
Readability Analysis of Programming Textbooks: Traditional Textbook or Trade Book?

Leah Schultz
lschult@tarleton.edu
Computer Information Systems
Tarleton State University
Stephenville, Texas 76401, United States

Abstract

This study looks at the readability of introductory programming textbooks compared to trade book publications aimed at IT professionals. Using three different readability algorithms, samples of text were analyzed from 40 different programming books to determine if there was a difference in readability measurements between the two types of books. Results indicate that there is a significant difference in readability between trade books and textbooks across all three measures with the readability levels of trade books being more difficult to read. An analysis of variance was also conducted to determine if there were differences in readability depending on the programming language taught in the book, but there were no statistical differences discovered. The results are discussed in terms of usefulness in textbook selection for programming classes and further research considerations are presented.

Keywords: programming instruction, textbook selection, readability

1. INTRODUCTION

As college textbooks become more expensive and more options become available in the form of electronic textbooks and course packets, many faculty members are rethinking their approach to textbook selection. One way to reduce cost in technical classes is through the selection of trade books that are generally less expensive and many times are more current than traditional textbooks. However, with the difference in audience for these textbooks, the selection of a trade book may not be doing students any favors beyond the price tag.

While the editing and procurement processes at textbook publishing companies are geared towards the higher education audience, trade books from technical publishers have a varied audience from the "dummies" to the IT professional. As a result, one might think that the approach to writing these books would vary according to the intended audience. This study examines whether one aspect out of many considerations, readability, differs between traditional textbooks and trade textbooks in

programming. This factor, along with other factors may provide information technology faculty additional information for selecting the programming textbook most appropriate for their classroom.

2. BACKGROUND

Textbook Pricing and Selection

Textbooks, which have been the foundation of higher education classrooms for decades, have undergone major changes over the past few decades. The growth of the e-book, the flat sales in textbook publishing due to the used book market, and rising costs of higher education in general have affected the choices available to faculty when selecting textbooks. (Lichtenberg, 1992). In recent years, the increasing cost of attendance in higher education has brought the rising cost of textbooks into the spotlight as well. In 2005, the Government Accountability Office issued a report on textbooks in higher education that found over the last two decades, the increase in the price of textbooks has doubled the rate of inflation and has been increasing at approximately 6% each

year. The release of this report led to the introduction of many state bills regarding textbook selection and encouraged many faculty to examine their own textbook selection process (Berry, Cook, Hill & Stevens, 2011).

Ask technology professors why they choose a certain textbook and you will get a myriad of different answers: subject coverage, appropriate level for student, inclusion of supplemental material, reputation of the author, and other reasons. Once a faculty member finds a textbook that they like and believe is effective, many will adopt that same textbook until either the instructor's criteria changes or perhaps is presented with a product that is perceived as "better" (Durwin & Sherman, 2008).

How good are faculty at selecting textbooks? Landrum, Gurung, and Spann (2012) argue that faculty are not necessarily the best judges when it comes to selecting textbooks based on pedagogical criteria. Their reasoning is that faculty were exceptional students during their educational career and do not necessarily represent the audiences they are teaching. In addition, their evaluation may also be hampered by their previous knowledge of the subject material, of which, most of their students are lacking (Durwin & Sherman, 2008).

One additional factor that might lead to better textbook selection is student preference. Many studies have documented the fact that reading the textbook leads to improved performance. With many studies showing a decrease in the amount of time students spend reading, any factor that will encourage students to read should be considered. (Berry et al., 2011; Clump, Bauer & Bradley, 2004) Many studies have shown that student preference is a factor that affects how much time students spend reading their textbook. In addition to that, a recent study also indicated that readability was one of the factors that lead to higher levels of student satisfaction with their assigned textbook. (Durwin & Sherman, 2008)

With all of these criteria in play, programming faculty may be tempted to substitute trade books about programming languages instead of using textbooks designed specifically for college level programming classes. These books tend to be less expensive and are updated frequently to keep track with new releases in programming environments. With titles that include phrases such as "for Dummies" or "in 24 hours", faculty

members might assume that these less formal titles may appeal to younger students. However, faculty may need to consider additional criteria in evaluation to ensure that the selected textbooks are appropriate for their students. One aspect of selection to consider would be readability of the material.

Readability

Readability is the concept of measuring the ease of reading content based on the structure of the writing. In the 1940s and 1950s formulas were developed to rate the readability of text based on the length of sentences, the number of syllables in words, and overall sentence structure. These formulas have been adapted over the years but are still used in many areas. The federal government has used readability formulas to ensure wide understanding in the citizenry (Longo, 2004; Risoldi, Gregory & Wilson, 2012, Reilly & Richey, 2011).

Many research projects have been conducted in the medical field to study readability of medical instructions, organ donation forms, and many other print and web based health information sites. (Yick, 2008; Paasche-Orlow, Brancati, Taylor, Jain, Pandit, & Wolf, 2013; Bennet, Drane, & Gilchrist, 2012; Wiener & Wiener-Pla, 2014). In addition, there have been additional studies in other areas such as academic research, adult literacy, and other areas in business such as investor reporting and disclosure (Gazni,, 2001; Lee, 2012; Sawyer, Laran, & Xu, 2008, Gray, 2012). There is also a growing body of literature assessing the readability of textbooks in business disciplines (Flory, Phillips, & Tassin, 1992; Tinkler & Woods, 2013; Plucinski, 2011,).

Readability research in some areas has also shown an inverse relationship with student performance. In a survey business course, Spinks and Wells (1993) found that students in courses with textbooks with higher readability scores had lower course averages. In addition to performance indicators, the authors also found a strong correlation between the difficulty of the course textbook and the number of withdrawals from the course. The readability of a textbook could become increasingly important because the numbers of students who come to college needing reading and writing remediation is increasing. Estimates in the mid-2000s indicate that nine percent of college students enrolled at four year universities took foundation courses in reading and double that number

requiring writing remediation (Attewell, Lavin, Domina & Levey, 2006). These numbers are even higher at two year schools. Even more challenging is the report from the National Endowment for Humanities which reports a 20% decline in reading proficiency in college graduates (Office of Research & Analysis NEA, 2007).

The following section will cover three common readability algorithms that were used in this study and many others.

Flesch Readability Scale and Flesch-Kincaid Grade Level Scale

In the 1940s reading expert Rudolph Flesch developed a formula to assess readability based on the average sentence length and the number of syllables per word. The scale has been validated numerous times and is one of the most commonly used scales to evaluate readability (Cochrane, Gregory, & Wilson, 2012). The formula produces a number from 0 to 100 indicating the readability of a written passage. A lower score indicates a higher level of difficulty in reading the passage. In the 1960s, the Flesch Reading scale was adapted by Peter Kincaid for use in assessing readability of Naval training material. The Flesch-Kincaid scale uses the foundation of the original formula developed by Flesch and assigns grade level reading categories based on the result. (Dubay, 2007). This test has become so common that it is included in current versions of Microsoft Word review functionality. However, these, and other algorithms have been criticized due to the lack of correlation between the score and actual reading comprehension scores (Moley, Bandre', & George, 2011).

Gunning Fog Test

Another commonly used test to evaluate readability is the Gunning Fog test developed by Robert Gunning. Gunning's test focuses on the use of multi-syllable words to indicate difficulty in reading. Gunning believes that clear writing avoids unnecessary words and that writers should choose simple over complex words. The formula is based on the percentage of long words compared to average sentence. Some critics of the measure argue that not all long words are necessarily difficult words and do not correlate to difficult reading. Nevertheless, the measurement is still commonly used (Spinks & Wells, 1993).

3. PROBLEM STATEMENT

The purpose of this study is to determine the readability of programming textbooks and trade books and compare the results. This data can assist faculty in selecting instructional material that is appropriate to their audience. When exploring the possibility of using trade books over traditional textbooks, faculty will be able to determine the level of writing and its potential impact on their students.

4. METHODS

In order to study the readability of traditional textbooks versus the trade books, 20 introductory programming textbooks designed for higher education use and 20 popular trade books on introductory programming topics were selected. The textbooks were selected from textbook vendors in four different areas: Java, C#, Visual Basic, and HTML/CSS. Five books on each subject were selected from current catalogs of textbook publishers. For the trade books, a similar sampling was selected with the assistance of Amazon.com. The trade books in the four areas were selected from the best-selling lists in each of these categories. Books that were written on advanced topics or specialized areas of the programming languages were not selected. Instead, titles that indicated they were general or beginner books in the areas were selected.

Once the books were selected in both categories, three samples were analyzed from each book using an online tool. Samples from textbooks were taken in one of two ways depending on access to the book. Books that were available physically were sampled in the middle of the second chapter, the middle of the middle chapter, and the middle of the next to last chapter. This allowed for samples that were more likely to be explaining concepts rather than introductory matter at the beginning of the book or at the beginning of chapters. For titles selected from Amazon, samples were taken using the random generator for previewing content. Any content that was in the beginning or ending chapters of the book was discarded for the reasons mentioned above. Both methods were employed to try to eliminate bias in the selection of the samples for analysis.

Using an online tool, samples were processed to determine the Flesch Readability score, the corresponding Flesch-Kincaid score and the Gunning Fog score. When inputting the samples, some minor variations were made from time to time to improve the accuracy of the score. In many of the programming samples, the periods in dot notation would indicate the premature end of a sentence in a normal analysis. To avoid this, a decision was made to separate these into two words because that is how most people would say them. For example, if the author was discussing the label.text property, the period should not indicate the end of a sentence. If left with the period in place, there would be a significant effect on the results of the analysis. In addition to these programming specific limitations, URLs appeared in some of the samples taken from the textbook. The decision was made to remove these completely from the sample in that they would affect the readability algorithm. If left alone, the periods would indicate the end of a sentence, resulting in many small sentences. If combined, the word would be considered a very long, multi-syllable word which would also affect the analysis.

5. FINDINGS

The average scores for all samples were analyzed using SPSS to determine if the average scores of the trade books differed significantly from the average readability scores of the textbooks.

Flesch Reading Ease (FRE)

An independent T-test was conducted using the samples to determine if the average Flesch Reading Ease scores of the two types of book were significantly different. The mean FRE for textbooks was 70.81 (SD=10.78) and the mean FRE for trade books was 65.99 (SD=11.22). The results of the T-test indicate there was a significant difference in the readability of the two groups, $t(60)=2.397$, $p<.05$. This difference indicates that on the FRE measurement, the trade books, with the lower score, were significantly harder to read than the textbooks.

Flesh-Kincaid Grade Level (FKGL)

The next set of data analyzed was the samples' ranking using the FKGL. The mean average grade level for programming textbooks was 7.15 (SD=2.42) and the mean grade level for trade books was 8.92 (SD=2.46). Using an

independent T-test to analyze the means, the results indicate that the difference between the two means is significant, $t(60)=-3.96$, $p<.05$. Again, the results mirror those of the FRE in that the readability of the trade books increases in difficulty by almost two grade levels.

Gunning Fog Test (Fog)

The final data analyzed between the two groups used the Fog scores of each sample. The mean average Fog score for the trade book sample was 11.79 (SD=2.82) and the mean score for the textbook sample was 9.72 (SD=2.99). The results of the T-test indicate that there also was a significant difference between these two means, $t(60)=-3.90$, $p<.05$. These results of the third readability measure support the idea that the trade books have more difficulty in reading than textbooks.

Differences Between Languages

An additional analysis was performed to determine if there were any readability differences among programming languages that might affect the overall readability of the groups. Samples from both types of books were analyzed and the means of each programming language, regardless of type of book, were compared. Results from the ANOVA test using the FRE score indicated no significant difference between the means of the four groups of programming languages, $F(3,116)=2.28$, $p=.08$. The same result was found for the FKGL test, $F(3, 116)=1.11$, $p=.35$, as well as the means for the Fog measure, $F(3,116)=1.00$, $p=.40$. This indicates it is unlikely that the programming language in the books chosen had little effect on the readability of the text.

6. CONCLUSIONS

Readability is just one of many factors that a faculty member may consider when selecting the instructional material to be used in their introductory programming language course. The results of this study indicate that there is a trade off in readability when choosing a professional trade book over a traditional textbook from an established textbook publishing house. With mixed messages about textbook usage showing that students spend less time reading their textbooks while expressing the importance of the textbook to their learning in the class overall, it is important to take many factors into consideration.

One could argue that not only does the trade book prepare students to use the types of resources they are more likely to use in professional practice, but it allows students to save money on their books during their time in school. Assuming that coverage of topics is the same, it appears to be a good option. However, with the results of this study, there is an indication that these books could be harder to understand due to language and the nature in which they are written. To further cloud programming concepts for introductory students, seems counterproductive to the learning process. A textbook that is densely written may be used even less than a traditional textbook and may leave students frustrated.

Faculty members who are well-versed in the programming concepts they are teaching, may not be cognizant of the difficulty in the readability due to their expertise. So should trade books written for professionals be excluded for all programming classes? The answer is probably "no", as readability is one of many factors to consider. However, faculty members may want to have students in their target audience review a trade text before adoption or may consider using one of the online readability tools to assess the difficulty of the chosen text. When reviewing textbooks of both categories, faculty members can easily take a few samples from a potential textbook and use the readability analysis tools built into Microsoft Word or any similar tools available online. Altman, Ericksen, and Pena Shaff (2006) outline a textbook selection process that involves many faculty peers and receives feedback from students about the textbooks selected for courses.

One other thing to keep in mind when reviewing the readability of potential programming textbooks, is the limitation of the readability algorithms. For example, a high score on the FOG scale may just indicate the presence of many multi-syllable words that may not necessarily be hard for students to understand. In a programming text, common multi-syllable words such as properties, object-oriented, superclass, and others may not necessarily make the text more difficult for the reader to understand. To a layperson reading a programming text of either category, these words could potentially increase the readability of the text significantly. However, once these words become part of a student's jargon, subsequent encounters with these words are

unlikely to continue causing problems of understanding. These formulas do not take into consideration the reader's background in the subject matter and can potentially underestimate the difficulty of the text (Burke & Greenberg, 2010).

Again, the purpose of this paper is not to put the concept of readability at the top of the list of criteria for selecting programming textbooks. Instead, the study sheds light on one way faculty can think about their instructional material when introducing concepts that may be difficult for students to comprehend. There is no reason to exacerbate the problem with unnecessarily dense, complicated writing. Instead, texts, whether they are from traditional or trade publishers, should present the necessary material in a straightforward, easy to read concept so that students can focus on the concepts they will need to build a foundation for programming. A difficult to read textbook runs the risk of not only impeding a student's learning but could potentially discourage them from exploring a lucrative and fulfilling career path.

7. FURTHER RESEARCH

As mentioned many times throughout this paper, readability is only one aspect of textbook selection to consider. Future research in this area may focus on relationships between the readability of a programming textbook and student's preference. For example, one might think that books written in a conversational manner popular in some trade book such as the Dummies series, would be more enjoyable to students. However, the readability of the titles included in this study was surprisingly difficult.

In the end, student success and learning is the most important outcome of an introductory programming class. Studies regarding textbook selection and use patterns of students in programming classes should be studied in relationship to student success. If readability is related to improved learning, then it should be a very important factor for faculty in textbook selection. Along this line, future studies may investigate whether the readability scores of the formulas presented correlate with faculty perceptions of the readability of chosen textbooks.

One other factor that may warrant further study is the inclusion or supplementation of readable

texts with illustrations and other graphical aids. As printing processes become less expensive and more interactive experiences are available in e-books, the combination of visual and textual information may work in tandem to make textbooks more readable.

8. References

- Altman, W. S., Ericksen, K., & Pena-Shaff, J. B. (2006). An Inclusive Process for Departmental Textbook Selection. *Teaching Of Psychology*, 33(4), 228-231.
- Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New Evidence on College Remediation. *Journal Of Higher Education*, 77(5), 886-924.
- Bennett, D. M., Drane, E., & Gilchrist, A. (2012). Readability of CAMHS clinical letters. *Child & Adolescent Mental Health*, 17(3).
- Berry, T., Cook, L., Hill, N., & Stevens, K. (2011). An Exploratory Analysis of Textbook Usage and Study Habits: Misperceptions and Barriers to Success. *College Teaching*, 59(1), 31-39.
- Burke, V., & Greenberg, D. (2010). Determining Readability: How to Select and Apply Easy-to-Use Readability Formulas to Assess the Difficulty of Adult Literacy Materials. *Adult Basic Education & Literacy Journal*, 4(1), 34-42.
- Clump, M. A., Bauer, H., & Bradley, C. (2004). The Extent to which Psychology Students Read Textbooks: A Multiple Class Analysis of Reading across the Psychology Curriculum. *Journal Of Instructional Psychology*, 31(3), 227-232.
- DuBay, W.H. (2007). The Classic Reability Studies. Retrieved from ERIC database (ED506404).
- Durwin, C., & Sherman, W. M. (2008). Does Choice of College Textbook Make a Difference in Students' Comprehension?. *College Teaching*, 56(1), 28-34.
- Flory, S.M., Phillips, T.J., & Tassin, M.F. (1992). Measuring Readability: A Comparison of Accounting Textbooks. *Journal of Accounting Information*, 10, 151-161.
- Gazni, A. (2011). Are abstracts of high impact articles more readable? Investigating the evidence from top research institutions in the world. *Journal of Information Science*, 37(3), 273-281.
- Gray, C. J. (2012). Readability: A Factor in Student Research? *Reference Librarian*, 53(2), 194-205.
- Landrum, R., Gurung, R. R., & Spann, N. (2012). Assessments of Textbook Usage and the Relationship to Student Course Performance. *College Teaching*, 60(1), 17-24.
- Lee, Y. (2012) The Effect of Quarterly Report Readability on Information Efficiency of Stock Prices. *Contemporary Accounting Research*, 29(4), 1137-1170.
- Lichtenberg, J. (1992). The new paradox of the college textbook. (Cover story). *Change*, 24(5), 10.
- Longo, B. (2004). Toward an Informed Citizenry: Readability Formulas as Cultural Artifacts. *Journal Of Technical Writing & Communication*, 34(3), 165-172.
- Moley, P. F., Bandre, P. E., & George, J. E. (2011). Moving Beyond Readability: Considering Choice, Motivation and Learner Engagement. *Theory Into Practice*, 50(3), 247-253.
- Office of Research and Analysis, National Endowment for the Arts (2007). To Read or Not to Read: A Question of National Consequence. (Research Report #47) Retrieved from <http://arts.gov/sites/default/files/toRead.pdf>
- Paasche-Orlow, M. K., Brancati, F. L., Taylor, H. A., Jain, S., Pandit, A., & Wolf, M. S. (2013). Readability of Consent Form Templates: A Second Look. *IRB: Ethics & Human Research*, 35(4), 12-19.
- Plucinski, K. J. (2011). Readability of Cost Accounting Textbooks. *Business Education Innovation Journal*, 3(1), 15-20
- Reilly, S., & Richey, S. (2011). Ballot Question Readability and Roll-Off: The Impact of Language Complexity. *Political Research Quarterly*, 64(1).

-
- Risoldi Cochrane, Z., Gregory, P., & Wilson, A. (2012). Readability of Consumer Health Information on the Internet: A Comparison of U.S. Government-Funded and Commercially Funded Websites. *Journal Of Health Communication*, 17(9), 1003-1010.
- Sawyer, A., Laran, J., & Xu, J. (2008). The Readability of Marketing Journals: Are Award-Winning Articles Better Written?. *Journal Of Marketing*, 72(1), 108-117.
- Spinks, N., & Wells, B. (1993). Readability: A textbook selection criterion. *Journal Of Education For Business*, 69(2), 83.
- Tinkler, S., & Woods, J. (2013). The Readability of Principles of Macroeconomics Textbooks. *Journal Of Economic Education*, 44(2), 178-191.
- U.S. Government Accountability Office (2005). College Textbooks: Enhanced Offerings Appear to Drive Recent Price Increases. Report to Congressional Requesters. (GAO Publication No. GAO-05-806). Retrieved from <http://www/gao.gov/products/GAO-05-806>.
- Wiener, R. R., & Wiener-Pla, R. (2014). Literacy, Pregnancy and Potential Oral Health Changes: The Internet and Readability Levels. *Maternal & Child Health Journal*, 18(3), 657-662.
- Yick, A. G. (2008). Evaluating Readability of Domestic Violence Information Found on Domestic Violence State Coalitions' Websites. *Journal Of Technology In Human Services*, 26(1).