

Readability Levels of Health-Based Websites: From Content to Comprehension

Mary Schutten, P.E.D¹; Allison McFarland, PhD, M.B.A²

Author¹ is affiliated with CLAS Dean's Office at Grand Valley State University. Author² is affiliated with Bethel College. **Contact author:** Mary Schutten, Grand Valley State University, 4-232 MAK, 1 Campus Dr, Allendale, MI 49401, USA. **Phone:** 616 331 2495; **Fax:** 616 331 3675; **Email:** schutmar@gvsu.edu.

Submitted March 16, 2009; Revised and Accepted June 18, 2009

Abstract

Three of the national health education standards include decision-making, accessing information and analyzing influences. WebQuests are a popular inquiry-oriented method used by secondary teachers to help students achieve these content standards. While WebQuests support higher level thinking skills, the readability level of the information on the website is often overlooked. Readability refers to the ease with which a text can be read and understood. If an individual's reading skill level is significantly below that of the readability level of the document, then it can be reasonably assumed that the individual was not able to fully comprehend what s/he read. This can adversely impact the health literacy and, ultimately, the education of the child. A recent review of 105 randomly chosen health-related website pages representing the six risk factors for youth was performed. The analysis yielded an average reading level of grade 11 and low correlation between target audience and readability grade level. This grade level far surpasses the reported grade level reading of the average adult in the United States and most secondary students. The first line of action for health educators is to determine the readability characteristics of WebQuest materials used by their students and implement suggestions to enhance reader comprehension.

Key Words: *Readability, Health Literacy, Patient Education, WebQuests.*

Introduction

In a position statement issued by the American Association for Health Education (AAHE) related to health literacy in 2008, health literacy was defined as the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions.¹ The vital importance of developing health literacy to the overall quality of life nation is further supported by some of the *Healthy People 2010* objectives, specifically, objective 11-2 as referenced by AAHE.¹ In addition, National Health Education Standards consist of seven skills and one knowledge component that are designed to create a health literate child.¹ Three of these standards are decision-making, accessing information and analyzing influences; skills that are essential to the development of health literacy.

One key factor that impacts health literacy is readability. Readability refers to the ease with which a text can be read and understood. If an individual's reading skill level is significantly below that of the readability level of the document, then it can be reasonably assumed that the individual was not able to fully comprehend what s/he read. This can adversely impact the health literacy and, ultimately, the education of the child. This paper focuses on the concept of readability; calculation of readability, and suggestions to enhance the readability of student assignments.

The increasing popularity of online information resources has provided many new methods for the delivery of health instructions. Implementation of technology for health instruction is actively encouraged by AAHE and is enthusiastically embraced today's tech-savvy students.¹ WebQuests are a popular inquiry-oriented method used by secondary teachers to help students access information on the web, analyze influences from a variety of viewpoints and develop decision making skills within a single activity. Enter the word "WebQuest" in any Internet search engine, and you soon discover thousands of online lessons created by teachers from around the world.

According to Bernie Dodge, the educational technologist who originated this learning tool, WebQuest is an inquiry-oriented online tool for learning.² More specifically, it is a classroom-based lesson in which most or all of the information that students access, read, and evaluate comes from the World Wide Web. WebQuests have evolved into

activities that can be as short as a single class period or as long as a month-long unit. WebQuests usually (though not always) involve group work with division of labor among students who take on assigned roles or perspectives and are built around resources that are pre-selected by the teacher. Students spend their time *using* information, not *looking* for it. The Resources section of a WebQuest consists of a list of the resources (bookmarked Web sites, print resources, etc.) that students will access to complete the task. In older WebQuests, resources were listed in a section of their own. More recent WebQuests have the resources embedded within the Process section, to be accessed at the appropriate time.² Since students spend the majority of their time using information, educators should be certain that the resources they select can be comprehended by their students.

While WebQuests support higher level thinking skills, contribute to developing media literate students, and enhance decision making, they often overlook the concept of the readability level of the Internet information used in the activity. What should health educators know about the concepts of readability and text comprehension related to WebQuests? More importantly, what additional steps should be taken to ensure that participants in WebQuests fully understand the information they access and use to make personal health decisions?

Methods

The Concept of Readability

Simply defined, readability is the study of matching reader and text.² In its broadest sense, readability is the sum total of all those elements within a given document that affect the extent to which individuals understand, read at optimum speed, and find text interesting.²⁻⁵ The primary goal of readability assessment tools is to estimate the reader's understanding of the material as a function of the reader's language competence, the subject matter of the text, and the syntactic complexity of the passage.⁶⁻⁸ Although both quantitative and qualitative methods of assessment may be utilized to determine the readability level of a written document, readability research generally centers on the use of formulae that estimate the relative difficulty of a passage by assessing word length/difficulty and sentence length.

Originally developed for use by educators, readability formulas and text comprehension assessments have

been used to evaluate a myriad of other documents. Recent uses of readability instruments included the assessment of informed consent forms used by exercise and sport professionals,^{9,10} business communications,¹¹ patient education materials in Mental Health,¹² patient medication leaflets,¹³⁻¹⁵ advertisements used in consumer publications,¹⁶ product warranties,¹⁶ and collegiate textbooks and printed advertisements.¹⁷⁻¹⁹

Since the initial introduction of reading level assessment in the 1950s, numerous reasons have been suggested as causes of reading difficulty. Among these causes, two factors consistently predicted the difficulty of a text and impacted comprehension: vocabulary and sentence length.^{20,21} More recent reviews of readability factors supported and reinforced the use of word length/difficulty and sentence length formulas for making predictions about readability.^{22,23} Although unable to adjust for word derivations; slang, style, and syntax; previously existing contextual knowledge; and personal interest in the subject matter, readability formulas have become a widely accepted method of estimating the average comprehension of a text by an average reader.^{8,22} It is important to note that other factors such as document legibility, length, print size, use of graphics, primary language different than English, and cultural relevance are also critical factors influencing reader comprehension.^{9,24}

Readability Formulas

The most frequently used tool for determining readability is a readability formula. Readability formulas measure certain features of text which can be subjected to mathematical calculations and can provide predictive information regarding how easily a text will be understood by the average reader.^{5,23,25} Previously calculated by hand, computer programs now dissect passages in less time and with a higher degree of precision.²⁶ Readability levels are achieved by typing multiple passages from the document (most formulas require a minimum of 100 words per passage) into a software program designed for this purpose. Depending on the formula selected, passages of text will be identified by the computer program by grade or reading level. Word processing programs such as Microsoft Word also include the capability of assessing the readability level of any highlighted text (see Table 1).

Included among popular formulas for estimating readability levels are the New Dale-Chall Readability Formula,²³ the Spache Readability Formula,²⁷ the Fry Readability Graph,²⁵ the Flesch Reading Ease

Formula,²⁸ and the Gunning Fog Index.²⁹ Readability measures are primarily based on factors such as the number of words in the sentences and the number of letters or syllables per word (i.e., as a reflection of word frequency). From these counts, sentence length and word length are combined within each formula to compute the estimated readability level.

The Gunning Fog Index and the Flesch Reading Ease Formula are the most often used methods to measure the readability of health-related communications and adult level documents.¹⁸ Flesch Reading Ease scores can range from zero, for extremely difficult reading, to 100, for very easy reading. A score of 60 is considered plain English: about 20 words per sentence and 1.5 syllables per word. Today most states within the United States (U.S.) require insurance documents to have a score between 40 and 50.

The Gunning-Fog Index is a rough measure of how many years of schooling it would take someone to understand the content of the document in question. Scores are reported in equivalent grades from 0-12 where the lower the number, the more understandable the content. Negative results are reported as zero, and numbers over twelve are reported as twelve. A score of seven would indicate that the reader must demonstrate a 7th grade reading level for adequate comprehension. Although computer programs require less time, hand-calculation formulas may also be used (See Table 2).

Literacy Rates in the United States

Literacy is defined by the U.S. Department of Education as using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential.³⁰ Consequently, literacy is a measure of everyday functioning, rather than simply academic achievement. The U.S. Department of Education noted that an average person must have an eighth grade reading level to meet the literacy demands of American society.³¹ Recent estimates of the average adult reading grade level range from grade eight to grade four and that an increasing number of children are reading *below* their grade level.³² Given these parameters, recent estimates suggest that over 40-44 million Americans may be functionally illiterate while an additional 50 million are marginally literate, able to locate and assimilate information in simple text forms only.³²

Additional data about health literacy suggested that literacy proficiency is strongly related to levels of

formal schooling.³³ In general, proficiency on all dimensions of health literacy is lowest for individuals who have not graduated from high school, higher for high school graduates and GED holders, and highest for individuals who have attended postsecondary schooling. This pattern is found for African-American, Hispanic, and White populations; for males and females; and for adults in all age ranges. Additional findings have suggested that Hispanic adults demonstrate lower average health literacy rates than any other racial/ethnic groups.³³ Also, living below the poverty level had lower health literacy than living above it.³³

Central to the application of these statistics is the following inquiry: Did the individual who just embarked on your problem solving assignment of a WebQuest possess sufficient reading skills that would allow him/her to read the page and to comprehend what s/he was reading? If a patron's individual reading skills are significantly below that of the determined readability level of the document, it can be reasonably assumed that the individual was not likely to comprehend what they had just read.³⁴

Results

Example: Middle School Target Audience

After a terrorist attack, it's a normal struggle with the emotional impact of the devastating violence and loss of life. You are likely to experience variety of feelings, ranging from shock to anger, fear, anxiety, and a sense of helplessness. You may start to feel unsafe and insecure in situations and places you used to enjoy.

The health content in this example related to violence prevention is relevant and specific. What about its readability for the target audience? The three sentences provided above consist of about 19 words per sentence and 57 total words. The Microsoft Word readability statistics were: Flesch Reading Ease Score = 30 (difficult); and Grade Level = 11. Its major limitation is that the functional literacy for the information posted makes it unlikely that the average U.S. middle school student will comprehend what he/she read. How many other Internet based educational materials would reflect the same limitation?

A recent review of 105 random chosen health-related website pages representing the Center for Disease Control and prevention (CDC) identified six risk factors for youth (sexual activity, intentional and

unintentional injury, physical inactivity, nutrition, alcohol and other drugs, and tobacco) was performed using the Microsoft Word readability function (Table 1). The analysis yielded an average reading level of grade 11. This level far surpasses the literacy level of the average adult in the United States and most students in our health classes. A second analysis performed on health websites whose information was targeted for specific grade levels (grades 6-12, N = 85). There was an average readability level of grade 11 for all sites with a weak correlation between readability level and grade level of the intended audience ($r = .34$). This data suggests that there was little relationship between the targeted grade level and actual readability grade level of the information.

Conclusion and Recommendations

Before developing a WebQuest, the first line of action for health educators is to determine the readability characteristics of the materials to be used by their students. From a strategic point of view it would appear most effective to adapt the readability to the lowest level of literacy of the population with which you wish to use the website. However, given the specifics and content requirements of Internet documents, this strategy may not always be feasible. There are, nevertheless, practical steps that can be taken to increase the comprehension and retention of the health information contained in the text.

Before embarking on a WebQuest activity, health educators should evaluate both their written instructions and the potential websites for text clarity, and reading difficulty. Both quantitative and qualitative approaches should be used for this type of investigation. Equipped with this information, the difficulty of the text should then be judged in terms of the participant's background, educational level, reading confidence, capabilities, and interests.⁷ Data regarding educational levels for geographic regions is readily available from the U.S. Census website [www.census.gov], and a few minutes spent with those who will be reading the information will provide valuable insight into their document comprehension. It is critical to remember that the higher the readability level of the document, the more time must be spent ensuring that students fully understand the activity in question.

When designing WebQuests and health communications, several strategic changes can be implemented to enhance reader comprehension. These include (1) increasing the size of font, (2)

avoiding or simplifying technical language, (3) including anecdotes to illustrate important concepts, (4) utilization of headings, subheadings, and short paragraphs, and (5) inclusion of graphics.^{7,9} Short sentences in the active voice with bulleted lists also contribute to better comprehension.³²

Special consideration should be given to the age (and reading level) of the reader, and the amount of time allocated for reading the document. It is reasonable to conclude that regardless of their literacy level the more rushed a student is to complete the task, the less content they can comprehend. Teaching pre-service teachers how to perform a readability assessment of documents used in lessons will also help sensitize them to the needs of their students, and to develop their own activities that will optimize the comprehension and critical thinking skills of their students.

The primary objective of readability assessments has been to identify the level of skill needed for readers to comprehend printed documents. Therefore, examining popular internet sites for readability level and text coherence is an appropriate and necessary use of this educational tool. An approach that considers the literacy level of the readers and the readability of the text should prove beneficial in the development of a health literate student.

References

1. American Association for Health Education. Health literacy: A position statement of the American Association for Health Education (AAHE). *HExtra*. 2008; 33(3):5-5-7.
2. Gilliland J. *Readability*. London: University of London Press, Ltd; 1972.
3. Chall JS. *Readability: An appraisal of research and application*. Columbus, OH: The Ohio State University Press; 1958 Accessed Mar, 2007.
4. Klare GR. *The Measurement of Readability*. Ames, IA: Iowa State University Press; 1963.
5. Fry E. Reading formulas: Maligned but valid. *Journal of Reading*.1989; 32:292-292-297.
6. Hittleman DR. Readability, readability formulas, and cloze: Selecting instructional materials. *Journal of Reading*.1978; 22:117-122.
7. Gunning TG. The role of readability in today's classrooms. *Topics in Language Disorders*. 2003; 3:175-175-189.
8. Stahl SA. Vocabulary and readability: How knowing word meanings affects comprehension. *Topics in Language Disorders*.2003; 2:241-241-247.
9. Cardinal BJ, Martin JJ, Sachs ML. Readability of written informed consent forms used in exercise and sport psychology research. *Research Quarterly for Exercise and Sport RQES*. 1996; 67:360-360-363.
10. Cardinal BJ. Reduced readability has no effect on research participants' ability to understand an informed consent form: Implications for exercise and sport science research. *Research Quarterly for Exercise and Sport RQES*. 2000; 71(1):A51.
11. Curtis JK. Fry, SMOG, lix and rix: Insinuations about corporate business communications. *The Journal of Business Communication*.1987; 24(2):19-19-27.
12. Adkins AD, Singh NN. Reading level and readability of patient education materials in mental health. *Journal of Child and Family Studies*.2001; 10(1):1-1-8.
13. Kirkpatrick MF, Mohler C. Using the readability assessment instrument to evaluate patient medication leaflets. *Drug Information Journal*.1999; 33:557-563.

14. Slaten D, Parrott R, Steiner C. Readability of skin cancer prevention brochures targeting parents of young children. *Journal of the American Academy of Dermatology*.1999; 40:997-997-999.
15. Wilson FL. Patient education materials nurses use in community health. *Western Journal of Nursing Research*.1996; 18:195-195-205.
16. Shuptrine FK, Moore EM. Even after the magnuson-moss act of 1975, warranties are not easy to understand. *Journal of Consumer Affairs*.1980; 13:393-393-404.
17. Gallagher DJ, Thompson GR. A readability analysis of selected introductory economics textbooks. *Journal of Economic Education*.1981; 12(2):60-60-63.
18. Clark LC, Kaminski PF, Brown G. The readability of advertisements and articles in trade journals. *Industrial Marketing Management*.1990; 19:251-251-260.
19. Flory SM, Phillips TJ, Tassin MF. Measuring readability: A comparison of accounting textbooks. *Journal of Accounting Education*.1992; 10(1):151-151-158.
20. Kearl B. A closer look at readability formulas. *Journalism Quarterly*.1948; 25:344-344-348.
21. Stolurow LM, Newman JR. A factorial analysis of objective features of printed language presumably related to reading difficulty. *Journal of Educational Research*.1959; 52:243-243-251.
22. Meyer B. Text coherence and readability. *Topics in Language Disorders*.2003;3: 204-224.
23. Chall JS, Dale E. *Manual for the New Dale-Chall Readability Formula*. Cambridge, MA: Brookline Books; 1995.
24. Doak CC, Doak LG, Root JH. *Teaching Patients with Low Literacy Skills*. Philadelphia, PA: Lippincott; 1985.
25. Fry E. Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of Reading*.1968; 21:242-242-252.
26. Mailloux SL, Johnson ME, Fisher DG, Pettibone TJ. How reliable is computerized assessment of readability. *Computers in Nursing*.1995; 13:221-221-225.
27. Spache GS. *Good Reading for Poor Readers*. 9th ed. Champaign, IL: Garrard; 1974.
28. Flesch R. *The Art of Readable Writing*. New York: Collier Books; 1949.
29. Gunning R. *The Technique of Clear Writing*. New York: McGraw-Hill; 1968.
30. Lippman L. *The Health Literacy of America's Adults: Results from the 1998 National Assessment of Adult Literacy*. Washington, D.C.: U.S. Department of Education; National Center for Education Statistics; 2001.
31. United States Department of Educational Statistics. *Digest of Educational Statistics*. Washington, D.C.: United States Department of Education; 1986.
32. Bar-Yam NB. Low health literacy: A problem for all in the health care circle. *International Journal of Childbirth Education*.2002; 17(3):21-21-25.
33. Kutner M, Greenberg E, Jin Y, Paulsen C. *The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy (NCES 2006-483)*. Washington, D.C.: U.S.

Department of Education; National
Center for Education Statistics;
2006.

34. Stone LA. A demonstration of different estimations of readability for several forms of the miranda warnings and associated waivers of rights. *Electronic Journal of Forensic Psychonomics* [serial online]. August 2000: Mar 2007.

Table 1. Step for Determining Readability Level Using Microsoft Word 2007

- Step 1: Open Microsoft Word document you wish to analyze
- Step 2: Click on the Microsoft icon in upper right corner
- Step 3: Choose “Word Options”
- Step 4: Click on “Proofing”
- Step 5: Under the Grammar and Spelling section, check the “Show Readability Statistics” box
- Step 6: Click the “Okay” box
- Step 7: Highlight a section of the document that includes at least 100 words. Avoid the introductory and concluding paragraphs.
- Step 8: Initiate the spelling and grammar check (under Tools in the top menu).
- Step 9: At the completion of the spell check, Microsoft Word will present a box that includes a word count for your selection, as well as Flesch Reading Ease and Flesch-Kinkaid grade level readability statistics.

**With the readability statistics box checked, Microsoft Word will provide an estimated readability level for the document or highlighted section every time spell check initiate is initiated.*

Table 2. Hand Calculation Formulas for Estimating Readability Levels Using the Flesch Reading Ease and Gunning Fog Index

Sample Sentence (Note: a passage is more accurate reflection): *After a terrorist attack, it's a normal struggle with the emotional impact of the devastating violence and loss of life.*

- A. To manually calculate the Flesch Reading Ease Readability Estimate Example
1. Count the number of words in the document. 20
 2. Count the number of syllables in the document. 33
 3. Count the number of sentences in the document. 1
 4. Calculate the average sentence length (ASL) value. To do this, divide the number of words in the document by the number of sentences in the document. 20/1
 5. Calculate the average number of syllables per word (ASW) value. To do this, divide the number of syllables in the document by the number of words in the document. 33/20= 1.65
 6. Use the following formula to calculate the Flesch Reading Ease value of your document: $206.835 - (1.015 \times \text{ASL}) - (84.6 \times \text{ASW})$

$$206.835 - (1.015 \times 20) - (84.6 \times 1.65) = 206.835 - 20.3 - 139.59 = 46.945 = 47$$

A score of 60 is considered plain English.

Sample Passage: *After a terrorist attack, it's a normal struggle with the emotional impact of the devastating violence and loss of life. You are likely to experience variety of feelings, ranging from shock to anger, fear, anxiety, and a sense of helplessness. You may start to feel unsafe and insecure in situations and places you used to enjoy.*

- B. To manually calculate the Gunning Fog Index Readability Estimate
1. Select a short passage and count the number of words. 57
 2. Count the number of sentences within the passage 3
 3. Count the number of big words (words with three or more syllables). Exclude words in which "es" or "ed" form the third or final syllable; hyphenated words like "state-of-the-art"; and compound words like "dishwasher". 7
Using technical terms or jargon tends to increase Fog Index scores. If these are words with which the audience is familiar, don't count them as big words.
 4. Calculate the average sentence length by dividing the number of words by the number of sentences. 19
 5. Calculate the percentage of big word by dividing the number of big words by the number of words, and multiplying by 100. 7/57X100= 12
 6. Add the average sentence length to the percentage of big words and multiply that result by 0.4. (19+12)X 0.4= 12
-