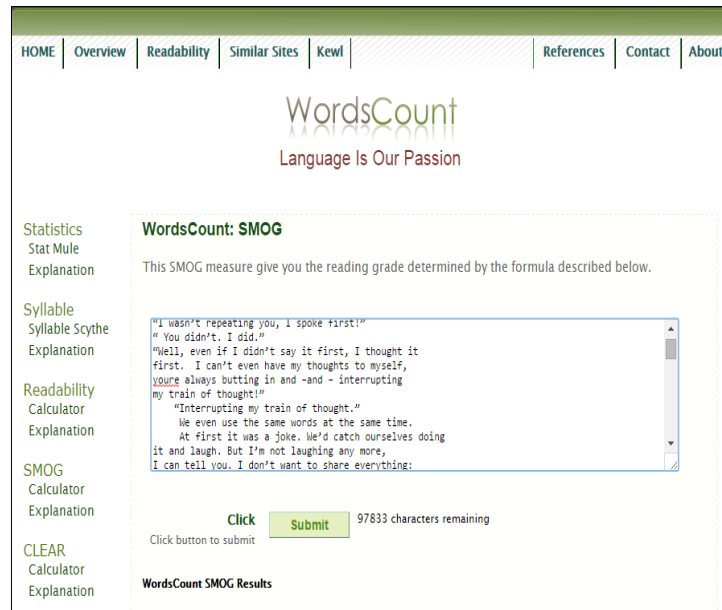


Figure 1. Words count website.

The results of the readability indices are displayed on the next screen.

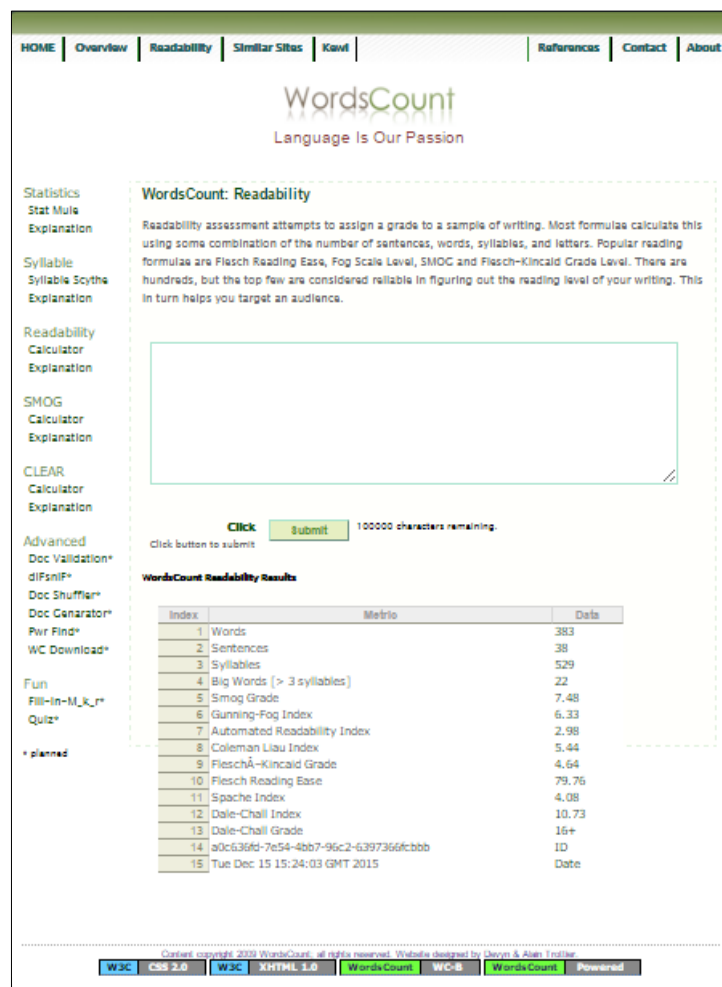


Figure 2. The result of the readability indices.

4. RESULTS

The results on predictor variables show that there are 307 words, 47 sentences, and 337 syllables whereas in the Form One students' prose from 'Fair's Fair'. As for the Form Two students' prose from 'Cheat', there are 569 words, 48 sentences, 708 syllables, and 34 polysyllable words. The indices of the readability formulas on the text 'Fair's Fair' show 0.0 on SMOG formula, Fog 2.61, Flesch-Kincaid -0.09, and Dale-Chall 9.52. The readability score on the text 'The Cheat' shows 6.6 on SMOG formula, Fog 5.1, Flesch-Kincaid 4.5, and Dale-Chall 6.6.

The readability index scores obtained from each readability measurements for each prose form were graded according to their level of difficulty derived from US grade level because it is the grade level, a common procedure used in judging the readability of texts. Based on the indices, they clearly show that the scores are inconsistent with the formulas. Therefore, it is suggested by Davison and Bolt (1981) that an average of the indices is calculated to determine a reliable readability index. It is calculated by adding the indices on the four formulas and the total is divided by four. It is because there are four readability formulas used in this study to judge the index score. The average readability index calculated for chapter three from the prose 'Fair's Fair' is 3.01 which, falls at grade three. The average of the readability index for the Form Two prose indicates 5.7 points at fifth grade and is appropriate for pupils at the age of ten.

The indices of both literary texts fall under a low reading index. It can imply that the literary texts are simple to comprehend. The American grade level shows the target age group for these literary texts 'Fair's Fair' and 'Cheat' are for lower primary pupils aged eight years old, and upper primary pupils aged ten years old. Converting these grades to the Malaysian context, these literary texts are intended to be read and comprehended by lower secondary students, ages 13 and 14. There is a significant age difference upon comparing with the American grade level. The significant difference in grade level does not imply that Malaysian students are at the elementary level in the English language; instead, the literary texts assigned are tailored to the students' various abilities taking into account different individuals who have different schemas and prior experience when reading these literary texts. Indeed, students at the advanced and intermediate levels of language proficiency would, without a doubt, regard these texts as simple, but not weak students. The weak students need assistance and guidance from their teachers in exploiting the literary text. The teaching practices that are appropriate for the student's level of language proficiency, understanding, and experience can aid them in comprehending and analyzing literary texts in the ESL classroom. Indeed, the text selected must be appropriate in any way to meet their needs.

5. DISCUSSION

Readability is a measure of the ease or difficulty of text (DuBay, 2004). We can calculate the complexity of words and sentences in any given content using readability formulas. Since readability formulas differ depending on the intent and measurement, most authors use no more than three identical formulas to assess a document's readability. In this study, the researcher used four formulas to get the reading index, and the findings of the study showed there was no association between the formulas in

terms of reading indices. Hence, it was necessary to find an average within the readability formulas. The reading indices for both literary texts were low, which did not correspond to the age group of Malaysian students. The target age group for reading and comprehending the literary texts 'Fair's Fair' and 'Cheat' is between the age of eight and ten. Both of these literary texts have low reading indices, suggesting that they are easy for students in lower secondary school. The use of these readability formulas has the advantage of providing a solid figure on the text difficulty. This information aids educators in text selection and pedagogical practices for getting the best out of literary texts to reach out to their students.

However, when it comes to determining how well the target audience understands the text, readability formulas are not very useful (Zamanian & Heydari, 2012). Readability formulas only take into account certain aspects of a text, such as vocabulary and a sentence length. These two variables are only accurate and correlate with text complexity, and this has been acknowledged (Harrison, 1980). Klare (1976) reviewed 36 experimental studies on the ability to predict the comprehension of readability variables and established that 19 studies were statistically significant, 11 were not statistically significant, and six were mixed (had some significance). The findings cannot be used to advocate the use of readability formulas. These results indicate that readability formulas cannot be used with much confidence about their success in predicting reading comprehension. Bailin and Grafstein (2001) reexamined the linguistic criteria that constituted the basis of readability ratings, arguing that the criteria frequently utilized in readability formulas do not provide a sufficient basis for judging reading difficulties.

Additionally, the readability formulas cannot determine the literary value of a book (Abdullah & Hashim, 2007). However, this does not negate the importance of critically evaluating literary texts before recommending them to a certain group of readers. According to Schulz (1981), careful selection of literary books based on their linguistic difficulties can help to avoid "frustrated" reading in a foreign language while also increasing comprehension, appreciation, and love of literature. The appeal of utilizing formulas to assess readability is from the notion that they can objectively and quantitatively assess the difficulty of written material without taking into account the characteristics of readers. A readability formula can also produce a numerical score, giving the user the impression of knowing just how tough a text is (Bailin & Grafstein, 2001). Concerning the basic concepts of formulae, DuBay (2004) stated, "The variables utilized in the readability formulas show us the skeleton of a text" (p. 61). He insisted that it is up to us to add tone, content, organization, coherence, and design to that skeleton. Readability formulae were created with the intention of assisting educators in the selection of acceptable texts for children of various ages by ranking school textbooks in terms of difficulty (Hargis, 2000). Hence, teaching and learning practices can be devised in ways to suit students' level of language ability, interest, prior knowledge, and cultural background to engage them throughout literature lessons.

6. CONCLUSIONS

The readability index judged the Form One and Two prose forms easy. It portrayed that there was no match between the reading index calculated and students'

reading age on the two prose forms. The readability indices for the two prose forms indicated a low reading level as with students' reading age.

According to Carrell (1987), lower readability levels of reading materials allow lower reading abilities of students to understand a text easily. Thus, in the Malaysian context, a low readability index text would cater to all students of diverse abilities in the second language and background. Lower ability students could read and understand the text easily, and for students with higher proficiency, this would be an advantage because their degree of performance would be higher with low reading index text.

There are many methods in the 21st century in judging the reading index of reading materials. However, the study used the traditional readability measurement to find out the reading index of the prose forms prescribed in the literature component for the lower secondary students. This is because the process of finding out the reading index using the Allain Trottier website was not complex. Additionally, readability formulas cannot measure everything that contributes to how readable a text is for a student. They cannot reflect the interaction between text-reader factors in a reading activity. Perhaps, future research on the area of readability can be on new ways of assessing text difficulty that reflects the interaction between text-reader factors.

REFERENCES

- Abdullah, I. H., & Hashim, R. S. (2007). Readability analysis of Malaysian short stories in English. *Jurnal e-Bangi*, 2(2), 1-11.
- Bailin, A., & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Language & Communication*, 21(3), 285-301. [https://doi.org/10.1016/S0271-5309\(01\)00005-2](https://doi.org/10.1016/S0271-5309(01)00005-2)
- Bailin, A., & Grafstein, A. (2016). *Readability: Text and content*. Macmillan.
- Brown, J. D., Janssen, G., Trace, J., & Liudmila, K. (2012). A preliminary study of cloze procedure as a tool for estimating English readability for Russian students. *Second Language Studies*, 31(1), 1-22.
- Byram, M. (2014). Twenty-five years on-from cultural studies to intercultural citizenship. *Language Culture and Curriculum*, 27(3), 209-225. <https://doi.org/10.1080/07908318.2014.974329>
- Carrell, P. (1987). Readability in ESL. *Reading in a Foreign Language*, 4, 21-40.
- Chin, E. (1994). Redefining "context" in research on writing. *Written Communication*, 11, 445-481. <https://doi.org/10.1177/0741088394011004002>
- Crossley, S. A., Skalicky, S., & Dascalu, M. (2019). Moving beyond classic readability formulas: New methods and new models. *Journal of Research in Reading*, 42(3-4), 541- 561. <https://doi.org/10.1111/1467-9817.12283>
- Dale, E. & Chall, J. S. (1949). The concept of readability. *Elementary English XXVI*(1), 19-26.
- Dale, E. & Chall, J. S. (1995). *Readability revisited: The new Dale-Chall readability formula*. Cambridge.
- Dass, R., Chapman, A. & O'Neill, M. (2012). Literature in English: How students in Singapore secondary schools deal with the subject. *Malaysian Journal of ELT Research*, 8(2), 1-20.

- Davison, A. & Bolt, B. (1981). *Readability: Appraising text difficulty* [Reading education report No. 24]. <https://files.eric.ed.gov/fulltext/ED208340.pdf>
- Doak, L. G., & Doak, C. C. (2010). Writing for readers with a wide range of reading skills. *American Medical Writers Association Journal*, 25(4), 149-154.
- DuBay, W. H. (2004). *The principles of readability*. Impact Information.
- Goodman, K. (1969). Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly*, 5(1), 9-30.
- Gunning, R. (1952). *The technique of clear writing*. McGraw-Hill.
- Hamid, A. R. A., Hashim, M. F., Hasan, N. A., & Azhan, N. M. (2020) Readability of COVID-19 information by the Malaysian Ministry of Health. *Journal of Nusantara Studies*, 5(2), 170-191. <https://doi.org/10.24200/jonus.vol5iss2pp170-191>
- Hargis, G. (2000). Readability and computer documentation. *ACM Journal of Computer Documentation*, 24(3), 122-131. <https://doi.org/10.1145/344599.344634>
- Harrison, C. (1980). *Readability in the classroom*. Cambridge University Press.
- Harrison, C. (1984). Readability in the United Kingdom. *Journal of Reading*, 29(6), 521-529.
- Heydari, P., & Riazi, M. A. (2012). Readability of texts: Human evaluation versus computer index. *Mediterranean Journal of Social Sciences*, 3(1), 177-190.
- Ismail, A., Yusof, N. & Yunus, K. (2016). The readability of Malaysian English children's books: A multilevel analysis. *International Journal of Applied Linguistics & English Literature*, 5(6), 214. <http://dx.doi.org/10.7575/aiac.ijalel.v.5n.6p.214>
- Janan, D. (2011). *Towards a new model of readability* [Unpublished doctoral dissertation]. University of Warwick.
- Kassim, R. (2010). Literacy through children's literature: A Malaysian perspective. *Malaysian Journal of Media Studies*, 12(2), 1-9.
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count, and Flesch reading ease formula) for Navy enlisted personnel* [Research Branch Report 8-75, Chief of Naval Technical Training]. Naval Air Station Memphis.
- Klare, G. R. (1963). *The measurement of readability*. Iowa State University Press.
- Klare, G. R. (1976). A second look at the validity of readability formulas. *Journal of Reading Behavior*, 8(2), 129-152. <https://doi.org/10.1080/10862967609547171>
- Kondru, J. (2006). *Using part of speech structure of text in the prediction of its readability* [Unpublished master's thesis]. University of Texas.
- McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading*, 22, 639-664.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M. & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments and Computers*, 36(2), 193-202. <https://doi.org/10.3758/BF03195564>
- Nor, A. K. M. (1997). *Understandability of chairman's address in annual reports: A Malaysian context* [Unpublished master's thesis]. Universiti Teknologi MARA.
- Oakland, T., & Lane, H. B. (2004). Language, reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing*, 4(3), 239-252. https://doi.org/10.1207/s15327574ijt0403_3

- Owu-Ewie, C. (2014). Readability of comprehension passages in Junior High School (JHS) English textbooks in Ghana. *Ghana Journal of Linguistics*, 3(2), 35-68. <https://www.doi.org/10.4314/gjl.v3i2.3>
- Pikulski, J. (2002). *Readability: A definition*. Houghton Mifflin Company.
- Rahmawati, Y. I., & Lestari, L. A. (2014). *The readability of reading texts in the English textbooks used by tenth grade students* [Unpublished bachelor's thesis]. State University of Surabaya.
- Rosenblatt, L. (1978). *The reader, the text and the poem*. Southern Illinois University Press.
- Schultz, R. A. (1981). Literature and readability: Bridging the gap in foreign language reading. *The Modern Language Journal*, 65(1), 43-53. <https://doi.org/10.1111/j.1540-4781.1981.tb00952.x>
- Srisunakrua, T. & Chumworatayee, T. (2019). Readability of reading passages in English textbooks and the Thai National Education English test: A comparative study. *Arab World English Journal*, 10 (2), 257-269. <https://dx.doi.org/10.2139/ssrn.3418124>
- Tabatabaei, E., & Bagheri, M. S. (2013). Readability of reading comprehension texts in Iranian senior high schools regarding students' background knowledge and interest. *Journal of Language Teaching and Research*, 4(5), 1028-1035. <https://doi.org/10.4304/jltr.4.5.1028-1035>
- Vučković, D. (2017, March 30-31). *Literature as a core of integrated curriculum* [Paper presentation]. The 5th International Conference, Kolegji Universitar Bedër, Tirana, Albania.
- Zamanian, M., & Heydari, P. (2012). Readability of texts: State of the art. *Theory and Practice in Language Studies*, 2(1), 43-53.