# READABILITY AND ITS EFFECTS ON STANDARDIZED ASSESSMENT OUTCOMES

**April Ward; Mara Charles; Amanda Barlow; Crystal Randolph, PhD, CCC-SLP; Corine Myers-Jennings, PhD, CCC-SLP; Ruth Renee Hannibal, PhD, CCC-SLP**

**Valdosta State University**

## Introduction and Background

Readability is the ease at which written text can be understood by the reader. Readability is especially important when creating standardized assessments and screenings that are used to classify students that may have reading difficulties and/or identify students who may be at risk for developing reading disorders. Research completed on readability concludes it is an important factor to consider when analyzing standardized test scores(Homan & Hewitt, 2004). Readability is one of many factors that may influence the validity of an assessment. Consequently, it is important that the readability of such assessments is calculated prior to administration.

In a recent article discussing the reform of reading achievement tests, Schutz (2013) analyzed the reading passages from reading achievement tests and found that most were not written on the appropriate grade-level. To offset, the variability in grade-level appropriateness, test creators adjust the complexity of comprehension questions, which often accompany reading passages. When the readability of passages on standardized measurements do not match the intended grade level, results cannot be considered accurate. Students' knowledge of a particular content may be misinterpreted, which could cause them to be retained in their current grade level or they may receive an incorrect diagnosis concerning reading ability. Because speech-language pathologists often use standardized measurements to diagnose literacy deficits, it is important to ascertain if the readability of these measurements are valid and reliable.

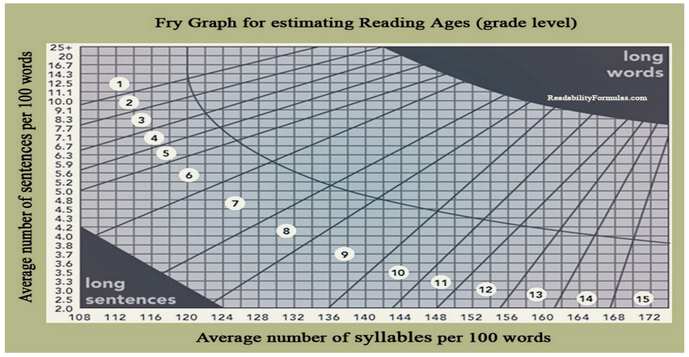
In a previous article, Klare (1974-1975) discusses a number of formulas for calculating the ease at which a reader can understand a text or passage. Assessing readability can be completed by using one of several formulas including Fry, SMOG Index, and Flesch-Kincaid among others. The varying methods used for each formula explains the differences in readability scores when calculated. For example, the Fry formula uses a combination of the total number of sentences and syllables in a 100-word sample to find the resultant grade level. The SMOG Index accounts for polysyllables in sentences; whereas Flesch-Kincaid formulas accounts for the number of words in a sentence and the number of syllables in a word.

The purpose of this study is to evaluate the readability level of passages on three standardized measurements: the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), the Gray Oral Reading Test (GORT), and the Georgia Milestones Assessment System (GMAS). Three readability tools will be used to measure readability levels: Microsoft Word, an online readability calculator, and calculation by hand using a readability formula.


## Methodology

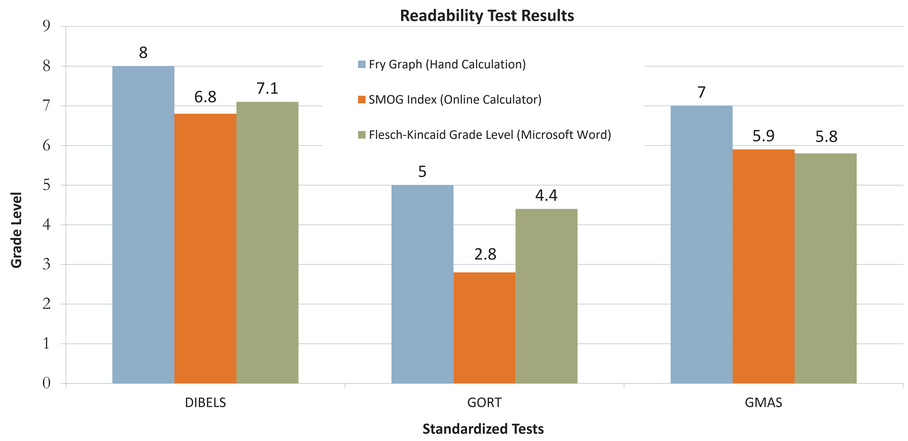
To determine readability levels, three calculation methods were used on the DIBELS, GORT, and GMAS reading passages. The researchers completed the calculations by hand, using an online readability calculator, and Microsoft Word. The Fry formula was used for the hand calculation. Each passage was divided into three sections of 100 words. If the one-hundredth word fell in the middle of a sentence, the next segment would begin with a new sentence. The number of sentences and syllables within each section was counted and the totals were recorded. When calculating the total amount of sentences in a 100 word section, the last sentence was estimated to the nearest 1/10th. The average number of sentences and syllables were gathered and then used to find the corresponding readability level using the Fry Graph for Estimating Reading Ages.

The researchers conducted the online readability calculation by inserting the reading passages from each assessment onto www.readabilty-score.com. The website reported several sets of results using various readability formulas; however, the researchers used the numbers from the SMOG Index. Finally, the researchers used Microsoft Word to calculate readability. The researchers then copied each passage onto a word document and then clicked on the *"File"* tab → *"Options"* → *"Proofing"* . Under the heading, *"When correcting spelling and grammar in Word,"* the researchers checked the box next to *"Show readability statistics."* Next, each researcher clicked on the *"Review"* tab → *"Spelling & Grammar."* The results appeared in a new window, and included the Flesch-Kincaid Grade Level. To ensure reliability of readability data, two of the three researchers completed readability calculations for each reading measure. Comparison of the calculations for each reading measure revealed 100% agreement.




## Results

The readability results for the DIBELS were Grade 8 on the Fry Graph, 6.8 according to the SMOG Index, and 7.1 as the Flesch-Kincaid Grade Level. The readability results for the GORT were Grade 5 on the Fry Graph, 2.8 according to the SMOG Index, and 4.4 as the Flesch-Kincaid Grade Level. The readability results for the GMAS were Grade 7 on the Fry Graph, 5.9 according to the SMOG Index, and 5.8 as the Flesch-Kincaid Grade Level.




## Discussion

The data gathered from this research study proves that there is a discrepancy in the readability level on the previously mentioned tests. Prior to calculating the readability levels of the DIBELS, GORT, and GMAS, our assumption was that each assessment's reading passage would be at the appropriate reading level for a student in grade five. All three of the readability formulas indicated that the reading passages on the DIBELS were above a fifth grade reading level. The passages from the GMAS were calculated to be appropriate for a fifth grade student according to the SMOG Index and Flesch-Kincaid formulas, but too high based on the Fry formula. Contrary to the appropriate reading level score produced by the Fry Graph, the results from the SMOG Index and the Flesch-Kincaid formula indicated that the GORT passages were below a fifth grade reading level. The Fry Graph consistently yielded higher readability scores for each assessment when compared to the SMOG Index and Flesch-Kincaid Formula.

## Implications

A readability level that is too high on tests such as the GMAS, could lead to a student failing a high stakes test and cause them to be retained. A readability level that is too low may result in high test results. Consequently, the student could potentially be placed in a class too difficult for his academic level causing him to struggle. Implications for the DIBELS, a universal screening and progress monitoring tool used at the elementary level, suggest that high readability levels could place students in at risk categories when they may have grade appropriate reading levels. Low readability level for standardized tests such as the GORT may result in under identification of students with reading difficulties.

When drafting reading passages for standardized assessments, test developers should ensure that the readability levels are grade-appropriate prior to publication. It is imperative for high stakes tests to be as valid and reliable as possible because of their use as a measure of student performance in the classroom (Hewitt & Homan, 2004).

In a future study, it would be beneficial to examine student performance on these assessments in addition to calculating readability level. By factoring in student performance, we would have a more accurate representation of the effects of readability on test scores.

## References

Hewitt, M. A. & Homan, S.P. (2003). Readability level of standardized test items and student performance: The forgotten validity variable. *Reading Research and Instruction*, 43 (2), 1-16.

Klare, G.R. (1974-1975). Assessing readability. *Reading Research Quarterly, 10 (1)*, 62-102.

Schutz, D. (2013). Toward educational testing reform: Inside reading achievement tests. Educational Policy Analysis Archives, 21 (90), 1-32.