

Readability of Financial Accounting Principles Textbooks

Wei-Chih Chiang
Texas A&M International University

Ted D. Englebrecht
Louisiana Tech University

Thomas J. Phillips, Jr.
Louisiana Tech University

Ying Wang
Montana State University-Billings

Abstract

This study investigates whether there are readability differences among leading principles of accounting textbooks. Seven financial accounting principles textbooks with a 2002 or later publishing date are reviewed. Readability scores are obtained using the following four readability evaluation methods: Flesch Reading Ease, Flesch-Kincaid Grade Level Index, Gunning's Fog Index, and the SMOG procedure. Three sets of comparisons are made: overall readability comparisons, chapter comparisons, and topic comparisons. The Kruskal-Wallis test is used for multiple group comparisons. For significant differences, the Wilcoxon Rank-Sum test with a Bonferroni adjustment for multiple comparisons is used in further comparisons. Results indicate significant differences for all three sets of comparisons, although the conservative Bonferroni adjustment substantially reduces the number of differences that are shown here and could even reduce textbook differences that were found in some previous studies. Evidence suggests that readability among textbooks varies, but readability within textbooks is generally consistent.

Introduction

Reduced numbers of accounting students lead to concerns of reforming accounting curriculums and attracting students to the profession (Albrecht and Sack, 2000; Diller-Haas, 2004; Russell et al., 2000). The first course in accounting is of significance for both those considering a career in accounting and those who have selected the option to major in accounting (AECC, 1992). Generally, prior research (e.g., Spinks and Wells, 1993) finds that the higher the readability level of the textbooks, the lower the grade averages are found in these courses. Specifically, Peterson (1982) concludes that the readability of accounting textbooks has a significant impact on students' learning achievement. In addition, performance in the first accounting course could influence students' decisions to major in accounting (Cohen and Hanno, 1993; Geiger and Ogilby, 2000). Hence, the readability of accounting principles textbooks may affect students' decisions to major in accounting. After taking this into account, it is not surprising

that comprehensibility to students is ranked as the most important criterion in accounting faculty's textbook selection procedures (Smith and DeRidder, 1997).

The potential usefulness of the first accounting course provides the impetus to study the readability of seven leading principles of accounting textbooks.¹ Four readability evaluation methods, the Flesch Reading Ease formula (Flesch, 1948), the Flesch Grade Level Index (Sticht, 1973), Gunning's Fog Index (Gunning, 1952), and the SMOG method (McLaughlin, 1969) are used to compare readability.² While the current study neither recommends an appropriate readability level nor which accounting textbooks should be selected (e.g., see Flory et al., 1992), it does focus on whether these textbooks differ in readability. Naturally, this information can offer assistance during the textbook selection process.

This study differs from prior research in several aspects. First, readability of four selected chapters is compared within and among textbooks. The chapter comparisons can provide additional information, especially to the instructors who are concerned with the readability of hard-core chapters. Second, chapters with similar topics in each textbook are grouped based on content, then readability levels of these groups are compared. The topic comparisons can assist instructors who intend to teach the financial accounting principles in several integrated segments. Furthermore, the homogeneity of the overall readability within each textbook is examined. Finally, to protect against false significant results, the Bonferroni adjustment for multiple comparisons is applied when performing the Wilcoxon Rank-Sum test. The adjustment for multiple comparisons has not been performed in the readability context. See Tables 1A and 1B for a summary and comparison of studies that evaluate readability of financial accounting textbooks.

The remainder of the article is structured as follows: in the following section, we provide background about readability evaluation methods and prior studies in readability of accounting textbooks; section two covers hypothesis development; section three describes research methods; section four provides the testing results; and the last section discusses the results, noting limitations and future research opportunities.

Background

Readability and Understandability

Readability is an attribute of the text, but understandability (comprehensibility) is an attribute of the reader which is a broader concept than readability (Harrison, 1980; Jones, 1997). That is, readability indicates the textual difficulty level of a message; on the other hand, understandability is concerned with the reader's ability to comprehend the materials presented in a text. Specifically, readability focuses on the textual content, but understandability is an interaction among text (e.g., text structure, topic content, word familiarity, and sentence length), task (e.g., mode or rate of presentation), reader (e.g., verbal ability and word knowledge, education, age, and interest), and strategy variables (e.g., rereading and underlining) (Meyer, 2003). Therefore, estimating readability of a text is not equivalent to predicting a reader's comprehension. The focus of this paper is limited to the attribute of readability.

Readability Evaluation Methods

Readability Formulas. Readability formulas (largely, mathematical formulas) are designed to assess the suitability of books for particular grade levels or ages. Readability formulas may be performed manually by counting and performing mathematical calculations or by referring to a chart or graph. Usually, these applied

¹ The first author's last name will be used to indicate the textbook in this study.

² Computer software, Readability Calculations (Application Version 6.3), developed by Micro Power & Light Co. is employed to perform the four methods. This software was introduced in the early 1980's and is reviewed by several articles (e.g., Dewey, 1984; Chan, 1985). Also, it has been used in liberal arts literature (e.g., Peterson and Coniglio, 1987) as well as in accounting research (e.g., Marsh et al., 2004; Reinstein and Houston, 2004).

routines count the length of sentences and the number of syllables in words. They do not take into account sentence/word order. Among readability standards, the Flesch Reading Ease formula, the Flesch Grade Level Index, Gunning's Fog Index, and the SMOG procedure, have been applied in the accounting research arena. Various formulas provide a numerical value equivalent to grade level. Today, readability scores can be easily ascertained through the use of computerized grammar and style checking software programs.

Cloze Method. The Cloze method (Taylor, 1953) can provide valid measures of the comprehension abilities of readers as well as the comprehension difficulties of messages (Adelberg and Razek, 1984). It has been used in accounting contexts (e.g. Adelberg and Razek, 1984 ; Raabe et al., 1984). The Cloze procedure consists of deleting words in a text and asking the reader to fill in the appropriate or a similar word. Usually every fifth word is deleted (Adelberg and Razek, 1984; Stevens et al., 1993).

Readability in Accounting Textbooks

Several prior studies examine readability in financial accounting textbooks. Razek et al. (1982) use the Flesch Reading Ease formula to compare the readability of six most often used intermediate accounting textbooks and six most often used advanced accounting textbooks. Using the student T test, they find that some books have significantly different mean Flesch scores. Adelberg and Razek (1984) use the Cloze method to compare the level of understandability of four intermediate accounting texts and conclude that the level of understandability varies significantly among the four textbooks and within the four textbooks.³ Flory et al. (1992) also examine the readability across seven intermediate accounting textbooks, but consider the scores derived from the Flesch Reading Ease Formula and Gunning's Fog Index as ordinal data. The Kruskal-Wallis test and Mann-Whitney U test are used to test for the difference of readability across the textbooks. They find that one textbook is more difficult to read than the others.

Regarding the readability of accounting principles textbooks, Traugh et al. (1987) use Gunning's Fog Index to evaluate three categories of accounting principles textbooks (successive editions of books which have been published for many years, first edition books, and books that fall in-between). An analysis of variance is used and the conclusion is that there is no significant difference in the readability of the textbooks within and among categories. Sullivan and Benke (1997) measure the readability of thirty-three introductory financial accounting textbooks. Principles textbooks with both financial and managerial accounting are not included in their study. They use WordPerfect's Grammatik, a well-known grammar software package that employs the Flesch-Kincaid Grade Level Index to evaluate complete chapters covering the same or very similar materials. The introductory chapter and the chapter on long-term assets are selected. Although the readability scores of each chapter in the selected textbooks are reported, no statistical comparison is done in their study.

Instead of reporting the Flesch scores, Davidson (2005) uses Grammatik to assess the trend of word complexity and sentence complexity separately for introductory, intermediate, and advanced accounting textbooks over the past one hundred years. He finds that all three groups of textbooks gathered for this study show decreases in sentence complexity and increases in word complexity over the relevant time periods. Introductory textbooks have the lowest levels of both sentence and word complexity compared with intermediate and advanced accounting textbooks.

Hypothesis Development

Even though studies on accounting textbook readability have been performed, several issues still need to be addressed. First, regarding to the readability of the accounting principles textbooks, Traugh et al. (1987) find that there is no significant difference in the readability level of the accounting principles textbooks selected in their research. However, other studies (e.g., Flory et al., 1992; Razek et al., 1982) show that readability among either intermediate accounting textbooks or advanced accounting textbooks varies significantly. Meanwhile, the number of

³ Adelberg and Razek (1984) evaluate understandability rather than readability.

published accounting principles textbooks has increased and their presentation styles are quite diverse over the past decades. Based on the presentation attribute, Sullivan and Benke (1997) classify principles accounting textbooks into five categories: conventional, moderately conventional, transitional, moderately revolutionary, and revolutionary. Textbooks in different categories may have different readability levels. Nevertheless, there is no particular reason to expect that readability among principles accounting textbooks should be consistent. We expect that some principles accounting textbooks have different readability levels. Re-examination is needed to investigate whether the current readability among accounting principles textbooks is still homogeneous. Therefore, hypothesis one is shown as:

H₁: Some accounting principles textbooks have different readability levels from the others.

Second, Sullivan and Benke (1997) indicate some readability differences between two selected chapters of accounting principles textbooks in their study; however, they do not compare the readability of the selected textbooks. Since the readability of different chapters could provide additional information for textbook selection decision-making, this paper compares readability of four selected chapters (including chapters covering long-term assets, long-term debt, stockholders' equity, and statement of cash flows) within and among seven textbooks. The selected chapters contain the hard-core knowledge of principles of financial accounting. Because co-authorship or variance of the contents may result in readability differences (Adelberg and Razek, 1984), we expect readability of the selected chapters varies within and among the seven textbooks. Hence, hypotheses two and three are:

H₂: Readability of the four selected chapters within each textbook varies significantly.

H₃: Readability of the four selected chapters across textbooks varies significantly.

Moreover, to provide a broader perspective than that given by chapter comparison, this paper groups the contents of each textbook into four topics for comparing topic readability within and among the seven textbooks. When information is presented in a logically organized manner, people can read faster and comprehend more (Meyer, 2003). Some accounting instructors may want to integrate materials into several topics and present them in a logical order. Test results allow comparisons of the readability level of each topic so that instructors can consider the most appropriate level by topic. Due to authors' various presentation styles and variance of contents, we expect the readability level of topics varies within and among the seven textbooks. Thus, hypotheses four and five are as follows:

H₄: Readability of the four topics within each textbook varies significantly.

H₅: Corresponding readability of the four topics across textbooks varies significantly.

Research Method

Why Readability Formulas Are Chosen

We do not choose the Cloze method for several reasons. First, recent research criticizes that Cloze is a measure of understandability rather than a measure of readability (Jones, 1997).⁴ Second, Davies and Vincent (1972) suggest that Cloze is inappropriate for measuring text or reader's abilities in languages other than their native language. The results of Cloze testing reflect the reader's basic intuition about the structure and vocabulary of the target language – and that does not exist for the non-native language speakers. Third, the Cloze method is typically costly and time consuming to perform (Flory et al., 1992). Adelberg and Razek (1984, p. 117) provide only two passages and Raabe et al. (1984, p. 70) offer one passage for subjects to read. However, the small number of passages may not faithfully represent the true readability of lengthy textbooks, especially where co-authors are involved (Flory et al., 1992). On the other hand, readability formulas are easier to implement, particularly with computerized software packages now available that incorporate formulas. The user friendly implementation would benefit instructors in assessing the readability of their potential textbooks. Furthermore, readability formulas are objective. They are valid and easy to replicate, allowing for the selection of numerous reading passages without the complications of the Cloze procedure. Although readability formulas are questioned to some extent (e.g., Davison and Kantor, 1982; McConnell, 1983),⁵

⁴ For additional criticisms on the Cloze method applied in accounting research, see Jones (1997).

⁵ McConnell (1983), for example, argues that readability formulas are not appropriate for technical material where understanding depends more upon terminology. However, it may be disputable whether accounting is technical in

Fry (1989, p. 294) recognizes readability formulas as “one of the best documented educational tools that we have.” Finally, since Traugh et al. (1987) uses readability formulas to evaluate the principles of financial accounting textbooks, this paper also employs formulas to revisit the issue.

Four readability formulas – the Flesch Reading Ease formula, the Flesch Grade Level Index, Gunning’s Fog Index, and the SMOG procedure – are applied in this study. The first three formulas have been used in prior research to measure the readability of financial accounting principles textbooks. This paper extends the readability measure to the SMOG which has been employed in accounting research arena (e.g., Marsh et al., 2004) but not readability of accounting textbooks. Furthermore, multiple formulas allow us to evaluate the correlation among them. If a high correlation is found among the formulas, then accounting educators might only use one of the four methods.

Readability Comparisons

Three sets of comparisons are used to test the hypotheses. First, there is an “overall” comparison where the readability of each textbook is compared to the other textbooks to investigate whether there is a significant difference. Thus, the overall comparisons test hypothesis one.

The second set of comparisons is “chapter” comparisons. Four similar chapters are selected, including the chapters covering long-term assets, long-term debt, stockholders’ equity, and statement of cash flows. The reason for selecting these four comparable chapters is that all of the seven textbooks have these four chapters of which the content is reasonably consistent. Furthermore, these chapters convey the hard-core knowledge of accounting principles, which is critical for students to understand. The readability of the selected chapters within each textbook and across the seven textbooks is compared. Hence, the chapter comparisons test hypotheses two and three.

The third set of comparisons is “topic” comparison. The chapters in Hermanson are divided into several segments, which center on different perspectives of financial accounting principles, such as introduction, accounting information process and policy, management’s perspectives, and financial statements. Therefore, according to the presentation method and the technical level of the contents, chapters in each textbook are assembled into four topic groups: introduction to accounting (hereafter “Topic One”); basic concepts of accounting such as the double entry system, the accrual method, and other similar concepts (hereafter “Topic Two”); more technical contents of accounting such as the inventory system, assets, liabilities, and equity (hereafter “Topic Three”); and financial statements (hereafter “Topic Four”). The readability of the four topics is compared to assess whether there are any significant differences in the readability among topics within each textbook and across the seven textbooks. These topic comparisons test hypotheses four and five.

Adjustment for Multiple Comparisons

Since readability formulas may lose information through the averaging of words per sentence and number of syllables, the measured scores are treated as ordinal data rather than interval data (Flory et al., 1992; Lewis et al., 1986). Therefore, non-parametric techniques are employed in this study.⁶

While using multiple comparisons, prior research (e.g. Flory et al., 1992; Razek et al., 1982; Traugh et al., 1987) does not adjust the significance level for multiple comparisons. That is, the experiment-wise error is not controlled. In this paper, the Wilcoxon Rank-Sum test with a Bonferroni adjustment is used for controlling the experiment-wise error at the 0.05 level (Higgins, 2004, p. 93). As a result, two tests are followed. First, the Kruskal-Wallis test is used to test for the experiment-wise difference. Second, if a significant difference is found, the Wilcoxon Rank-Sum test with a Bonferroni adjustment for multiple comparisons is used to address which of the groups differ.

nature. Also, Flory et al. (1992) note that when the purpose of study is not to determine the absolute reading level of textbooks, this limitation is not applicable. For an overview of readability formulas, see Klare (1988).

⁶ The results are not essentially changed when parametric methods are employed in this study. However, because of the measurement scale, we strictly use nonparametric results when reporting in the tables.

Sample Selection

Seven leading principles of accounting textbooks summarized in Table 2 are chosen for comparisons.⁷ None of the seven textbooks includes both financial and managerial accounting materials (Sullivan and Benke, 1997). Sixteen pages are determined to be the necessary sample size at the 0.05 level according to the following formula (Snedecor and Cochran, 1967, p. 58):⁸

$$n = 4\sigma^2/L^2,$$

where n is the sample size, σ is the estimated population variance,⁹ and L is the required accuracy.

For overall comparisons, the sample of sixteen pages is randomly selected from each textbook (throughout the whole text). For chapter or topic comparisons, the sample of sixteen pages is randomly selected from each selected chapter or topic group in each textbook, respectively.¹⁰

Only pages containing main contents are included as potential sampling pages. That is, the pages containing only charts, graphs, key terms, (review) problems, indices, or appendix are not selected.¹¹ This criterion is similar to that in Sullivan and Benke (1997). All of the complete sentences in the pages chosen are evaluated by the four readability evaluation methods.

Test Results

Overall Readability Comparison

Table 3 shows a statistically significant difference at the 0.05 level among the overall readability of the seven textbooks, using the Flesch Grade Level Index, Gunning's Fog Index, and SMOG. Since the overall comparison using the Flesch Reading Ease Index is not significant, the Flesch Reading Ease Index is excluded from the Wilcoxon Rank-Sum test with a Bonferroni adjustment. The test results of the Wilcoxon Rank-Sum test with a Bonferroni adjustment are reported in Table 4. The pair-wise p -value should be no greater than 0.0024 to be considered significant when using a Bonferroni adjustment.

Table 4 shows that the overall readability of Harrison is significantly different from that of Hermanson, Horngren, Needles, and Phillips. Therefore, hypothesis one is confirmed. After checking the raw data, we find that Harrison has easier readability compared to these four textbooks. This finding is similar to those of Razek et al. (1982) and

⁷ The seven texts in the sample are potential textbooks considered by the School of Accountancy-Assessment Committee at a traditional southern university for the introductory financial accounting course. Sullivan and Benke (1997) categorize Harrison, Hermanson, and Needles as "conventional," Horngren as "moderately conventional," and Edmonds as "transitional," but Kimmel and Phillips are not on their list. Nevertheless, at least four of the seven textbooks are classified as "conventional/moderately conventional," which by definition generally relies on a debit and credit presentation. Meanwhile, Traugh et al. (1987) address the issue of whether readability differs among successive editions of accounting principles textbooks. However, they do not find significant readability difference among different versions of textbooks. Therefore, it would be appropriate to compare current textbooks in different editions.

⁸ The Flesch Reading Ease Method shows scores on a scale between 0 and 100, but the other three methods give a grade level, usually less than 20. Since the Flesch Reading Ease Method uses a unit different from others, its required accuracy is adjusted correspondingly. Therefore, five is assigned to L for the Flesch Reading Ease Method, but one was assigned to L for the Flesch Grade Level Index, Gunning's Fog Index, and SMOG. A sample size of sixteen pages is equal to the largest number calculated from these four methods.

⁹ The estimated population variance is obtained via a pilot test pursuant to the procedure recommended by Cochran (1977, p. 19). Specifically, five pages from each book are randomly selected and the pooled variance is calculated.

¹⁰ Some chapter samples have less than sixteen eligible pages. In this circumstance, all eligible pages of the selected chapter are included in the sample.

¹¹ Pages with a combination of text and charts, graphs, etc. are included, when they contain main contents that exceed 100 words (Traugh et al., 1987). Nevertheless, only the main contents are evaluated.

Flory et al. (1992), but inconsistent to that of Traugh et al. (1987). The inconsistency may arise from the different textbooks selected in Traugh et al. and in this study, especially that Harrison is included in this study but not in Traugh et al. Furthermore, the diversity in attributes of accounting principles textbooks published after the 1990s could result in the inconsistency (Sullivan and Benke, 1997). Nevertheless, Hermanson, Horngren, and Needles are categorized as conventional or moderately conventional as well as Harrison, but the later has easier readability. While Edmonds is classified as transitional, its readability level is similar to Harrison's. Consequently, the impact of diverse presentation attributes of accounting principles textbooks on readability is unclear. Future research is needed to address this issue. Meanwhile, because accounting educators rank comprehensibility as the most important criterion in textbook selection (Smith and DeRidder, 1997), textbooks with comparable popularity may have similar readability levels. The popularity of the sampled textbooks in this study may explain why only Harrison has a different readability level.

This finding provides some evidence that the overall readability among accounting principles textbooks is different.¹² Hence, accounting instructors should consider the overall readability of their principles textbook choices when making decisions regarding which textbook is adopted.

Chapter Readability Comparisons within Textbook

Table 5 reports the multiple comparison results of the readability of the four selected chapters within each textbook using the Kruskal-Wallis test, thereby confirming hypothesis two. However, only two textbooks, Hermanson and Phillips, are significant at the 0.05 level, suggesting that the readability of the four selected chapters in the textbook is consistent in general. Hence, readability differs little through the contents of the four selected chapters.

Further investigations are performed on Hermanson and Phillips, and results are shown in Table 6, in which the pair-wise p-value should be no greater than 0.0083 to be considered significant when using a Bonferroni adjustment. After analyzing the raw data, we find that the long-term assets and the statement of cash flows chapters are less readable than the long-term debt in Hermanson. Likewise, the long-term assets and the statement of cash flows chapters are less readable than the stockholders' equity chapters in Phillips. Moreover, in Phillips the long-term debt chapter is easier to read, compared to the long-term asset chapter. The difference might arise from co-authorship or the variance of the contents (Adelberg and Razek, 1984). Also, writing structure might affect the readability. For example, a bullet point format which usually uses a short sentence is more readable than an essay format. However, while we look at the corresponding chapters in Hermanson and Phillips, no apparent difference in writing structure was found within each book.

Chapter Readability Comparisons among Textbooks

Table 7 suggests that the readability of the corresponding selected chapters among textbooks is significantly different. Therefore, hypothesis three is confirmed. Further investigations lead to the results reported in Tables 8-11, and the pair-wise p-value in these tables should be no greater than 0.0024 to be considered significant when using a Bonferroni adjustment. The interpretations of these tables are discussed in turn.

First, generally speaking, Tables 8-11 indicate the readability of the four chapters in Harrison is easier than that of the corresponding chapters in Hermanson, Horngren, Needles, and Phillips. This is consistent with the previous finding that the overall readability of Harrison is easier than that of Hermanson, Horngren, Needles, and Phillips. Second, Tables 8 and 11 show that the readability of the long-term assets and the statement of cash flows chapters in Kimmel is easier than that in Phillips. This is also consistent to the previous finding that the long-term assets and the statement of cash flows chapters are less readable than the long-term debt and the stockholders' equity chapters in

¹² Without the Bonferroni adjustment we would have shown additional differences in textbooks. Under the Bonferroni adjustment, for example, Table 4 indicates that four book sets have different overall readability. However, without the Bonferroni adjustment the number of significant book sets increases to eleven. That is, eleven of the twenty one book sets are significantly different.

Phillips. Even though neither the overall readability of Kimmel and Phillips in Table 4 nor the overall readability of the four selected chapters in Kimmel in Table 5 is significantly different, this does not preclude the readability of the long-term assets chapter in Kimmel from being easier than that in Phillips.

In addition, Table 9 shows that the readability of the long-term debt in Hermanson is higher than that in Horngren. Using the same rationale arguing for the relation of chapter-readability between Kimmel and Phillips, this result is consistent to the previous finding that in Hermanson the long-term assets and the statement of cash flows chapters are less readable than the long-term debt and the stockholders' equity chapters.

Table 11 shows that the readability of the statement of cash flows chapter in Needles is significantly lower than that in Horngren and Kimmel. The raw data show the overall readability of Needles is slightly less than that of Horngren and Kimmel, but the differences are not significant as shown in Table 4. Furthermore, the data indicate that the readability of the statement of cash flows chapter in Needles is slightly less than that of the other three selected chapters, but the difference is not significant either. Also, from examining the data, it seems that the chapter of statement of cash flows in either Horngren or Kimmel is the easiest chapter among the four selected chapters, whereas the readability difference is not significant. In summary, the readability of the compatible chapters among the textbooks (excluding Harrison) is consistent in general.

Topic Readability Comparisons within Textbook

Table 12 shows that significant differences exist within Harrison, Needles, and Phillips, thereby confirming hypothesis four. Further group comparisons using the Wilcoxon Rank-Sum test with a Bonferroni adjustment are performed for these three textbooks, and the results are shown in Table 13. The pair-wise p-value should be no greater than 0.0083 to be considered significant. Table 13 indicates that the readability of Topic One and Topic Four is significantly different from that of Topic Two in Needles. Raw data indicate that the readability of Topic Two is easier than that of Topic One and Topic Four in Needles. Nevertheless, the readability of the four topic groups within each textbook (except for Needles) is generally consistent.

Naturally, the heterogeneous topic readability might arise from co-authorship, the variance of the contents (Adelberg and Razek, 1984), writing structure, or inappropriate grouping. One possible explanation for the heterogeneity may be the writing structure in Needles. We scrutinize the related chapters in Needles and feel that Needles uses relatively more bullet points in Topic Two than in Topic One and Topic Four, resulting in easier readability of Topic Two.

Topic Readability Comparisons among Textbooks

Table 14 reports the overall comparison results of the readability of each topic among the seven textbooks. Significant differences are found when using all four readability evaluation methods. Thus, hypothesis five is confirmed. Further group comparisons using the Wilcoxon Rank-Sum test with a Bonferroni adjustment are performed with results reported in Tables 15-18. The pair-wise p-value in these tables should be no greater than 0.0024 to be considered significant.

In Tables 15-18, most of the significant differences occur where the readability of topics in Harrison is compared to that in other textbooks, especially in Hermanson, Horngren, Needles, and Phillips. This finding is consistent to the previous result that the overall readability of Harrison is significantly different from that of Hermanson, Horngren, Needles, and Phillips.

Table 15 shows that the readability of Topic One in Edmonds is significantly different from Hermanson and Needles. Raw data indicate that the readability of Topic One in Edmonds is higher than Hermanson and Needles. We realize that Needles usually summarizes the illustration events in very short sentences which may increase its readability. Apparently, the summarization writing style cannot explain Needles' lower readability of Topic One, compared to Edmonds. Further inquiry is needed to interpret these differences. Nevertheless, the readability of topics among textbooks, excluding Harrison, is reasonably consistent.

Interpretation for Readability within Each Textbook

Except for Adelberg and Razek (1984) who test for homogeneity within selected intermediate accounting textbooks, prior research does not address this issue. Although it is not essential that readability within a textbook be similar, some might prefer consistent readability within a textbook, making homogeneity of readability within textbooks of interest.

To test the homogeneity issue within each textbook, Adelberg and Razek (1984) divide textbooks into two parts and compare the Cloze scores between the parts. Since each textbook in this study is divided into four topics, we are able to test the homogeneity of readability within each textbook.

As mentioned above, the results show that the readability of the four topic groups within each textbook (other than Needles) is generally consistent. Thus, the equal readability of the four topics within each textbook implies that the overall readability of each selected textbooks is homogeneous.

Inter-Relationship of the Four Readability Evaluation Methods

Correlations of the readability scores obtained using the four evaluation methods are also analyzed. We use all of the samples chosen from each textbook (for either overall comparison, chapter comparison, or topic comparison), and analyze the relationship of the readability scores obtained using the four evaluation methods for the entire sample and each textbook. The Spearman correlations are reported in Table 19.

For all seven textbooks, readability scores obtained using the four readability methods are highly correlated ($p < .0001$). Thus, the four readability evaluation methods chosen seem consistent with each other. The consistency implies that accounting faculty could use only one of the four methods to evaluate the readability of the accounting principles textbooks and have reliable results. Nevertheless, some variations of test results using the four evaluation methods may occur. For example, Table 11 shows some variations among the four evaluation methods.

Discussion

This study shows that overall readability is different among accounting principles textbooks. Unlike the results in Traugh et al. (1987), we find some differences in readability among accounting principles textbooks. First, the selected sample textbooks are not the same in these two studies. Also, textbooks published after the 1990s are more diverse. However, most of the textbooks in this study are from the conventional or moderately conventional categories. As a result, if the sample textbooks in this study were more diverse, the findings might indicate more differences. Finally, the current study controls for experiment-wise error (by using the Bonferroni adjustment), resulting in statistical significance that reduces the probability of falsifying a true null hypothesis and reduces the number of differences found in this study.¹³ This adjustment has not always been used in previous studies.¹⁴

In comparing textbooks, some chapters might be more difficult to read than others. For example, the first chapter (usually the introductory chapter) is easier for students to read than the chapter containing the double-entry system in which the concept of “debit/credit” is first considered.¹⁵ In addition to the approach of introducing concepts,

¹³ See footnote 12.

¹⁴ Adelberg and Razek (1984) use Scheffe procedure to control for experiment-wise error.

¹⁵ Although Sullivan and Benke (1997) find several textbooks in which the readability grade level is higher for the introductory chapter than for the depreciation chapter, this situation happened in only 8 out of 33 textbooks in their sample. Further, statistical tests are not used to check the readability difference in their study. Only 2 of the 8 textbooks have the difference of mean grades for the two chapters larger than one. Thus, not majority of financial accounting principles textbooks have higher readability grade level for the introductory chapter than for the depreciation chapter.

accounting instructors should be aware of readability levels. Although chapter readability is generally consistent within the textbooks, several chapters are found to be more difficult to read than others.

Kopp and Phillips (2005) find that the organization of topics influences how students structure their knowledge in memory, and they suggest that the structure of organizing topics presented in classes, courses, and curricula has an important effect on student learning. Although the current study finds topic readability is generally consistent within and among textbooks, some differences are found. When accounting educators realize that students experience difficulties in particular topics, readability level should be evaluated as a potential cause. Preferable, perhaps, is considering topic readability in advance.

Accounting textbooks may have a major influence on the nature and type of learning activities in preparing students as new accounting professionals (Davidson and Baldwin, 2005). Further, as the first accounting course, it is desirable that principles of financial accounting attract students to major in accounting. Thus, an appropriate readability level of financial accounting textbooks is important. However, the easiest readability is not always the solution because the use of more complex written grammar and the ability to work with complex problems are significantly related (Davidson et al., 2000). Further, if the mental models developed by accounting students are too simplistic, they may not be able to work with today's complex business situations (Davidson, 2005).

This study has several limitations. First, readability formulas have some inherent weaknesses. For example, readability formulas only consider the word and sentence complexity, but omit other qualitative variables such as vocabulary level, syntax, text coherence, the complexity of ideas, reinforcement through restatement and repetition, and writing styles (McConnell, 1983). Therefore, users of formulas should carefully interpret the scores (Klare, 1988). Second, the study focuses on word-level and sentence-level features of the narrative and not the whole-text. To date, no alternative text-focused method is widely accepted. Sydserff and Weetman (1999) suggest a texture index that is a potentially powerful tool for analysis of accounting narratives and associated testing, but future research of this approach is needed. Third, the textbooks used in this study are not randomly selected. There are more than 30 financial accounting principles textbooks available in the market (Sullivan and Benke, 1997), but only seven are chosen in this study. Thus, the finding of general consistency of chapter and topic readability might be biased because of our selection. Finally, the grouping of topics is somewhat subjective. Different researchers might categorize topics in different ways and, hence, the results of topic comparison might differ from those in the current study.

Future research can be pursued on the following issues. First, Sullivan and Benke (1997) classify financial accounting principles textbooks into five categories. However, whether readability within and across these categories varies is unknown. Second, textbooks which include both financial and managerial accounting materials are excluded in Sullivan and Benke (1997) and this study. Nevertheless, whether these two types of textbooks differ in terms of readability has not yet been addressed. Third, Kopp and Phillips (2005) indicate two new structures that emerge for organizing accounting by topics: by business process or by stage of the business lifecycle. Further research is needed to inquire whether readability of textbooks with and without these course structure changes is different.

Conclusion

All five hypotheses are confirmed in this study. In the overall comparison among textbooks, we find that Harrison's readability is significantly easier than the other selected textbooks. For chapter comparisons, there are few significant differences in the four selected chapters within each textbook. Chapter comparisons among books reveal that Harrison shows significant differences from other textbooks. For topic comparison, within book topic comparison did not indicate many significant differences. All significant differences of within book topic comparisons are related to the second topic. The second topic includes basic accounting concepts, the accrual method, the double entry system, and related items. Some books employ more journal entries and worksheets for illustration. Users should choose according to their preference or particular need(s). However, it is interesting to note

that crucial subject material such as that covered in the second topic area might have a different readability level than other topics. The topic comparisons among books reveal that most of the significant differences are related to Harrison.

Further analysis of the four readability evaluation methods indicates that they are highly related with each other. This provides evidence that readability formulas give consistent and objective results.

This study focuses on the readability difference within and among textbooks. The study does not attempt to answer questions such as the appropriate readability grade level or which textbook results in the best overall subject comprehension. We do not make recommendations concerning textbook selection. Davison and Kantor (1982) point out that a more complex text may be easier to understand than a simplified version. Further research is needed to determine the proper readability level for financial accounting principles textbooks. However, accounting educators should at least be aware of readability level differences in choosing a textbook for financial accounting principles.

REFERENCES

- AECC (Accounting Education Change Commission)(1992). The First Course in Accounting: Position Statement No. Two. *Issues in Accounting Education*, 7(2), 249-251.
- Adelberg, A. H., and Razek, J. R. (1984). The Cloze Procedure: A Methodology for Determining the Understandability of Accounting Textbooks. *The Accounting Review*, 59(1), 109-122.
- Albretch, W.S. and Sack, R.J. (2000). *Accounting Education: Charting the Course Through a Perilous Future*, Sarasota, FL: American Accounting Association.
- Chan, J. M. T. (1985). Computer Software. *Journal of Reading*, February, 470-471.
- Cochran, W. G. (1977). *Sampling Techniques*, 3rd edition, New York: John Wiley & Sons.
- Cohen, J. and Hanno, D. (1993). An Analysis of Underlying Constructs Affecting the Choice of Accounting as a Major. *Issues in Accounting Education*, 8(2), 219-238.
- Davidson, R. A. (2005). Analysis of the Complexity of Writing Used in Accounting Textbooks over the Past 100 Years. *Accounting Education: an International Journal*, 14(1), 53-74.
- Davidson, R. A. and Baldwin, B. A. (2005). Cognitive Skills Objectives in Intermediate Accounting Textbooks: Evidence from End-of-Chapter Material. *Journal of Accounting Education*, 23, 79-95.
- Davidson, R. A., Slotnick, S. A., and Baldwin, B. A. (2000). Using Linguistic Performance to Measure Problem-Solving. *Accounting Education*, 9(1), 53-66.
- Davison, A., and Kantor, R. (1982). On the Failure of Readability Formulas to Define Readable Texts: a Case Study from Adaptations. *Reading Research Quarterly*, 17(2), 187-209.
- Davies, P. and Vincent, D. (1977). The Development of Cloze Tests, in *Reading: Research and Classroom Practice* edited by Gilliland, J., London, UK: United Kingdom Reading Association, 244-255.
- Dewey, P. (1984). Software for Libraries. *Wilson Library Bulletin*, November, 222-223.
- Diller-Haas, A. (2004, April). Time to Change Introductory Accounting. *The CPA Journal*, 60-62.
- Flesch, R. (1948). *The Art of Readable Writing*, New York: Harper & Brothers Publishers.
- Flory, S. M., Phillips, T. J. Jr. and Tassin, M. F. (1992). Measuring Readability: A Comparison of Accounting Textbooks. *Journal of Accounting Education*, 10, 151-161.
- Fry, E. B. (1989). Reading Formulas – Maligned But Valid. *Journal of Reading*, 32, 292-297.
- Gallagher, D. J. and Thompson, G. R. (1981). A Readability Analysis of Selected Introductory Economics Textbooks. *Journal of Economic Education*, Summer, 60-63.
- Geiger, M. A. and Ogilby, S. M. (2000). The First Course in Accounting: Students' Perceptions and Their Effect on the Decision to Major in Accounting. *Journal of Accounting Education*, 18, 63-78.
- Gunning, R.(1952). *The Technique of Clear Writing*, New York: McGraw-Hill.

- Harrison, C. (1980). *Readability in the Classroom*, New York, NY: Cambridge University Press.
- Higgins, J. J. (2004). *Introduction to Modern Nonparametric Statistics*, Pacific Grove, CA: Brook/Cole-Thomson Learning.
- Jones, M. J. (1997). Methodological Themes: Critical Appraisal of the Cloze Procedure's Use in the Accounting Domain. *Accounting, Auditing & Accountability Journal*, 10(1), 105-120.
- Klare, G. R. (1988). The Formative Year, in *Readability: Its Past, Present, and Future* edited by Zakaluk, B. L. and Samuels, S. J., Newark, DE: International Reading Association, 14-34.
- Kopp, L. S. and Phillips, F. (2005). Integrating Accounting Topics within or across Functions: Effects on Students' Structure and Use of Knowledge. *Journal of Accounting Education*, 23, 170-188.
- Lewis, N. R., Parker, L. D., Pound, G. D., and Sutchliffe, P. (1986). Accounting Report Readability: The Use of Readability Techniques. *Accounting and Business Research*, 16, 199-213.
- Marsh, T. L., Montondon, L. G., and Daniels, J. (2004). Readability of Award Winning Popular Annual Financial Reports. *Journal of Accounting and Finance Research*, Summer I, 32-39.
- McConnell, C. (1983). Readability: Blind Faith in Numbers? *Journal of Economic Education*, Winter, 65-71.
- McLaughlin, H. (1969). SMOG Grading – a New Readability Formula. *Journal of Reading*, 22, 639-646.
- Meyer, B. J. F. (2003). Text Coherence and Readability. *Topics in Language Disorders*, 23(3), July-September, 204-224.
- Peterson, E. A. (1982). A Study of the Relationship of Academic Achievement in Accounting Course to the Readability of Required Textbooks and the Reading Ability of Two-Year Accounting Students, Doctoral dissertation, Oregon State University.
- Peterson, L. and Coniglio, J. W. (1987). Readability of Selected Academic Library Guides. *RQ*, Winter, 232-239.
- Raabe W. A., Stevens K. C., and Stevens W. P. (1984). Tax Textbook Readability: An Application of the Cloze Method. *The Journal of American Taxation Association*, Fall, 66-73.
- Reinstein, A. and Houston, M. (2004). Using the Securities and Exchange Commission's "Plain English" Guidelines to Improve Accounting Students' Writing Skills. *Journal of Accounting Education*, 22, 53-67.
- Razek, J. R., Hosch, G. A. and Pearl, D. (1982). Readability: Readability of Accounting Textbooks. *Journal of Business Education*, October, 23-26.
- Russell, K.A., Kulesza, C.S.B., Albrecht, W.S., and Sack, R.J. (2000). Charting the Course through a Perilous Future. *Management Accounting Quarterly*, Fall, 4-11.
- Smith, K. J. and DeRidder, J. J. (1997). The Selection Process for Accounting Textbooks: General Criteria and Publisher Incentives – A Survey. *Issues in Accounting Education*, 12(2), 367-384.
- Snedecor, G. W. and Cochran, W. G. (1967). *Statistical Methods*, 6th edition, Ames, Iowa: The Iowa State University Press.

- Spinks, N. and Wells, B. (1993, November/December). Readability: A Textbook Selection Criterion. *Journal of Education for Business*, 83-88.
- Stevens, K. C., Stevens, K. T. and Stevens, W. P. (1993). A Response to "Measuring Readability: A Comparison of Accounting Textbooks". *Journal of Accounting Education*, 11, 287-292.
- Sticht, T. G. (1973). Research towards the design, development and evaluation of a job-functional literacy training program for the US Army. *Literacy Discussion*, 4(3), 339-369.
- Sullivan, M. C. and Benke, R. L. Jr. (1997). Comparing Introductory Financial Accounting Textbooks. *Journal of Accounting Education*, 15(2), 181-220.
- Sydserrff, R. and Weetman, P. (1999). A Texture Index for Evaluating Accounting Narratives: An Alternative to Readability Formulas. *Accounting, Auditing & Accountability Journal*, 12(4), 459.
- Taylor, W. L. (1953). Cloze Procedure: A New Tool for Measuring Readability. *Journalism Quarterly*, 30, 415-433.
- Traugh, H. M., Powers O. S. and Adedokun A. J. (1987, January). Readability of Accounting Principles Texts. *Journal of Education for Business*, 159-162.

Table 1A: Summary and Comparison of Studies Evaluating Readability of Financial Accounting Textbooks

	Razek et al. (1982)	Adelberg & Razek (1984)	Traugh et al. (1987)	Flory et al. (1992)
Type of textbooks	Intermediate; Advanced	Intermediate	Principles	Intermediate
Evaluation Methods	Flesch Reading Ease	Cloze method	Fog	Flesch Reading Ease; Fog
Statistical Methods	T-test	F-test; Scheffe procedure	ANOVA	Kruskal-Wallis test; Mann-Whitney U test
Test homogeneity of each textbooks	No	Yes (significant difference)	No	No
Type of readability comparison	Overall	Overall	Overall	Overall
Within/Cross textbook type comparison* and findings	Inter.-Inter. (no significant difference); Adv.-Adv. (significantly different); Inter.-Adv. (significantly different)	Inter.-Inter. (significantly different)	Prin.-Prin. (no significant difference)	Inter.-Inter. (significantly different)
Test correlation of evaluation methods	No	No	No	No

*Prin.: Principles; Inter.; Intermediate; Adv.: Advanced

Table 1B. Summary and Comparison of Studies Evaluating Readability of Financial Accounting Textbooks

	Sullivan and Benke. (1997)	Davidson (2005)	Current Study (2008)
Type of textbooks	Principles	Principles Intermediate Advanced	Principles
Evaluation Methods	Flesch-Kincaid	Flesch-Kincaid	Flesch Reading Ease Flesch Grade Level Fog SMOG
Statistical Methods	No	ANOVA	Kruskal-Wallis test Wilcoxon Rank-Sum test Bonferroni Adjustment
Test homogeneity of each textbooks	No	No	Yes (significant difference)
Type of readability comparison	Chapter (but no statistical comparison)	Overall (word and sentence complexity)	Overall Chapter Topic
Within/Cross textbook type comparison* and findings	N/A	Prin.-Inter.-Adv. (significantly different)	Prin.-Prin. (significantly different)
Test correlation of evaluation methods	No	No	Yes (significantly correlated)

*Prin.: Principles; Inter.; Intermediate; Adv.: Advanced

Table 2. Summary of Selected Textbooks

Authors [†]	Edmonds Edmonds McNair Olds	Harrison Horngren	Hermanson Edwards Herrmann Sellers Thomas Wetzel	Horngren Sundem Elliott Philbrick	Kimmel Weygandt Kieso	Needles Powers	Phillips Libby Libby
Publisher	McGraw- Hill/Irwin	Pearson Prentice Hall	Authors Academic Press	Pearson Prentice Hall	Wiley	Houghton Mifflin	McGraw- Hill/Irwin
Publication Year	2004	2005	2002	2004	2004	2005	2005
Edition	5	6	8	9	3	9	1

[†] The first author's name is used thereafter to indicate the textbook in this paper.

Table 3. Overall Readability among the Textbooks (Kruskal-Wallis Test)

Overall p-value	
FREAD	0.1181
FGRADE	<0.0001*
FOG	0.0025*
SMOG	0.0100*

* Significant at the 0.05 level

FREAD: Flesch Reading Ease Formula
FGRADE: Flesch Grade Level Index
FOG: Gunning's Fog Index
SMOG: Simple Measure of Gobbledygook

Table 4. Overall Comparisons between Textbooks (Wilcoxon Rank-Sum Test)

	Edmonds					
Harrison	0.0166					
	0.2060					
	0.0541	Harrison				
Hermanson	0.0331	<0.0001*				
	0.0734	0.0009*				
	0.1803	0.0022*	Hermanson			
Horngren	0.2267	0.0010*	0.2348			
	0.2350	0.0059	0.4394			
	0.5461	0.0127	0.4065	Horngren		
Kimmel	0.8210	0.0274	0.0434	0.3961		
	0.8064	0.2740	0.0764	0.2825		
	0.6105	0.2496	0.0895	0.3362	Kimmel	
Needles	0.0287	0.0003*	0.6781	0.1519	0.0301	
	0.0594	0.0008*	0.7918	0.2656	0.0620	
	0.0521	0.0012*	0.3265	0.1315	0.0862	Needles
Phillips	0.0247	<0.0001*	0.8063	0.3654	0.1009	0.6922
	0.0795	0.0006*	0.7202	0.3082	0.0734	0.9099
	0.3083	0.0078	0.9399	0.5717	0.1806	0.6108

FGRADE
FOG
SMOG

0.05/21=0.0024
Bonferroni adjustment

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 5. Chapter Comparisons within Each Textbook (Kruskal-Wallis Test)

Overall p-values							
Method	Edmonds	Harrison	Hermanson	Hornngren	Kimmel	Needles	Phillips
FREAD	0.0706	0.1618	0.0002*	0.0950	0.6637	0.1186	0.0051*
FGRADE	0.1198	0.0779	0.0065*	0.4959	0.2718	0.1463	0.0260*
FOG	0.4949	0.3228	0.3927	0.5725	0.0766	0.1399	0.1826
SMOG	0.3442	0.6785	0.1180	0.5184	0.2691	0.1070	0.0515

* Significant at the 0.05 level

Table 6. Readability Comparisons of Chapters within Each Textbook (Wilcoxon Rank-Sum Test)

FREAD	0.05/6=0.0083
FGRADE	Bonferroni adjustment

Hermanson	CH1		
CH2	0.0004*		
	0.0019*	CH2	
CH3	0.1172	0.0182	
	0.3552	0.0499	CH3
CH4	0.1969	0.0003*	0.0234
	0.8950	0.0037*	0.2350

Phillips	CH1		
CH2	0.0314		
	0.0052*	CH2	
CH3	0.0034*	0.4845	
	0.0331	0.5091	CH3
CH4	0.9699	0.0435	0.0061*
	0.8210	0.0477	0.1932

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0083)

CH1: The Long-Term Asset Chapter
 CH2: The Long-Term Debt Chapter
 CH3: The Stockholders' Equity Chapter
 Ch4: The Statement Of Cash Flows Chapter

**Table 7. Overall Readability Comparisons of the Selected Chapters among Textbooks
(Kruskal-Wallis Test)**

Overall Readability Comparisons of	Overall p-value			
	FREAD	FGRADE	FOG	SMOG
The Long-Term Assets Chapter	0.6842	0.0002*	0.0002*	0.0003*
The Long-Term Debt Chapter	0.0132*	0.0009*	0.0406*	0.0179*
The Stockholders' Equity Chapter	0.4642	0.0095*	0.2011	0.2196
The Statement of Cash Flows Chapter	0.0010*	<0.0001*	<0.0001*	<0.0001*

* Significant at the 0.05 level

Table 8. Readability Comparisons of the Long-Term Assets Chapter between Textbooks (Wilcoxon Rank-Sum Test)

	Edmonds					
Harrison	0.1315					
	0.1267					
	0.1048	Harrison				
Hermanson	0.2579	0.0184				
	0.0859	0.0024*				
	0.5971	0.0158	Hermanson			
Horngren	0.7917	0.0794	0.1744			
	0.2736	0.0092	0.6779			
	0.5586	0.0213	0.7771	Horngren		
Kimmel	0.3361	0.5087	0.0397	0.2996		
	0.7484	0.1220	0.0416	0.1129		
	0.6239	0.1742	0.1993	0.2496	Kimmel	
Needles	0.0969	0.0014*	0.3754	0.0300	0.0047	
	0.0248	0.0008*	0.4737	0.2995	0.0214	
	0.0273	0.0004*	0.0378	0.0475	0.0052	Needles
Phillips	0.0194	0.0003*	0.0895	0.0025	0.0006*	0.4970
	0.0066	0.0003*	0.1515	0.0593	0.0019*	0.4617
	0.0193	0.0003*	0.0270	0.0432	0.0074	0.7626

FGRADE
FOG
SMOG

0.05/21=0.0024
Bonferroni adjustment

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 9. Readability Comparisons of the Long-Term Debt Chapter between Textbooks (Wilcoxon Rank-Sum Test)

	Edmonds					
Harrison	0.2199					
	0.0455					
	0.1089					
	0.0378	Harrison				
Hermanson	0.0162	0.4376				
	0.4281	0.1414				
	0.9399	0.0969				
	0.3855	0.2274	Hermanson			
Horngren	0.0472	0.0049	0.0008*			
	0.0127	0.0004*	0.0083			
	0.0970	0.0006*	0.0703			
	0.0889	0.0006*	0.0247	Horngren		
Kimmel	0.5454	0.1087	0.0147	0.2342		
	0.1994	0.0093	0.1089	0.2910		
	0.5717	0.0236	0.5589	0.2996		
	0.3266	0.0049	0.1314	0.3850	Kimmel	
Needles	0.7771	0.1627	0.0589	0.4502	0.9549	
	0.1006	0.0026	0.0381	0.9699	0.4622	
	0.4176	0.0301	0.4068	0.5717	0.7344	
	0.5213	0.0522	0.3176	0.5587	0.8950	Needles
Phillips	0.8353	0.2060	0.0535	0.0857	0.6640	0.6505
	0.1086	0.0020*	0.0417	0.4395	0.7060	0.5976
	0.5973	0.0128	0.5092	0.1804	0.9699	0.7061
	0.5715	0.0150	0.1932	0.2128	0.7342	0.7774

FREAD
FGRADE
FOG
SMOG

0.05/21=0.0024
Bonferroni adjustment

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 10. Readability Comparisons of the Stockholders' Equity Chapter between Textbooks (Wilcoxon Rank-Sum Test)

FGRADE	0.05/21=0.0024					
	Bonferroni adjustment					
	Edmonds					
Harrison	0.1174	Harrison				
Hermanson	0.5844	0.0050	Hermanson			
Horngren	0.8356	0.0521	0.4973	Horngren		
Kimmel	0.8504	0.0398	0.6509	0.8505	Kimmel	
Needles	0.0826	<0.0001*	0.4622	0.1808	0.1869	Needles
Phillips	0.2348	0.0004*	0.7628	0.2581	0.3856	0.8209

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 11: Readability Comparisons of the Statement of Cash Flows Chapter between Textbooks (Wilcoxon Rank-Sum Test)

	Edmonds					
Harrison	0.0148					
	0.0016*					
	0.0049					
	0.0034	Harrison				
Hermanson	0.7625	0.0147				
	0.3458	0.0022*				
	0.6508	0.0053				
	0.5841	0.0057	Hermanson			
Horngren	0.0062	0.2353	0.0062			
	0.1266	0.0157	0.0378			
	0.5093	0.0046	0.2502			
	0.2820	0.0167	0.1173	Horngren		
Kimmel	0.0329	0.4713	0.0247	0.8948		
	0.0078	0.2016	0.0033	0.0932		
	0.0098	0.3902	0.0083	0.0088		
	0.0236	0.1499	0.0184	0.1219	Kimmel	
Needles	0.5460	0.0066	0.8801	0.0016*	0.0103	
	0.0175	0.0002*	0.1688	0.0014*	0.0001*	
	0.0214	0.0003*	0.1746	0.0098	0.0001*	
	0.0213	0.0002*	0.1463	0.0017*	0.0002*	Needles
Phillips	0.6920	0.0202	0.4735	0.0222	0.0496	0.3855
	0.2132	0.0008*	0.7061	0.0193	0.0011*	0.4621
	0.2422	0.0053	0.5716	0.2132	0.0056	0.4505
	0.1361	0.0009*	0.4618	0.0260	0.0014*	0.4065

FREAD
FGRADE
FOG
SMOG

0.05/21=0.0024
Bonferroni adjustment

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 12. Overall Readability Comparisons of Topics within Each Textbook (Kruskal-Wallis Test)

Overall p-values							
Method	Edmonds	Harrison	Hermanson	Horngren	Kimmel	Needles	Phillips
FREAD	0.4753	0.5497	0.1645	0.2293	0.8199	0.0570	0.8903
FGRADE	0.2098	0.1328	0.4548	0.6766	0.9090	0.0706	0.7608
FOG	0.0682	0.1506	0.0766	0.5944	0.8514	0.0091*	0.0483*
SMOG	0.1431	0.0385*	0.3635	0.5319	0.9971	0.0195*	0.2571

* Significant at the 0.05 level

Table 13. Readability Comparisons of Topics within Each Textbook (Wilcoxon Rank-Sum Test)

Harrison [†]	TP1			Needles	TP1		
TP2	-			TP2	0.0053*		
	0.6509				0.0135		
TP3	-	-		TP3	0.1049	0.0828	
	0.0416	0.0794			0.0731	0.1131	
TP4	-	-	-	TP4	0.6783	0.0070*	0.1519
	0.0286	0.0399	0.5213		0.8950	0.0128	0.1868

Phillips [‡]	TP1		
TP2	0.0109		
	-		
TP3	0.1314	0.3960	
	-	-	
TP4	0.6647	0.0331	0.2426
	-	-	-

FOG	0.05/6=0.00833
SMOG	

Bonferroni adjustment

[†] Only SMOG is used in Harrison.

[‡] Only FOG is used in Phillips.

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0083)

TP1: Topic One
 TP2: Topic Two
 TP3: Topic Three
 TP4: Topic Four

Table 14. Overall Readability Comparisons of Four Topics among Textbooks (Kruskal-Wallis Test)

Overall Readability Comparisons of	Overall p-value			
	FREAD	FGRADE	FOG	SMOG
Topic One	0.0001*	<0.0001*	<0.0001*	<0.0001*
Topic Two	0.0058*	0.0001*	0.0039*	0.0025*
Topic Three	0.0399*	0.0002*	0.0005*	0.0011*
Topic Four	0.0062*	0.0002*	<0.0001*	0.0001*

* Significant at the 0.05 level

Table 15. Readability Comparisons of Topic One between Textbooks (Wilcoxon Rank-Sum Test)

	Edmonds					
Harrison	0.0044					
	0.0001*					
	0.0041					
	0.0004*	Harrison				
Hermanson	0.0128	<0.0001*				
	0.0013*	<0.0001*				
	0.0135	<0.0001*				
	0.0044	<0.0001*	Hermanson			
Horngren	0.7915	0.0062	0.0193			
	0.1571	0.0002*	0.0398			
	0.2578	0.0019*	0.1315			
	0.2064	0.0006*	0.0794	Horngren		
Kimmel	0.5086	0.0078	0.0070	0.4500		
	0.5716	<0.0001*	0.0031	0.3363		
	0.5974	0.0194	0.0070	0.1746		
	0.9699	0.0004*	0.0087	0.2737	Kimmel	
Needles	0.1682	0.0001*	0.2200	0.3085	0.0826	
	0.0037	<0.0001*	0.7343	0.0898	0.0109	
	0.0062	<0.0001*	0.6783	0.0933	0.0047	
	0.0016*	<0.0001*	0.9398	0.0520	0.0055	Needles
Phillips	0.5842	0.0593	0.0092	0.5334	0.8651	0.1132
	0.0063	<0.0001*	0.4174	0.1686	0.0380	0.6107
	0.0331	0.0002*	0.9099	0.1808	0.0175	0.7061
	0.0416	0.0001*	0.7773	0.2276	0.0397	0.5215

FREAD
FGRADE
FOG
SMOG

0.05/21=0.0024
Bonferroni adjustment

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 16. Readability Comparisons of Topic Two between Textbooks (Wilcoxon Rank-Sum Test)

	Edmonds					
Harrison	0.0107					
	0.0062					
	0.0109					
	0.0047	Harrison				
Hermanson	0.0891	0.0002*				
	0.0203	<0.0001*				
	0.0732	0.0003*				
	0.0728	0.0001*	Hermanson			
Horngren	0.5329	0.0092	0.2418			
	0.3082	0.0039	0.4621			
	0.2132	0.0020*	0.8949			
	0.4732	0.0014*	0.3547	Horngren		
Kimmel	0.4278	0.0223	0.0328	0.3750		
	0.8799	0.0039	0.0157	0.2997		
	0.9850	0.0128	0.1047	0.2062		
	0.7485	0.0050	0.1517	0.6103	Kimmel	
Needles	0.7914	0.0344	0.0730	0.5842	0.9098	
	0.5459	0.0026	0.1628	0.9099	0.4737	
	0.9850	0.0436	0.1088	0.3555	0.9850	
	0.9099	0.0070	0.1931	0.7193	0.9850	Needles
Phillips	0.5966	0.0173	0.0272	0.3648	0.8502	0.8059
	0.0818	<0.0001*	0.7342	0.6105	0.0247	0.3266
	0.4175	0.0009*	0.2202	0.4853	0.3555	0.5212
	0.4734	0.0010*	0.3265	0.7628	0.4618	0.7485

FREAD
FGRADE
FOG
SMOG

0.05/21=0.0024
Bonferroni adjustment

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 17: Readability Comparisons of Topic Three between Textbooks (Wilcoxon Rank-Sum Test)

	Edmonds					
Harrison	0.0435					
	0.0044					
	0.2275					
	0.0345	Harrison				
Hermanson	0.2060	0.0142				
	0.0225	0.0020*				
	0.0122	0.0033				
	0.0380	0.0017*	Hermanson			
Horngren	0.1460	0.0031	0.6503			
	0.0732	0.0009*	0.2498			
	0.0121	0.0009*	0.5213			
	0.0346	0.0004*	0.5846	Horngren		
Kimmel	1.0000	0.0566	0.1307	0.1121		
	0.7628	0.0070	0.0301	0.1866		
	0.8357	0.1868	0.0204	0.0150		
	0.6642	0.0258	0.0619	0.0566	Kimmel	
Needles	0.6918	0.0260	0.3854	0.3642	0.6100	
	0.0074	0.0001*	0.4737	0.8949	0.0499	
	0.0115	0.0006*	0.5093	0.6921	0.0237	
	0.0414	0.0008*	0.5845	0.6104	0.1308	Needles
Phillips	0.5842	0.1804	0.1515	0.1006	0.5841	0.3550
	0.0702	0.0008*	0.4508	0.8950	0.1516	0.9850
	0.0795	0.0128	0.4068	0.4972	0.1515	0.5093
	0.3165	0.0069	0.4582	0.2995	0.4732	0.3855

FREAD
FGRADE
FOG
SMOG

0.05/21=0.0024
Bonferroni adjustment

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 18. Readability Comparisons of Topic Four between Textbooks (Wilcoxon Rank-Sum Test)

	Edmonds						
Harrison	0.0454				FREAD	FGRADE	
	0.0044						
	0.0087						
	0.0070	Harrison					FOG
Hermanson	0.0894	0.0135			SMOG	0.05/21=0.0024	
	0.0763	0.0014*					Bonferroni adjustment
	0.0128	0.0018*					
	0.0346	0.0016*	Hermanson				
Horngren	0.0893	0.3750	0.0055				
	0.2578	0.0363	0.0092				
	0.4283	0.0093	0.0098				
	0.2652	0.0135	0.0103	Horngren			
Kimmel	0.2131	0.6104	0.0436	0.9549			
	0.1412	0.1932	0.0274	0.4735			
	0.1363	0.4854	0.0088	0.1574			
	0.1410	0.1746	0.0150	0.3652	Kimmel		
Needles	0.1996	0.0088	0.7628	0.0073	0.0436		
	0.0301	0.0003*	0.4064	0.0055	0.0062		
	0.0248	0.0003*	0.7486	0.0098	0.0044		
	0.0544	0.0002*	0.8504	0.0092	0.0053	Needles	
Phillips	0.2345	0.3267	0.0183	0.5706	0.8062	0.0286	
	0.4280	0.0033	0.6647	0.1132	0.0594	0.2578	
	0.1465	0.0024*	0.4855	0.0862	0.0204	0.6922	
	0.2577	0.0033	0.4506	0.0731	0.0287	0.3855	

* Significant at the 0.05 level with a Bonferroni adjustment (i.e., amounts ≤ 0.0024)

Table 19: Spearman's Correlations of the Four Readability Evaluation Methods*

Text	Entire Sample			Edmonds		
Method	FREAD	FGRADE	FOG	FREAD	FGRADE	FOG
FGRADE	-0.86881 <.0001			-0.91654 <.0001		
FOG	-0.80903 <.0001	0.90464 <.0001		-0.82922 <.0001	0.86918 <.0001	
SMOG	-0.85717 <.0001	0.92450 <.0001	0.95164 <.0001	-0.86330 <.0001	0.92228 <.0001	0.95370 <.0001
Text	Harrison			Hermanson		
Method	FREAD	FGRADE	FOG	FREAD	FGRADE	FOG
FGRADE	-0.89775 <.0001			-0.93195 <.0001		
FOG	-0.77429 <.0001	0.80640 <.0001		-0.85308 <.0001	0.89922 <.0001	
SMOG	-0.80600 <.0001	0.87767 <.0001	0.94545 <.0001	-0.88747 <.0001	0.91635 <.0001	0.95365 <.0001
Text	Horngren			Kimmel		
Method	FREAD	FGRADE	FOG	FREAD	FGRADE	FOG
FGRADE	-0.88868 <.0001			-0.88574 <.0001		
FOG	-0.84896 <.0001	0.90191 <.0001		-0.76068 <.0001	0.88921 <.0001	
SMOG	-0.88820 <.0001	0.93113 <.0001	0.96033 <.0001	-0.81196 <.0001	0.92216 <.0001	0.92862 <.0001
Text	Needles			Phillips		
Method	FREAD	FGRADE	FOG	FREAD	FGRADE	FOG
FGRADE	-0.90285 <.0001			-0.84530 <.0001		
FOG	-0.85885 <.0001	0.92939 <.0001		-0.76443 <.0001	0.90418 <.0001	
SMOG	-0.89825 <.0001	0.93421 <.0001	0.96618 <.0001	-0.88363 <.0001	0.88683 <.0001	0.92056 <.0001

* All correlations are significant at the 0.05 level.