Stellenbosch Papers in Linguistics, Vol. 46, 2016, 155-179
doi: 10.5774/46-0-205

Using readability, comprehensibility and lexical coverage to evaluate the suitability of an introductory accountancy textbook to its readership

Gerrit-Jan Wissing

Directorate: Teaching and Learning with Technology, Tshwane University of Technology, Pretoria West Campus, South Africa
E-mail: wissingg@tut.ac.za

A Seugnet Blignaut

TELIT-SA, Faculty of Economic Sciences and Information Technology, North-West University, Vaal Triangle Campus, South Africa
E-mail: Seugnet.Blignaut@nwu.ac.za

Karien van den Berg

School of Languages, North-West University, Potchefstroom Campus, South Africa
E-mail: karien.hattingh@nwu.ac.za

Abstract
At universities, textbooks are still a primary source of course content. However, this can only be efficacious if the intended readers are able to comprehend the content of the textbooks adequately. This study investigated three possible approaches to determining whether the intended readership of a prescribed Introductory Accountancy textbook (Cornelius & Weyers 2011) will be able to make meaning of that textbook. Such an investigation has important implications for authors, publishers of textbooks and subject lecturers prescribing the texts. Readability of the textbook was determined by using the Flesch Reading Ease and Flesch-Kinkaid Grade Level indices, as well as the average of five conveniently calculated grade level reading indices. A Cloze procedure test was administered to a selection of students to determine their reading comprehension of a reading text. Finally, Nations' Vocabulary Size Test (Nation and Beglar 2007: 9, 11) was used to determine whether the vocabulary size of the selection of students provides adequate lexical coverage of the lexis used in the textbook to enable comprehension of the text. The findings were somewhat conflicting. The readability indices, and to a lesser extent the vocabulary size test, indicated suitability of the textbook to its intended readership. The Cloze test results suggested contradictory findings that users of the textbook will be reading at their frustration level. These conflicting findings are discussed.

Keywords: readability, reading comprehension, vocabulary size, higher education

1. Introduction

In academic and other contexts, textbooks are used as a primary source of course content, and courses are often conveniently structured around prescribed textbooks (Cline 1972: 33; Jones 2011: 29; McFall 2005: 72; Phillips and Phillips 2007: 25). Students, while regarding textbooks as valuable to the learning process are fearful that textbooks will be too complicated. They rely heavily on content delivered during lectures, referring to textbooks mostly when they still struggle with material after attending lectures (Jones 2011: 31; Phillips and Phillips 2007: 29). For students to be empowered by the content of their textbooks, they have to extract meaning from the content communicated by the textbooks (Smith and Taffler 1992: 84; Snyman 2004: 15).

In these textbooks, meaning is mostly conveyed by the vehicle of language. However, first-year students at the Tshwane University of Technology (TUT) have poor English literacy skills (Dockrat 2007: 11), owing to the fact that they are often not English First Language (EFL) speakers. It follows then that these students might have difficulty mastering the learning content provided in their textbooks when that content is expressed in the English language.

Authors of prescribed textbooks often gear their writing towards finding peer approval, rather than meeting the instructional needs of students (Cline 1972: 34). Typically the factors considered when selecting a textbook include (i) the pedagogical approach followed by the author(s); (ii) how well the required course material is covered and organised; (iii) illustrations and exhibits included; (iv) supplementary materials; and (v) the facilitator's previous experience with the textbook (Plucinsky, Olsavsky and Hall 2009: 119). However, authors such as Adelberg and Razek (1984: 109–110), Plucinsky et al. (2009: 119) and Smith and DeRidder (1997: 367) suggest that the ability of students to understand the learning content of the textbook should play a far greater role in textbook selection. In order to provide the ability to make meaning from text content with the consideration it deserves in the textbook selection process, factors that influence reading comprehension – such as readability, comprehensibility, the reader's knowledge of the vocabulary used in the text and the interrelationship of these factors with the reader's ability to make meaning of the texts – must be understood better.

To address this issue, this article reports one such an investigation. The aim of the study was to determine whether the textbook prescribed for a university module titled *Accounting for Marketers* presented at the TUT is written at a level that will enable the readers of the textbook to make meaning of the text.

The main research question that guided this investigation is: How do readability, understandability and readers' lexical coverage interrelate as measures for determining the suitability of a prescribed textbook to its intended readership?

This research question was operationalised in terms of three sub-questions:

1. To what extent is the prescribed textbook sufficiently readable as measured by a selected number of readability formulae?
2. To what extent is the prescribed textbook sufficiently understandable with reference to the scores achieved by its intended readership in a Cloze test drawn from the content of the text book?

3. How can the readers' lexical coverage of the lexis used in a textbook be used as indication of the readers' ability to adequately make meaning of the prescribed textbook?

2. Prior research

Previous studies have mainly relied on measures of readability and comprehensibility to assess the suitability of textbooks for the target readers. After a distinction is drawn between readability and comprehensibility, a synthesis of some of these earlier studies is provided, highlighting the strengths and weaknesses of the two approaches. A case is then made for investigating the appropriateness of lexical coverage as an alternative approach to determining the suitability of an academic text to its target readership.

2.1 Readability and comprehensibility

While Adelberg and Razek (1984: 109) do not distinguish between readability and comprehensibility (or understandability), others (Chiang, Englebrecht, Phillips and Wang 2008: 48; Jones 1997: 105; Smith and Taffler 1992: 85) make such a distinction, pointing out that while the two attributes are closely related, they are intrinsically different. Similarly, Smith and Taffler (1992: 85) suggest that it is not justified to treat 'readability' and 'understandability' as synonyms as there is a marked, measureable difference between the two concepts. Elucidating this difference, Wray and Dahlia (2013: 73) explain readability as a characteristic of the text itself and understandability as an indication of the readers' ability to make meaning of the text. While a text has to be readable to be understandable, comprehensibility is not only affected by syntactical difficulty, but also by "reader characteristics such as the reader's background, prior knowledge, interest and general reading ability" (Jones 1997: 106).

2.1.1 Readability

Readability refers to the linguistic characteristics of a text, which impacts the ease or difficulty with which a reader will be able to read and understand the text. Readability is distinct from legibility, the latter referring to the actual ease with which a text can be read. The readability level of a text is an indicator of the textual difficulty level of the text and the suitability of the text to readers of particular age groups or grade levels. It is fixed for a given text and is not varied by reader characteristics (Chiang et al. 2008: 48; Jones 1997: 105–106; Lee and French 2011: 694; McLaughlin 1969: 640; Plucinsky et al. 2009: 119).

Word difficulty and familiarity, along with sentence length, may be useful as indicators of reading difficulty. The difficulty of individual words used in a text influences a reader's ability to understand the text. Word difficulty depends on the length and familiarity of a word. The basic assumption is that longer, less familiar words are harder to read than shorter ones, though there are exceptions such as technical words that may be short, but unfamiliar. Word familiarity relates to a word's ranking in word frequency lists. A relatively large proportion of English text is made up of a relatively small number of English words, meaning that these words are very familiar. Frequency of use varies between different nationalities and different age groups, which consequently reduces the value of word frequency as indicator of word familiarity. In addition to these two factors, sentence difficulty also impacts readability. As a rule, longer sentences are harder to read than shorter ones. However, shorter sentences may contain concepts that are complex and difficult to understand, while longer sentences may provide more

helpful clues to the meaning being conveyed. Cohesion and coherence within a text may aid readability for those readers with sufficient reading skills (Klare 1974: 97–98; Stevens, Stevens and Stevens 1992: 368–369; Wray and Dahlia 2013: 74–76).

The extent to which the writer shares meaning with the reader can be enhanced when the writer takes the lexical, textual and background knowledge of the reader into consideration while composing the text (Snyman 2004: 16). However, this may be near impossible with texts published and prescribed globally. Therefore, the selection of the text must be conducted thoughtfully. At the same time, care should be taken not to oversimplify language in an attempt to improve readability. While shorter sentences and words with fewer syllables are considered easier to read, simple language might not foster the development of complexity in mental models where such complexity is necessary to deal with complex situations and course content. Absence of sentence complexity in the prescribed texts read by students may also have a negative impact on students' ability to convey complexities in their own writing (Davidson 2005: 71–72; Lee and French 2011: 695).

Spinks and Wells (1993) recommend that readability should be a prime measure for textbook selection. While there may be other influencing factors besides readability, academic performance and student retention decline as textbooks become more difficult to read. Peterson (1982: 2) found a significant relationship between text readability and academic achievement in Accounting. He concludes that readability may be used to predict which students might experience academic difficulty in technical subjects such as Accounting.

Davison and Kantor (1982: 189, 191) warn that readability formulae do not define actual readability and should not be used unguardedly as actual readability is not a simple function of objectively measureable properties such as word and sentence length. Syntax (sentence length and grammatical complexity) and semantics (difficulty of words measured in number of syllables) are commonly used in calculating indices of readability, but these calculated indices of readability take no account of whole-text aspects and reader characteristics such as skill, motivation and experience (Bargate 2012: 5; Chiang et al. 2008: 48–49; Sydes and Hartley 1997: 143; Sydserff and Weetman 1999: 459).

Readability is also influenced by a number of subjective factors, such as "the explicitness of connection between clauses, the extrasentential, pragmatic factors of discourse and sentence topic and focus, the inference load placed on a reader, the epistemological status of statements, and finally, the appropriateness of vocabulary for a particular audience reading with limited background knowledge" (Davison and Kantor 1982: 189, 191).

Courtis (1995:6), Fry (1989: 294), and McConnell and Paden (1983: 66) add that concept density, level of abstraction, complexity of ideas, extent to which these ideas are reinforced through repetition or restatement, the effect of the author's writing style on reader's interest and motivation, use of active voice, use of illustrations, and a number of other factors all have an influence on readability. They also mention the inappropriateness of using readability formulae where understanding relies heavily on whether or not readers are familiar with subject-specific terminology.

While all these points of criticisms are valid, the formulae remain useful for predicting a reader's reading comprehension, oral reading errors and willingness to carry on reading.

Readability formulae have been well researched as being indicative of whether a text will be understood by its intended readership and should be used, in conjunction with other factors, to aid in textbook selection (Courtis 1995: 6; Fry 1989: 294–296).

There are more than 200 objective, valid, easily administrable and repeatable reading indices (Chiang et al. 2008: 50; Fry 1989: 294; Lee and French 2011: 695; Shuptrine and Moore 1980: 397, 400). Reading indices such as the Flesch Reading Ease and Flesch-Kinkaid Grade Level indices are included in some word processing packages, making calculation easy. These indices are calculated with reference only to the average sentence length of a text and the average number of syllables per word of the text. The indices offer a pragmatic approach to determining a single, summary average readability score without requiring access to information of the characteristics of the eventual readers of a text (Courtis 1995: 6; Fry 1989: 294–296).

The Flesch Reading Ease index in particular is often used or referred to in research into the readability of texts (Bargate 2012: 9; Jones and Smith 2014: 191). This index scores the readability of text samples within a range of 0 to 100. Text with a readability score of 0 would be very difficult to read, while text with a readability score of 100 would be very easy to read. Text with a readability score of 90 to 100 indicates that a reader, who has completed Grade 4, should be able to correctly answer 75% of comprehension questions set over the text. Every 10 points of the scale increases the grade level of the text by one grade up to about Grade 7. Thereafter, grade levels are impacted to an increasing degree (Flesch 1948: 225).

The Flesch-Kincaid Grade Level index (FKGL) uses a simplified formula to directly predict the grade level for which the text is suitable (Kincaid, Fishburne, Rodgers and Chissom 1975: 19). In order to interpret her findings, Bargate (2012: 14) used readability scales adapted for use in South African educational grade levels. This adapted scale is displayed in Table 1, which also includes the type of text typically written at that readability level (Flesch 1948: 230).

Table 1: Seven-point General Reading Ease scale adapted for South Africa

Flesch Reading Ease Index	Reading level	Text type
90 – 100	Very easy (Grade 5)	Comics
80 – 89	Easy (Grade 6)	Pulp fiction
70 – 79	Fairly easy (Grade 7)	Slick fiction
60 – 69	Standard (Grades 8 and 9)	Digests
50 – 59	Fairly difficult (Grades 10 to 12)	Quality
30 – 49	**Difficult (Undergraduate)**	**Academic**
0 – 29	Very difficult (Postgraduate)	Scientific

Table 1 summarises characteristics of the text in terms of degrees of reading level. These levels are reflective of the reader's ability to make meaning of texts at each of these levels.

Much criticism has been levelled against the use of readability indices as indicator of readers' ability to understand a text. The main objection is that these indices measure qualities of the text, and not qualities of the reader (Bargate 2012: 5; Chiang et al. 2008: 48–49; Sydes and Hartley 1997: 143; Sydserff and Weetman 1999: 459). However, this study still included readability, as the aim was to establishing how readability indices compare with other measures of comprehensibility when evaluating the suitability of a text for its intended readership.

Objections against using measures of readability can possibly be addressed by using a measure of comprehensibility such as the Cloze test[1]. Cloze tests are used widely to measure reading comprehension objectively, reliably, validly, and with due consideration for reader qualities (Adelberg and Razek 1984: 109; Bormuth 1969: 358–363; Stevens, Stevens and Stevens 1993: 291; Taylor 1957: 20). Comprehensibility and the use of Cloze tests to measure comprehensibility is discussed next.

2.1.2 Comprehensibility

Readability, which is essentially fixed for a given text, contributes to, but is not equal to understandability of the text, which can vary among readers of the same text. For a text to be understood by a reader, it has to be readable by that reader. A text being readable does not guarantee that it will be understandable, although understandability of a text can at least partially be predicted by readability indices. (Flory, Phillips and Tassin 1992: 152; Jones 1997: 105–106; Plucinsky et al. 2009: 119; Smith and Taffler 1992: 94). Adelberg and Razek (1984: 109) define understandability as "the ability of readers to comprehend … textbooks and to complete the act of communication initiated by the writers of those textbooks." According to De Vos and Raepsaet (2010: 5) a text is understandable when the receiver receives the message as intended by the sender. Given this distinction between readability and understandability, it is necessary to think twice about using readability measures as indicators of understandability, as readability may not be directly related to understandability (Davidson 2005: 59; Smith and Taffler 1992: 85, 93).

Meyer (2003: 205) identifies four sets of variables that interact to influence understanding, viz. reader variables (such as verbal ability, word knowledge, education and age), strategy variables (such as structure strategy, rereading and underlining), text variables (such as structure, topic content, word familiarity and cohesion) and task variables (such as mode and rate of presentation, response mode and task requirements). Understandability concerns itself with the reader's ability to understand the content dealt with in a text, and is dependent on reader attributes such as the reader's background, prior knowledge, interests and reading skills. (Chiang et al. 2008: 48; Jones 1997: 105–106).

Comprehensibility measures are essential in ensuring appropriate text selection. To illustrate the importance of understandability of text, Wray and Dahlia (2013: 72) use the example of a test item with a readability level that exceeds the readers' reading ability. Such an item may no longer assess subject matter knowledge but rather their reading ability. Razek, Hosch and Pearl (1982: 23) point out that an easily understandable textbook enables independent self-study by students, thereby allowing for lecture time to be used for supplementary learning activities.

The Cloze procedure was initially introduced as a measure of readability, but its usefulness was soon extended to include application to understandability. In a Cloze test, a number of passages of equal length are selected from a text. Passages are then mutilated by deleting selected words and replacing the words with a standard sized blank space. The test is administered by requiring participants to guess the deleted words, gaining clues from words remaining in the passage. Where a high number of deletions are guessed accurately, the text is considered more understandable than

[1] Cloze procedure tests are constructed by deleting random words, significant words or every n^{th} word from a text paragraph, and replacing the deletions with spaces of equal length. Test subjects have to 'close' the gaps by guessing the missing words and inserting them into the blank spaces. The ability to correctly guess the missing words is thought to rely on the subject's ability to make sense of the remaining text in the paragraph (Taylor 1957: 19).

a text where only few deletions are guessed accurately (Taylor 1956: 81–83; Taylor 1957: 19–20). Where a distinction is made between readability and understandability, the Cloze procedure is superior to readability indices as indicator of text understandability (Jones 1997: 106). By testing reading skills, Cloze tests require reader-text interaction and so overcome many of the objections against readability indices, which rely on syntactic and vocabulary features of text as indicators of understandability (Bargate 2012: 7–8; Jones 1997: 106; Smith and Taffler 1992: 87; Taylor 1957: 20). A Cloze test measures reading comprehension objectively, reliably and validly (Adelberg and Razek 1984: 109; Bormuth 1969: 358–363; Stevens et al. 1993: 291).

Jones (1997: 106), among others, is critical of using Cloze procedures, contending that they do not necessarily measure reading comprehension; that validating results of Cloze tests against readability indices is problematic; that there is a lack of consensus about interpreting scores meaningfully; and that using Cloze tests for technical texts presents challenges. Flory et al. (1992: 152) also argue that Cloze tests are difficult to administer and time consuming for research subjects, possibly leading to researchers using only a small number of passage selections. Too small a number of passage selections may not be representative of the entire text, especially where more than one author contributed to the text. In refuting this argument, Stevens et al. (1993: 290–291) point out that three randomly selected passages are sufficiently representative of a text. Furthermore, it should be noted that some software applications such as Blackboard Learn™ now offer Cloze-type questions as a standard feature, making compilation and administration of Cloze tests somewhat easier.

An important aspect to be considered when using Cloze tests is the interpretation of scores. There is need for a frame of reference when relating a Cloze score to corresponding scores in a reading comprehension test. The rule of thumb for oral reading texts at the *instructional* level, suited to supervised textbook-based instruction, is that a student should be able to score at least 75% in a comprehension test covering the text. For *independent* level texts (reference and voluntary reading) the student should be able to score 90% in such a comprehension test. The related Cloze scores are 44% and 57%, respectively. Cloze scores of 43% and lower characterises understanding at the reader's *frustration* level and indicate that the text is too difficult for students (Bargate 2012: 16; Bormuth 1968b: 196; Bormuth 1971: 147). Rankin and Culhane (1969: 197) have determined corresponding required Cloze scores for instructional and independent level texts at 41% and 61%, respectively. In her study, Bargate (2012: 16) used Bormuth's (1968b: 196) guidelines to interpret the results of her Cloze test. These guidelines are set out in Table 2.

Table 2: Cloze comprehension levels

Cloze score	Level
0% – 43%	*Frustration level* – language is too difficult for readers to cope with
44% – 57%	*Instructional level* – readers able to cope, but some assistance required
58% – 100%	*Independent level* – readers able to cope with the language

While there is little consensus about how understandability should be measured, most recent research in readability of Accounting texts has focused on Cloze procedures. However, Cloze procedures may not measure understandability, but rather only the ability to infer missing words correctly (Jones and Smith 2014: 184–187; Jones 1997: 118). These and other limitations in current approaches to measuring understandability, such as the difficulty to administer Cloze

procedures (Flory et al. 1992: 152), require investigation into an alternative approach to assessing understandability. Direction might be found in the consistently strong positive correlation found between vocabulary knowledge and reading comprehension, almost irrespective of the research design (Stahl 2003: 241).

2.2 Vocabulary knowledge and reading comprehension

A strong, but not necessarily causal relationship exists between vocabulary knowledge and reading comprehension (Hu and Nation 2000: 404; Qian 1999: 299), to the extent that a reader's knowledge of words used in a text is the leading predictor of the reader's ability to understand that text (Stahl 2003: 241). Laufer (2013: 869–871) indeed suggests determining the difficulty of text for a reader with reference to the proportion of words used in that text which the reader understands (lexical coverage). This could be done by measuring the reader's vocabulary size, compiling frequency lists for the textbook and then determining the reader's lexical coverage of the prescribed textbook by expressing vocabulary size as a percentage of the number of words used in the text. This is somewhat similar to the approach followed by Nation (2006: 79) and Anderson (2013: 61), although both these studies simply measure the number of word families in a text against the benchmark of 8 000 to 9 000 words established by Hu and Nation (2000: 422) as the vocabulary size required by a typical reader to be able to understand a text. If the vocabulary of the students using the textbook adequately matches the word tokens used in the selected textbook, it is likely to aid those students' understanding of the content, in the process of assisting them to improve the level of their language skills and their academic performance.

Numerous studies have found a strong positive correlation between vocabulary size and reading comprehension (Baleghizadeh and Golbin 2010: 33; Carroll, Bowyer-Crane, Duff, Hulme and Snowling 2011: 2; National Institute of Child Health and Human Development 2000: 2.12). Stahl (2003: 241) adds that this correlation is usually stronger than 90%, and that the difficulty of the words used in a text is the foremost determinant of the difficulty of that text (Stahl 2003: 246). While empirical evidence of a causal relationship between vocabulary size and reading comprehension is not yet conclusive (Lubliner and Smetana 2005: 189; National Institute of Child Health and Human Development 2000: 4.15), the body of evidence which suggests that a student's ability to comprehend a text is influenced by the size of the student's vocabulary seems to be expanding (Stanovich 1986: 379).

Vocabulary knowledge is a multi-faceted construct. Qian (2002: 514–516) surveys a number of authors' criteria for *knowing* a word in proposing four dimensions of word knowledge, namely vocabulary breadth (or size), vocabulary depth, lexicon organisation and automaticity of receptive or productive knowledge. Studies of vocabulary have primarily focused on breadth of knowledge, referring to the number of words of which the meaning is at least superficially known, and depth of word knowledge, referring to how well a word is known. Qian (1999: 299) suggests a strongly positive association and interdependence between the breadth and depth dimensions of vocabulary knowledge.

In connection with depth of word knowledge, Nation's (2001: 27) model identifies form, meaning and use as general aspects of knowing a word. In the context of second language learning, Laufer, Elder, Hill and Congdon (2004: 206–207) differentiate between four degrees of word knowledge based on two distinctions. This classification is set out in Table 3.

Table 3: Types of vocabulary knowledge

	Recall	**Recognition**
Active or productive knowledge (retrieval of the correct word form)	Active recall	Active recognition
Passive or receptive knowledge (retrieval of the correct meaning of the word)	Passive recall	Passive recognition

A word is known actively (productively) when with or without a cue the correct L2 form of an L1 word can be retrieved. It is known passively (receptively) when an L2 word is provided and the L1 meaning can be retrieved. A word is recalled when its form or meaning can be provided, and recognised when its form or meaning can be selected from a set of options (Laufer et al. 2004: 206).

Research has mostly focused on the relationship between the number of words known (*breadth of vocabulary knowledge*) and reading comprehension, as it is easier to measure vocabulary breadth than to measure how well a word is known (*depth of vocabulary*). Instruments measuring breadth are better developed and perhaps as a consequence, more studies have explored the relationship between vocabulary breadth and reading comprehension (Qian 2002: 517). For practical reasons, then, this study focuses on the breadth dimension of vocabulary, hereafter referred to as vocabulary size.

Laufer and Ravenhorst-Kalovski (2010: 15–19) explain how the number of words that can be understood out of context (sight vocabulary) determines the percentage of total running words or tokens in a text that a specific reader can understand (lexical coverage). The sight vocabulary size required for sufficient lexical coverage to adequately understand a typical academic text is referred to as the lexical threshold. The threshold is probabilistic, meaning it is possible for one with a smaller sight vocabulary and consequent lexical coverage to understand the text adequately, but it is not likely.

However, adequate understanding is not a clearly defined term and may vary depending on the context. One could relate *adequate comprehension* to the level of comprehension required to achieve Cloze scores of 43% and above for *instructional* level texts and 57% and above for *independent* level texts (Bargate 2012: 16; Bormuth 1968b: 196). According to Biemiller (2001: 1), students will comprehend the meaning of a text if they understand the meanings of at least 95% of the words making up that text. Laufer (1989: 319–321) supports this estimate, while Laufer and Ravenhorst-Kalovski (2010: 15) suggest using 8 000 word families yielding 98% lexical coverage as an optimal threshold where *adequate comprehension* is meant to be synonymous with *independent comprehension.* They set 4 000 to 5 000 word families, yielding 95% coverage, as a minimal threshold where *adequate comprehension* is meant to mean *reading with some guidance and help.* Schmitt, Jiang and Grabe (2011: 26) as well as Hu and Nation (2000: 414–415) also find the 98% estimate more appropriate than 95%. In their study, Hu and Nation (2000: 414–415) defined *adequate comprehension* as the understanding required to score about 85% in a reading comprehension test using a fiction text where lexical coverage was 100%. They predict that most readers would already achieve this level of adequate unassisted comprehension where lexical coverage was 98%, which can be achieved at a probabilistic sight vocabulary threshold of 8 000 to 9 000 word families including proper nouns for written text (Hu and Nation 2000: 422; Nation 2006: 59). Regardless of whether one accepts the 95% or the 98% estimate, there certainly appears to be a strong relation between knowing the meaning of words used in a text, and comprehending that text.

Krashen's (2009: 21) input hypothesis furthermore proposes that students will improve their knowledge of a language when they are exposed to texts (input) that are just a little beyond their current ability (i+1) in that language. Krashen (2009: 21) refers to such texts as comprehensible input. Having existing knowledge of a sufficient number of words used in a text to adequately comprehend that text will allow students to derive the meaning of the unknown words from the context in which they appear. Improving vocabulary in such a way is strongly associated with improved reading comprehension. In terms of textbook selection, one could then recommend that chosen texts contain between 95% and 98% of words that are familiar to students to enable them to make meaning of the content itself, including the unknown words comprising the remaining 5% to 2% of the text.

In this section, three different approaches to evaluate the suitability of a prescribed Accountancy text book to its intended readership were discussed in detail. These approaches are (i) readability of a text as measured using a selection of readability indices; (ii) comprehensibility of the text as measured using a Cloze test; and (iii) readers' lexical coverage as determined with reference to the readers' vocabulary size and the lexis used in the text. Authors such as Davidson (2005: 59), and Smith and Taffler (1992: 93) advise against using readability indices as measures of understandability. Objections typically point out that they measure qualities of the text, and not qualities of the reader experience (Bargate 2012: 5; Chiang et al. 2008: 48–49; Sydes and Hartley 1997: 143; Sydserff and Weetman 1999: 459). Being widely accepted as measuring reading comprehension objectively, reliably and validly and with due consideration of reader qualities (Adelberg and Razek 1984: 109; Bormuth 1969: 358–363; Stevens et al. 1993: 291; Taylor 1957: 20), Cloze procedures have been touted as an alternative. A Cloze test has its own challenges, not the least of which concerns the difficulty level of administering Cloze tests (Flory et al. 1992: 152). A third alternative for determining the appropriateness of a textbook to its intended readership proposed by this article would be to match the readers' vocabulary size against the word tokens used in the prescribed textbook in an attempt to easily match reader characteristics to the challenges set to the reader by the text. This measurement, referred to as lexical coverage, has the potential to provide a more reliable yardstick with which to measure the readership's likelihood of being able to understand the meaning of the words used in the text and perhaps of the meaning of the text itself. The research methodology followed in the study is described next.

3. Research methodology

This paper provides a quantitative examination of the appropriateness of a specific prescribed text in terms of its readability, understandability and lexical coverage of students in the course for which the text is prescribed. As such, the design may be considered a case study.

The tools used were selected in order to show how the experiment could be repeated, using freely available electronic tools, by researchers without expert levels of linguistic knowledge. The following tools were used for the purpose of this study:

- Readability indices available from readability-score.com (Child 2014), read-able.com (Simpson 2013) and from within Microsoft Word™
- The Blackboard Learn™ *Fill in Multiple Blanks* question type, which allows for the construction of Cloze tests;

- Nation's Vocabulary Size Test (VST) (Nation, Chui, Chung, Nakata, Sasao, Quero et al. 2014) accessible from my.vocabularysize.com; and
- BNC 20k vocabulary profiler (Cobb 2013), available from lextutor.ca.

3.1 Participants in the study

All first-year students registered in 2013 for the National Diploma: Marketing at the TUT, with the exclusion of students registered for the extended curriculum programme, were invited to participate in the study. While *Accounting for Marketers* is a second-year subject, the first-year students were chosen as participants as the purpose of the study was to establish whether the textbook is suitable for use by students new to the subject. Second-year students might have encountered Accounting terminology in the classroom, which could have influenced test results. Participation in the study was voluntary. Students in the cohort who did not participate either chose not to participate, or were not present at the time the tests were administered.

3.2 Data collection

The assessment battery used in this study included Nation's VST (Nation and Beglar 2007), and a Cloze test based on text from the prescribed textbook. Students were allowed to complete the assessment battery at their own pace, but with an overall time limit of approximately two hours – the scheduled duration of the lecture period during which the assessments were administered. Participant results were organised and analysed according to student number. Participants were required to supply student numbers for the English Language Skills Assessment and the VST, and Blackboard Learn™, used for administering the Cloze test, automatically captured student numbers of participants. While comparative analyses of results required participants to be individually identifiable, the results were treated confidentially.

3.2.1 Passage selection

The prescribed textbook in question – *Accounting All-in-1* (Cornelius and Weyers 2011) – is used for the course *Accounting for Marketers* at the TUT. A digital copy of the text, in Microsoft Word™ format, was obtained from the publishers. Calculation of reading indices and construction of Cloze tests were facilitated by having access to an electronic copy of the text.

To allow for a comparison of readability and understandability across all chapters and between authors, the Cloze test selections were spread across text sampled from each of the 19 chapters in the textbook. A page was randomly chosen from each chapter, and a suitable paragraph selected from the page. Passages were selected randomly to ensure a representative sampling of different levels of textual difficulty within the textbook was examined. For a passage to be suitable for inclusion in the Cloze test it had to, apart from the first and last sentences, contain approximately 40 words as eight deletions of every fifth word was required. In cases where the selected page did not contain a suitable passage, another passage from a different page was chosen.

3.2.2 Readability

As previously mentioned, readability was included in the study to compare readability indices to indicators of understandability and lexical coverage as measure of readers' ability to make meaning of a text. Two readability indices were used: The Flesch Reading Ease index (FRE)

and the Flesch-Kincaid Grade Level index (FKGL). These indices were chosen following earlier, but similar studies (Bargate 2012; Chiang et al. 2008; Plucinsky et al. 2009). The indices are also convenient and easy to use, one measuring instrument being an embedded functionality of Microsoft Word™, and two others – Readability-Score.com and The Readability Test Tool – being freely available online (Child 2014; Simpson 2013). Using these two indices, readability was established for each of the 19 passages selected for the Cloze test.

In determining the indices using Microsoft Word™, no adjustments were made to the texts. A passage was simply selected, and the indices for the selected passage calculated. When calculating the indices using Readability-Score.com and The Readability Test Tool (Child 2014; Simpson 2013), certain minor text modifications were made in order to obtain the same result from both applications. Both these online tools provide a basket containing five grade level readability indices:

- Flesch-Kinkaid Grade Level (FKGL) index;
- Gunning-Fog Score index;
- Coleman-Liau index;
- Simple Measure of Gobbledygook (SMOG) index; and
- Automated Readability index.

An average grade level is automatically calculated by both applications from these five indices. As this average grade level is available with no additional effort it is also reported in the results, as it could provide a convenient alternative measure for evaluating the readability of a text.

3.2.3 Understandability

Cloze tests were chosen as measure of understandability for this study. While many questions about the validity of Cloze tests may be raised, evaluating the validity of Cloze tests and the assumptions underlying them are beyond the scope of this paper. The deletion pattern chosen – eight deletions of every fifth word from the passage selected from each of the 19 chapters – resulted in a total of 152 deletions, consistent with the number of deletions used by Bargate (2012: 13) and Baghaei (2011: 689).

The first sentence of each selected passage was left intact to provide context for the remainder of the passage (Bargate 2012: 13). A random number between one and five was used to determine the first deletion in the second sentence of the passage (Adelberg and Razek 1984: 113). Thereafter, every fifth word was deleted until eight deletions were made. Deleting every fifth word allows the greatest number of deletions possible per passage without compromising the reliability of the test. Increasing distance between deletions to more than five words has little effect on a reader's ability to restore deletions (Adelberg and Razek 1984: 113; Bargate 2012: 13; Bormuth 1968a: 432; Macginitie 1961: 129). According to Baghaei (2011: 688) more deletions per passage provide more reliable ability measures, but the ability scores themselves are not affected by the number of deletions.

The fifth word deletion pattern was only disrupted for duplicate words in the same passage (Blackboard Learn™ does not allow for duplications), proper nouns, amounts, and simple words, such as *an*, *the*, and *is*. In these cases, the immediately following word would be deleted.

The sentence following the one in which the eighth deletion occurred would be the last sentence of the passage, and would once again be left intact to provide context. A typical paragraph would look like this:

> **Sales returns and allowances**
> When customers purchase goods from a trading entity, there is always the possibility that they may not be satisfied with the goods they purchased. [(1) This] may be due to a [(2) number] of reasons; for example, the [(3) goods] may be incorrect or [(4) damaged]. The customer may then [(5) return] the goods to the [(6) entity] (sales returns). If the goods [(7) were] purchased for cash, the [(8) customer] will receive a cash refund and if the goods were purchased on credit, the customer's account will be credited by the entity. It is also possible that the customer may decide to keep the goods, albeit at a lower price (sales allowance).

In this example of a selected passage, the fifth word deletion pattern was disrupted for "a", "the", and "goods" (already selected in this passage). Word classes was not taken into account, because doing so would affect the objectivity and repeatability of the test. Furthermore, the measures used during this study was purposefully selected for not requiring specialised linguistic expertise to administer. This should make it easy for non-linguist subject matter experts to use in their own attempts to establish the suitability of texts for their own specific areas of study. Selecting every n^{th} word for deletion renders the test more objective and repeatable, and does not require specific linguistic competence.

The test was administered using the *Fill in Multiple Blanks* question type featured in the Blackboard Learn™ learning management system. When presented, deletions are replaced by evenly sized blank text boxes, which do not provide any clue to the length of deleted words (Bargate 2012: 13; Culhane 1970: 412). The Cloze scores were interpreted with reference to the levels described in Table 2: Cloze comprehension levels. Other than for minor misspellings, only exact replacements were accepted. Allowing synonyms would be cumbersome: it would require manual assessment of each test submission to evaluate whether the synonym was a suitable alternative to the exact word. An automated assessment of viable alternatives would similarly require all assessments to be examined for acceptable synonyms, so that these alternatives could be included in the software's marking rubric. However, previous studies have shown that the additional effort to allow synonyms would not lead to significantly different results (Bargate 2012: 15; Culhane 1970: 412; Hartley 2004: 931; Taylor 1957: 22).

3.2.4 Vocabulary size

The instrument used in this study to measure word knowledge is the online version of Nation's VST (Nation et al. 2014). This standardised instrument reliably, accurately and comprehensively measures receptive recognition of the 14 000 most frequently used English word families in the British National Corpus (BNC), and requires a moderately developed understanding of a word's full meaning for the word to be included in the measured vocabulary size (Nation and Beglar 2007: 9, 11). The test consists of 10 multiple choice items per 1 000-word list for a total of 140 items. The score achieved is multiplied by 100 to estimate the number of known word families in the participant's vocabulary.

Next, the number of word families used in the relevant texts were determined using a BNC 20k vocabulary profiler (Cobb 2013). The profiler produces a report of the cumulative percentage of word tokens in the text drawn from each group of thousand words in the BNC 20k list – starting from the 1 000 most frequently used words (K-1) and progressing to the 1 000 least frequently used words (K-20). The group of 1 000 words where the required 95% to 98% lexical coverage required for understanding the text is reached, indicates the approximate vocabulary size required to be able to read the text with adequate comprehension (Laufer and Ravenhorst-Kalovski 2010: 15) or independent comprehension (Hu and Nation 2000: 414–415). In order to be able to compare the readers' vocabulary size to the vocabulary required to make meaning of the text, this approach relies quite heavily on vocabulary being acquired sequentially, as Biemiller (2001: 2) suggests, from the most frequently used words to the least frequently used words.

4. Results and discussion

Results obtained based on the above-mentioned measurements are reported and discussed consecutively.

4.1 Readability scores

The results of the FRE, FKGL and the average score for a basket of the five other indices are shown in Table 4. The table shows the range of readability scores as well as the mean scores and standard deviation for the 19 passages. The scores were calculated using Microsoft Word™ (Legend = W), readability-score.com and read-able.com (Legend = R). The average score of the basket of readability scores calculated on readability-score.com and read-able.com is also shown. Readability of the passages are then discussed firstly with reference to FRE, then with reference to FKGL and finally with reference to the average of the basket of indices.

Table 4: Readability scores

	Tool	Flesch Reading Ease	Flesch-Kinkaid Grade Level
Range	W	18.3 – 64.7	8.7 – 18.7
	R	17.7 – 68.2	9.8 – 18.9
	Basket	8.8 – 19.1	
Mean	W	43.7 Difficult (Undergraduate)	12.2 Fairly difficult (Grades 10 – 12)
	R	46.2 Difficult (Undergraduate)	12.2 Fairly difficult (Grades 10 – 12)
	Basket	12.9 Fairly difficult (Grades 10 – 12)	
Standard deviation	W	12.4	2.1
	R	13.3	2.3
	Basket	2.5	
Legends	W	Microsoft Word™	
	R	Reconciled readability-score.com and read-able.com	
	Basket	Reconciled readability-score.com and read-able.com (basket of 5 indices)	

When using Microsoft Word™ to determine FRE for the 19 passages, readability ranged from 64.7 (*Standard*, suitable for Grades 8 and 9 students) to 18.3 (*Very difficult*, suitable only for

postgraduate students). The average of the values was 43.7 (*Difficult*, suitable for undergraduate students) with a standard deviation of 12.4, indicating the relatively wide variability within the calculated values. Corresponding values for the 19 passages were calculated using readability-score.com and read-able.com, the range being from 68.2 (*Standard*, suitable for Grades 8 and 9 students) to 17.7 (*Very difficult*, suitable only for postgraduate students). The average of the values was 46.2 (*Difficult*, suitable for undergraduate students), with a standard deviation of 13.3, again indicating the relatively wide dispersion of the calculated values.

When calculating FKGL, Microsoft Word™ shows readability to range from 8.7 (*Standard*, suitable for Grades 8 and 9 students) to 18.7 (*Very difficult*, suitable only for postgraduate students). Average readability was 12.2 (*Fairly difficult*, suitable for students in Grades 10 to 12) with a standard deviation of 2.1 grade levels. Scores calculated using readability-score.com and read-able.com ranged from 9.8 (*Standard* to *Fairly difficult*, suitable for Grades 9 to 10 students) to 18.9 (*Very difficult*, suitable only for postgraduate students). Average readability was also 12.2 (*Fairly difficult*, suitable for undergraduate students), the standard deviation from mean being 2.3 grade levels.

The basket of Grade Level indices confirms the FRE and FKGL measurements. The index basket average ranges from 8.8 (*Standard*, suitable for Grades 8 and 9 students) to 19.1 (*Very difficult*, suitable only for postgraduate students) with an average readability of 12.9 (*Fairly difficult*, suitable for students in Grades 10 to 12) and a standard deviation of 2.5 grade levels.

The readability analysis seems to indicate that the prescribed textbook will be suitable to undergraduate students. There were some instances where readability measured at a *very difficult* level, suitable only for postgraduate students. In these instances some rewriting might be required to make the text more accessible to the target readership of undergraduate students. From a purely readability point of view, this rewriting would entail reducing average sentence length and using words with fewer syllables.

4.2 Cloze scores

Results of the Cloze procedure test are shown in Table 5.

Table 5: Cloze scores

n = 58	**Score**	**Average level (24.6%)**
Range	2.0% – 53.3%	**0% – 43%: Frustration level** Language is too difficult for readers to cope with.
Mean	24.6%	
Standard deviation	13.2%	
n(<44%) = 54 (93%)	n(44% to 56%) = 4 (7%)	n(>56%) = 0 (0%)

As can be seen from Table 5, when reader abilities are taken into account with reference to Table 2: Cloze comprehension levels, it would appear as if the prescribed textbook is written at a level at which the vast majority of the target readership cannot make adequate meaning of the content. The mean score of 24.6% is well below what is required for readers to be able to cope with the text. Only four students would be able to cope with the content of the text if some assistance were provided – such as by a lecturer in a classroom – while the language used in

the text would be too difficult for the remaining readers to cope with. None of the students tested would be able to read the text independently.

Results of the Cloze test illustrate, as a number of studies have cautioned (Sydes and Hartley 1997: 143; Sydserff and Weetman 1999: 459; Wray and Dahlia 2013: 79–84), the difficulty in attempting to establish whether a text is suitable for readers using readability measures that do not take reader characteristics into account. This difficulty with readability measures is aggravated where the intended audience does not have the reading skills normally associated with their particular grade level, as might well be the case in this study. Such an explanation would be consistent with a study by Dockrat (2007) in which she reported that only 5% of the 2007 student intake at the TUT had English literacy skills at the level of Grade 12 and above.

4.3 Vocabulary size

The selected passages contained 1 781 word tokens in total. Of these words, 94.43% falls within the first 3 000 (K-1 to K-3) most used English words from the British National Corpus (BNC), while 97.58% falls within the first 4 000 (K-1 to K-4) most used words. This implies that a vocabulary size of 3 000 to 4 000 word families should be sufficient to achieve the lexical coverage of 95% and larger suggested by Biemiller (2001: 1) as being necessary to make meaning of the passages selected for the Cloze test. By the same reasoning, a vocabulary of 4 000 to 5 000 word families is required to achieve 95% coverage of the 45 776 word tokens in the book as a whole. This finding is consistent with an estimate by Laufer and Ravenhorst-Kalovski (2010: 15) of the vocabulary size required to be able to read a text "with some guidance and help".

Hu and Nation (2000: 414–415), as well as Laufer and Ravenhorst-Kalovski (2010: 15), set the lexical coverage required for independent comprehension at 98%, which in their study was achieved at a vocabulary size of 8 000 to 9 000 word families. In this study, a vocabulary size of 4 000 to 5 000 word families is required to achieve such comprehension for the Cloze passages, and 6 000 to 7 000 words for the book as a whole.

When examining the measured vocabulary sizes of the test group, the mean vocabulary size was 6 769 word families, with a standard deviation from the mean of 1 518 word families. Measured values were dispersed over a range from 4 100 families to 10 900 families. The lowest vocabulary size measured of 4 100 families should, for this book, provide the 95% coverage necessary to read the text with some assistance (Biemiller 2001: 1; Laufer and Ravenhorst-Kalovski 2010: 15), while the average vocabulary of 6 769 should provide the 98% lexical coverage required for being able to read the textbook independently (Hu and Nation 2000: 414–415).

Table 6 contains the word frequency profile for the text used in the passages selected for the Cloze test as well as for the textbook as a whole. It also shows descriptive data in respect of the VST administered during the experiment.

Table 6: Word frequencies and vocabulary sizes

Frequency level	Selected passages				Complete textbook			
	Families	**Types**	**Tokens (%)**	**Cumulative token %**	**Families**	**Types**	**Tokens (%)**	**Cumulative token %**
K-1 words	231	303	1 450 (81.41%)	81.41%	526	90	34 280 (74.89%)	74.89%
K-2 words	68	82	191 (10.72%)	92.13%	265	385	6 524 (14.25%)	89.14%
K-3 words	24	26	41 (2.30%)	94.43%	110	129	893 (1.95%)	91.09%
K-4 words	19	20	56 (3.14%)	**97.58%***	85	112	1701 (3.72%)	94.81%
K-5 words	8	9	15 (0.84%)	**98.42%***	41	55	777 (1.70%)	**96.50%***
K-6 words	3	4	16 (0.90%)	99.32%	29	33	638 (1.39%)	97.90%
K-7 words	1	1	2 (0.11%)	99.43%	15	17	295 (0.64%)	**98.54%***
K-8 words	1	1	1 (0.06%)	99.49%	17	17	43 (0.09%)	98.63%
K-9 words	1	1	5 (0.28%)	99.77%	6	8	452 (0.99%)	99.62%
K-10 words					6	6	87 (0.19%)	99.81%
K-11 words					3	3	4 (0.01%)	99.82%
K-12 words	1	1	1 (0.06%)	99.83%	3	3	13 (0.03%)	99.85%
K-13 words					1	1	1 (0.00%)	99.85%
K-14 words					2	2	2 (0.00%)	99.86%
K-15 words								
K-16 words					1	1	5 (0.01%)	99.87%
K-17 words					1	1	3 (0.01%)	99.87%
K-18 words								
K-19 words								
K-20 words								
Off-list words	3	3	3 (0.17%)	100.00%	16	16	58 (0.13%)	100.00%
Totals	360	445	1 781 (100%)		1127	879	45 776 (100%)	

Vocabulary size test (n = 42)							
Vocabulary size	**4 000 to 5 000 words**	**5 000 to 6 000 words**	**6 000 to 7 000 words**	**7 000 to 8 000 words**	**8 000 to 9 000 words**	**9 000 to 10 000 words**	**10 000 to 11 000 words**
Distribution	n = 3	n = 11	n = 9	n = 12	n = 5	n = 0	n = 2
Mean **Range** **Std. deviation**	6 769 word families 4 100 – 10 900 word families 1 518 from mean						

* Percentages in bold indicate at what K-level the 95% and 98% coverage levels are achieved, both for the selected passages and for the textbook as a whole.

The results of matching readers' vocabulary size to the vocabulary used in a text as indicator of the ability of readers to adequately comprehend that text, produced unexpected results. According to the results of this test, the textbook is suitable for the intended readership, perhaps with some support from lecturers. This largely contradicts the results of the Cloze test, which found the text to be too difficult for its intended readership. A possible explanation comes from a more detailed analysis of the vocabulary size test results, as indicated in Table 7.

Table 7: Vocabulary size per thousand words

Vocabulary size test (n = 42)															
Frequency level		**K-1**	**K-2**	**K-3**	**K-4**	**K-5**	**K-6**	**K-7**	**K-8**	**K-9**	**K-10**	**K-11**	**K-12**	**K-13**	**K-14**
Mean %		78.1	71.4	67.4	60.0	60.0	44.0	40.7	52.9	34.0	31.9	35.5	41.0	23.8	32.1
Range	**Low**	30	30	10	10	0	10	10	20	0	0	0	10	10	10
	High	90	100	100	90	100	80	70	100	80	70	90	60	50	70
Std. deviation		13.0	20.3	21.8	21.0	22.7	18.7	15.6	18.3	16.9	16.7	20.0	11.7	11.7	16.7

This table shows the average percentage of words known by participants for every grouping of 1 000 words from the first 14 000 most frequently used words from the BNC (K-1 to K-14). It also shows the lowest and highest percentage achieved per 1 000 words, and the standard deviation per 1 000 words. Strongly sequential vocabulary acquisition would be indicated by high average word knowledge for early groups of 1 000 most used words from the BNC, and low average word knowledge for the later groups. From this table it seems as if the vocabulary size of the test group does not display the strong evidence of being sequentially acquired that Biemiller (2001: 2) has found. While more words are known from the more frequently used groupings, the highest average percentage of known words per grouping of 1 000 words is 78.1%, which is well short of the 95% to 98% required for adequate comprehension (Biemiller 2001: 2; Hu and Nation 2000: 414–415; Laufer and Ravenhorst-Kalovski 2010: 15). The progression is also not linear. For example, K-8 shows a larger average word knowledge than K-6 and K-7.

Ideally, the three measures used in this study to establish readers' ability to make meaning of a text would have provided congruence in their results. Unfortunately, this proved not to be the case. The study did not find a definitive approach to establishing the comprehensibility of a text book. However, the results provided indication of the possible direction future studies have to take in order to provide congruence between measures of readability and understandability, as indicators of readers' ability to understand a text. These are discussed in the following section.

5. Limitations of the study and areas for further research

Readability indices were determined with reference to grade levels established in the USA. These grade levels might not be appropriate to levels generally encountered in the context in which this study was conducted, where students do not necessarily possess the ELS normally expected for a specific grade level. Research should be undertaken to establish grade levels more appropriate to the context of the study (Dockrat 2007: 11).

A relatively small sample size (n: Cloze tests=58; n: VST=42), selected from only one subject at one University of Technology was used in the study. The study should be undertaken using a larger sample from a more diverse readership of the textbook to improve the generalisability of the research findings related to the Cloze test and the VST.

A number of participants, when completing the Cloze test, filled in meaningless answers (e.g. *kk*, or *ergtt*) for some of the deletions. While this could be interpreted as the student legitimately not knowing the specific answer, it could also be an indication of the participant wanting to get the test over and done with without really trying to guess the correct word. If the latter explanation is the case, it would have an impact on the validity of the test results.

In line with previous studies (Adelberg and Razek 1984: 113; Bargate 2012: 13; Bormuth 1968a: 432), this study has followed the practice of deleting every fifth word when developing the Cloze procedures test. This practice traces its origin back to a study by Macginitie (1961: 129), confirmed by Alderson (1979), who found little positive effect of a context – the distance between deletions – of more than about five words. However, a context of five words is achieved by deleting every sixth word. Using a context of five words might have improved the results of the Cloze tests. A future study of this nature could be conducted to assess the impact of using a deletion rate of every sixth word.

This study did not show the strong sequential order in which vocabulary is acquired that Biemiller (2001: 2) has found. The order in which English vocabulary is acquired by students similar to the participants in this study should be investigated. Once this sequence is established, a future study using readers' vocabulary size as predictor of readers' ability to understand a text might be of great value.

6. Conclusion

The usefulness of textbooks to students is conditional upon the students' ability to understand the contents of those textbooks (Smith and Taffler 1992: 84; Snyman 2004: 15). The present study examined three measurements viz. readability indices, comprehensibility and lexical coverage for their usefulness to gauge the suitability of a prescribed text to its intended readers' abilities.

The results of this study were contradictory in that two measures – readability and vocabulary size – point to the textbook being appropriate to its intended readership of undergraduate students newly entering into higher education while the third measure – understandability – seems to indicate that the readership may be reading the textbook at their frustration level – ideally a textbook should allow for independent study (Bormuth 1968b: 196).

It is unsurprising that results should differ between readability measures on the one hand and understandability measures on the other hand, as they measure different things. Readability formulae measure qualities of text, while understandability measures reflect reader characteristics, specifically the readers' ability to interact meaningfully with text (Jones 1997: 105). Furthermore, readability formulae such as the Flesch formulae were developed about 70 years ago in the USA (Flesch 1948: 221), while the population in this study are South Africans with poor English literacy skills (Dockrat 2007: 11). The specific readability formulae used might therefore not be valid in the South African context.

However, one would have expected a closer match between the results of the assessment of suitability of a text to the intended readership when using understandability and lexical coverage, as both these measures are determined with reference to reader characteristics. The discrepancy can possibly be explained by the fact that general vocabulary acquisition for the test group, outlined in Table 7, was not as strongly sequential in the order that Biemiller (2001: 2) suggests, resulting in a vocabulary measurement which does not strongly predict lexical coverage of the text sufficient for adequate comprehension.

It might be more appropriate to use a vocabulary size test made up of test items drawn from the word families used in the specific text, rather than a test of general vocabulary size such as Nation's VST (Nation et al. 2014). The lexical coverage of a text determined for a specific student might then more closely reflect students' understanding of the meaning of that text. This should be investigated in a further study. Further investigation into the order in which vocabulary is acquired by participants from similar contexts as in this study should also be considered, as a clearer understanding of this order would enable authors to better suit vocabulary used in textbooks to the vocabulary of target readers. It could also be worthwhile to conduct a comparative study for a given text between a Cloze test and a comprehension test standardised for readers such as the participants in this study. Such a comparison might give indication of the validity of a Cloze test for use in similar contexts as measure of understandability of the text.

That vocabulary acquisition did not seem to occur in the same strong sequential order as previously believed for this group of participants, has important implications for classroom practice. Attention should be paid to direct priming vocabulary instruction of not only the Academic Word List and subject-specific jargon, but also, to a larger extent than previously considered necessary, of K1 and K2 words. Such an approach to expanding vocabulary would assist readers to gain better lexical coverage of the lexis used in their prescribed text books, and aid them in improving their understanding of the content thereof.

References

Adelberg, A.H. and J.R. Razek. 1984. The Cloze procedure: a methodology for determining the understandability of accounting textbooks. *The Accounting Review* 59(1): 109–122.

Alderson, J.C. 1979. The effect on the doze test of changes in deletion frequency. *Journal of Research in Reading* 2(2): 108–119.

Anderson, M. 2013. *The Readability and Usability of Building User Guides*. Wellington, NZ: Victoria University – Master (Building Science)).

Baghaei, P. 2011. Do C-Tests with Different Number of Gaps Measure the Same Construct? *Theory and Practice in Language Studies* 1(6): 688–693.

Baleghizadeh, S. and M. Golbin. 2010. The effect of vocabulary size on reading comprehension of Iranian EFL learners. *Linguistic and Literary Broad Research and Innovation* 1(2): 33–46.

Bargate, K. 2012. The readability of managerial accounting and financial management textbooks. *Meditari Accountancy Research* 20(1): 4–20.

Biemiller, A. 2001. Teaching Vocabulary: Early, Direct, and Sequential. *American Educator* 25. Available online: http://www.aft.org/newspubs/periodicals/ae/spring2001/biemiller.cfm (Accessed 6 January 2012).

Bormuth, J.R. 1968a. The Cloze Readability Procedure. *Elementary English* 45(4): 429–436.

Bormuth, J.R. 1968b. Cloze Test Readability: Criterion Reference Scores. *Journal of Educational Measurement* 5(3): 189–196.

Bormuth, J.R. 1969. Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability. *Reading Research Quarterly* 4(3): 358–365.

Bormuth, J.R. 1971. *Development of Standards of Readability: Toward a Rational Criterion of Passage Performance*. *Final Report.* Illinois: University of Chicago.

Carroll, J.M., C. Bowyer-Crane, F.J. Duff, C. Hulme and M.J. Snowling. 2011. Theoretical Framework: Foundations of Learning to Read. In J.M. Carroll, C. Bowyer-Crane, F.J. Duff, C. Hulme and M.J. Snowling. *Developing Language and Literacy: Effective Intervention in the Early Years*. Chichester, UK: John Wiley and Sons, Ltd. pp. 1–16).

Chiang, W.-C., T.D. Englebrecht, T.J. Phillips Jr and Y. Wang. 2008. Readability of financial accounting principles textbooks. *The Accounting Educators' Journal* 18(2008): 48–80.

Child, D. 2014. *Readability-Score.com*. Available online: https://readability-score.com/ (Accessed 24 April 2014).

Cline, T.A. 1972. Readability of Community College Textbooks. *Journal of Reading* 16(1): 33–37.

Cobb, T. 2013. *The Compleat Web VP!* Available online: http://www.lextutor.ca/vp/bnc/ (Accessed 24 April 2014).

Cornelius, L. and M. Weyers, M. 2011. *Accounting All-in-1*. 2nd ed. Johannesburg: LexisNexis.

Courtis, J.K. 1995. Readability of annual reports: Western versus Asian evidence. *Accounting, Auditing and Accountability Journal* 8(2): 4–17.

Culhane, J.W. 1970. CLOZE Procedures and Comprehension. *The Reading Teacher* 23(5): 410–464.

Davidson, R.A. 2005. Analysis of the complexity of writing used in accounting textbooks over the past 100 years. *Accounting Education* 14(1): 53–74.

Davison, A. and R.N. Kantor. 1982. On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly* 17(2): 187–209.

De Vos, M. and S. Raepsaet. 2010. *Taalkundige Analyse van de Complexiteit in Accountinghandboeken*. Gent: University of Gent – Masters in Business Economics.

Dockrat, S.Y. 2007. *Language Profile: January 2007 Intake* (Internal Presentation). Pretoria: Tshwane University of Technology.

Flesch, R. 1948. A new readability yardstick. *Journal of Applied Psychology* 32(3): 221–233.

Flory, S.M., T.J. Phillips Jr and M.F. Tassin. 1992. Measuring readability: A comparison of accounting textbooks. *Journal of Accounting Education* 10(1): 151–161.

Fry, E.B. 1989. Reading Formulas: Maligned but Valid. *Journal of Reading* 32(4): 292–297.

Hartley, J. 2004. Designing instructional and informational text. In M. Spector, M.D. Merrill, J. Elen and M.J. Bishop (Eds.) *Handbook of research on educational communications and technology*. pp. 917–945.

Hu, H.-C.M. and P. Nation. 2000. Unknown Vocabulary Density and Reading Comprehension. *Reading in a Foreign Language* 13(1): 403–430.

Jones, J.J. 2011. Enhancing Student Learning: An Examination of the Student Use of Textbooks in Financial Accounting. *American Journal of Business Education* 4(1): 29–36.

Jones, M. and M. Smith. 2014. Traditional and alternative methods of measuring the understandability of accounting narratives. *Accounting, Auditing and Accountability Journal* 27(1): 183–208.

Jones, M.J. 1997. Methodological themes: Critical appraisal of the Cloze procedure's use in the accounting domain. *Accounting, Auditing and Accountability Journal* 10(1): 105–128.

Kincaid, J.P., R.P. Fishburne Jr, R.L. Rogers and B.S. Chissom. 1975. *Derivation of new readability formulas* (automated readability index, fog count and Flesch reading ease formula) for navy enlisted personnel: DTIC Document.

Klare, G.R. 1974. Assessing Readability. *Reading Research Quarterly* 10(1): 62–102.

Krashen, S.D. 2009. *Principles and Practice in Second Language Acquisition*. Available online: http://sdkrashen.com/Principles_and_Practice/index.html. (Accessed 5 October 2011)

Laufer, B. 1989. What percentage of text-lexis is essential for comprehension. In C. Laurén and M. Nordman (Eds.) *Special language: From humans thinking to thinking machines.* pp. 316–323.

Laufer, B. 2013. Lexical Thresholds for Reading Comprehension: What They Are and How They Can Be Used for Teaching Purposes. *TESOL Quarterly* 47(4): 867–872.

Laufer, B., C. Elder, K. Hill and P. Congdon. 2004. Size and strength: do we need both to measure vocabulary knowledge? *Language testing* 21(2): 202–226.

Laufer, B. and G. Ravenhorst-Kalovski. 2010. Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language* 22(1): 6.

Lee, S. and N. French. 2011. The readability of academic papers in the Journal of Property Investment and Finance. *Journal of Property Investment and Finance* 29(6): 693–704.

Lubliner, S. and L. Smetana. 2005. The Effects of Comprehensive Vocabulary Instruction on Title I Students' Metacognitive Word-Learning Skills and Reading Comprehension. *Journal of Literacy Research* 37(2): 163–200.

Macginitie, W.H. 1961. Contextual Constraint in English Prose Paragraphs. *The Journal of Psychology* 51(1): 121–130.

McConnell, C. and D.W. Paden. 1983. Readability: Blind Faith in Numbers? *Journal of Economic Education* 14(1): 65–71.

McFall, R. 2005. Electronic textbooks that transform how textbooks are used. *The Electronic Library* 23(1): 72–81.

McLaughlin, G.H. 1969. SMOG Grading-a New Readability Formula. *Journal of Reading* 12(8): 639–646.

Meyer, B.J. 2003. Text coherence and readability. *Topics in Language Disorders* 23(3): 204–224.

Nation, I. 2001. *Learning vocabulary in another language*. Ernst Klett Sprachen.

Nation, I. 2006. How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/La revue canadienne des langues vivantes* 63(1): 59–82.

Nation, P. and D. Beglar. 2007. A vocabulary size test. *The Language Teacher* 31(7): 9–13.

Nation, P., J. Chiu, T. Chung, T. Nakata, Y. Sasao, B. Quero, et al. 2014. *VocabularySize.com*. Available online: http://my.vocabularysize.com/ (Accessed 24 April 2014).

National Institute of Child Health and Human Development. 2000. *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC.

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*. Oregon: Oregon State University – PhD.

Phillips, B.J. and F. Phillips. 2007. Sink or Skim: Textbook Reading Behaviors of Introductory Accounting Students. *Issues in Accounting Education* 22(1): 21–44.

Plucinsky, K.J., J. Olsavsky and L. Hall. 2009. Readability of Introductory Financial and Managerial Accounting Textbooks. *Academy of Educational Leadership Journal* 13(4): 119–127.

Qian, D.D. 1999. Assessing the Roles of Depth and Breadth of Vocabulary Knowledge in Reading Comprehension. *Canadian Modern Language Review* 56(2): 282.

Qian, D.D. 2002. Investigating the Relationship Between Vocabulary Knowledge and Academic Reading Performance: *An Assessment Perspective. Language Learning* 52(3): 513–536.

Rankin, E.F. and J.W. Culhane. 1969. Comparable Cloze and Multiple-Choice Comprehension Test Scores. *Journal of Reading* 13(3): 193–198.

Razek, J.R., G.A. Hosch, and D. Pearl. 1982. Readability of Accounting Textbooks. *The Journal of Business Education* 58(1): 23–26.

Schmitt, N., X. Jiang and W. Grabe. 2011. The percentage of words known in a text and reading comprehension. *The Modern Language Journal* 95(1): 26–43.

Shuptrine, F.K. and E.M. Moore. 1980. Even After the Magnuson-Moss Act of 1975, Warranties Are Not Easy to Understand. The Journal of Consumer Affairs 14(2): 394.

Simpson, D. 2013. *The Readability Test Tool*. Available online: http://read-able.com/ (Accessed 24 April 2014).

Smith, K.J. and J.J. DeRidder. 1997. The selection process for accounting textbooks: general criteria and publisher incentives–a survey. *Issues in Accounting Education* 12(2): 367–384.

Smith, M. and R. Taffler. 1992. Readability and understandability: different measures of the textual complexity of accounting narrative. *Accounting, Auditing and Accountability Journal* 5(4): 84–98.

Snyman, M.E. 2004. Using the printed medium to disseminate information about psychiatric disorders. *South African Psychiatry Review* 7(4): 15–20.

Spinks, N. and B. Wells. 1993. Readability: A textbook selection criterion. *Journal of Education for Business* 69(2): 83–87.

Stahl, S.A. 2003. Vocabulary and Readability: How Knowing Word Meanings Affects Comprehension. *Topics in Language Disorders* 23(3): 241–247.

Stanovich, K.E. 1986. Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly* 21(4): 360–407.

Stevens, K.C., K.T. Stevens and W.P. Stevens. 1993. A response to "measuring readability: A comparison of accounting textbooks". *Journal of Accounting Education* 11(2): 287–292.

Stevens, K.T., K.C. Stevens and W.P. Stevens. 1992. Measuring the Readability of Business Writing: The Cloze Procedure Versus Readability Formulas. *Journal of Business Communication* 29(4): 367–382.

Sydes, M. and J. Hartley. 1997. A Thorn in the Flesch: Observations on the Unreliability of Computer-Based Readability Formulae. *British Journal of Educational Technology* 28(2): 143–145.

Sydserff, R. and P. Weetman. 1999. A texture index for evaluating accounting narratives: An alternative to readability formulas. *Accounting, Auditing and Accountability Journal* 12(4): 459–488.

Taylor, W.L. 1956. Recent Developments in the Use of "Cloze Procedure". *Journalism and Mass Communication Quarterly* 33(1): 42–99.

Taylor, W.L. 1957. "Cloze" readability scores as indices of individual differences in comprehension and aptitude. *Journal of Applied Psychology* 41(1): 19–26.

Wray, D. and J. Dahlia. 2013. Exploring the Readability of Assessment Tasks: The Influence of Text and Reader Factors. *REMIE: Multidisciplinary Journal of Educational Research* 3(1): 69–95.