

## Readability of professional accounting examinations – evidence from South Africa

Masibulele Phesa <sup>a</sup>, Tesfaye T. Lemma <sup>b</sup> and Reon Matemane <sup>c</sup>

<sup>a</sup>Discipline of Accounting, Economics & Finance, University of KwaZulu-Natal, Durban, South Africa;

<sup>b</sup>Department of Accounting, Towson University, Towson, United States of America; <sup>c</sup>Department of Financial Management, University of Pretoria, Pretoria, South Africa

### ABSTRACT

The South African Institute of Chartered Accountants (SAICA), the preeminent professional accounting body in South Africa, administers two staggered sequential qualifying examinations – the Initial Test of Competency (ITC) and the Assessment of Professional Competency (APC) – to licence candidates as chartered accountants. Analysing the readability of the exams using the Gunning Fog Reading Index (GFRI) and Flesch Readability Ease Index (FREI), we find that both exams are difficult to read. We also find that the ITC exams are more difficult to read than the APC exams. These findings hold even when different parts of the exam are considered separately. While a difficult level of readability might be necessary to at least some extent due to the technical nature of the content being assessed, our findings raise concerns about potential construct-irrelevant variance (CIV), which could impact the validity, fairness, and reliability of the qualifying exams. Considering that a substantial portion of the candidates are non-native English speakers, these findings highlight the potential role of readability in ensuring equitable access to the chartered accountancy profession.

### ARTICLE HISTORY

Received 17 April 2023  
Revised 11 December 2023;  
2 September 2024;  
14 December 2024;  
13 February 2025  
Accepted 12 March 2025

### KEYWORDS

Professional accounting examinations; examination readability; construct-irrelevant variance; examination validity and reliability; access to chartered accountancy; South African Institute of Chartered Accountants

## Introduction

This study examines the readability of the Initial Test of Competency (ITC) and Assessment of Professional Competency (APC) exams, the two qualifying examinations administered by the South African Institute of Chartered Accountants (SAICA).<sup>1</sup> Specifically, it investigates whether these exams are difficult to read, which may potentially affect examinees' comprehension and performance. It also explores whether significant differences exist between the readability of the ITC and APC exams, and if they exist, the nature of those differences.

Accounting researchers have long sought to identify factors that influence performance in professional accounting examinations (e.g. Espahbodi et al., 2023; Koh & Koh, 1999; Rodrigues et al., 2018). While exam readability has been shown to impact exam performance (Lamb, 2010; Walkington et al., 2018), studies specifically investigating

**CONTACT** Masibulele Phesa  PhesaM@ukzn.ac.za  Discipline of Accounting, Economics & Finance, University of KwaZulu-Natal, Durban, South Africa

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

the readability of professional accounting examinations remain scant. This is a significant void, given that readability has direct implications for validity, fairness, and reliability in high-stakes assessments (Plake, 1988), particularly in contexts where the language used as the medium in the exams is not necessarily the examinees' first language. Although this study focuses on SAICA's qualifying exams, its findings are relevant to accounting educators and professional bodies worldwide.

We analysed the readability of SAICA's ITC and APC exams using data drawn from examinations administered between 2014 and 2022. Employing the Gunning Fog Reading Index (GFRI) and Flesch Readability Ease Index (FREI), we sought to establish whether the exams are difficult to read. The aim was also to evaluate whether our findings would align with the expectation that professional accounting exams are likely to impose a high cognitive load on examinees (Sackett et al., 2001), as they tend to assess advanced knowledge and skills using complex scenarios, problem-solving tasks, and in-depth analyses (Barac, 2012), which in turn might increase the difficulty of reading the exams. Such findings would be consistent with the notion that professional licensure examinations, as gatekeepers to professional practice, deliberately entail a high cognitive load (Dawrant & Han, 2021; Petchauer, 2015) and are therefore likely to be less readable.

While the ITC exam focuses on assessing a candidate's technical expertise through complex calculations, the APC exam employs a scenario-based approach to evaluate a candidate's ability to make judgements and decisions (Maroun, 2017). This difference in focus between the two qualifying examinations is likely to have ramifications for their readability. In this regard, Sweller (1988) suggests that exams requiring quantitative calculations and problem-solving, due to their step-by-step calculations and application of logical reasoning, are likely to impose a higher cognitive load on candidates. Thus, we posit that the ITC exam's focus on core technical proficiency, involving complex calculations, is likely to render it less readable than the APC exam which focuses on the application of concepts and principles in a defined field to form judgements and make decisions. Consistent with this argument, our findings, based on both the GFRI and FREI, show that the ITC exams are more difficult to read than the APC exams.

Recent years have witnessed continuing and extreme shortages of qualified chartered accountants (CAs) in South Africa, partly due to the exodus of qualified accountants leaving the country in search of greener pastures and the decline in pass rates at the qualifying exams (Businesstech, 2021; Lubbe et al., 2020; Ramalepe, 2022; Ryan, 2019). Prior studies on professional accounting certification exams identified factors such as a candidate's age and gender, number of exam attempts, first language, attendance at board courses, and textbook reading habits as key determinants of exam performance (Franklin et al., 2017; Roos, 2009; van Wyk, 2011). This study aims to add to this corpus of knowledge by evaluating the readability of the Initial Test of Competency (ITC) and Assessment of Professional Competency (APC) exams – the two qualifying exams administered by SAICA.

Readability refers to the ease with which written text can be understood (Oakland & Lane, 2004). When an exam is difficult to read, it risks assessing not only the intended construct but also a candidate's reading ability (Badgett, 2010). This phenomenon, known as construct-irrelevant variance (CIV), arises when extraneous factors, such as reading difficulty, affect assessment outcomes, leading to an inaccurate measurement of the intended competency (Cronbach, 1980; Plake, 1988). CIV is widely recognised

as a significant threat to validity, particularly in constructed-response assessments that require candidates to engage with rich, contextualised information, highlighting the critical role of readability in such exams (Arthur & Everaert, 2012; Godev et al., 2002; Zhai et al., 2021).

Professional accounting licensure exams are likely to be complex, as they assess technical expertise through calculations and problem-solving (Barac, 2012). The challenge, therefore, is to ensure that readability does not become an unnecessary barrier, while maintaining the rigor necessary to evaluate professional competence. In contexts where examinees exhibit varying levels of English proficiency, exams with excessive reading difficulty may disproportionately disadvantage candidates whose first language differs from the language of the exam, raising concerns about fairness and accessibility (Shealy & Stout, 2012). Thus, in high-stakes professional certification exams, such as those administered by SAICA, it is important to minimise reading difficulty to ensure the assessments measure only the intended constructs (Woo et al., 2009).

Our decision to focus on the readability of SAICA's two qualifying exams was underpinned by three key factors. First, a recent survey conducted on behalf of Chartered Accountants Worldwide shows that CAs in South Africa are global leaders in terms of professional trust (SAICA, 2023a), which contributes to South Africa's global competitiveness (World Competitiveness Center [WCC], 2023). Nonetheless, the continually decreasing number of newly qualified CAs, coupled with the exodus of already qualified CAs searching for better opportunities elsewhere, threatens South Africa's position as a global leader in accounting and auditing (Stephaniewix, 2017). We contend that a better understanding of the readability of SAICA's qualifying exams would lay the foundation for exploring readability as a possible driver of low pass rates, which will enlighten policy debates on the continuing shortage of CAs in South Africa.

Second, although the number of examinees of Black African origin – many of whom are not native English speakers – has grown substantially in recent years (SAICA, 2019),<sup>2</sup> their performance in the qualifying exams remains disappointingly low (Mofokeng & Mosala, 2021). SAICA's own data indicate that Black African candidates perform significantly worse than their White counterparts in the ITC exam, with a performance gap as high as 43 percent (SAICA, 2023b).<sup>3</sup> The relative 'closure' of access to the highly exclusive chartered accountancy profession in South Africa to people of Black African ancestry (Barac, 2015) remains a challenge to the country's socio-economic transformation agenda. While language proficiency is not the sole determinant of exam success, prior studies suggest that it plays a non-trivial role in the underperformance of Black African candidates (Barac, 2015; Weil & Wegner, 1997). Investigating the readability of SAICA's qualifying exams could therefore provide additional insights into the impediments Black African candidates face in navigating SAICA's certification process.

Third, exam readability plays a critical role in shaping student performance globally (see, for example, Lamb, 2010; Rezaee & Norouzi, 2011; Walkington et al., 2018). While numerous studies examine the readability of accounting textbooks (Bargate, 2012; Chiang et al., 2008; Flory et al., 1992; Plucinski, 2011; Plucinski & Hall, 2012; Razek et al., 1982; Traugh et al., 1987), research on the readability of professional accounting exams remains scarce. A notable exception is Phillips et al. (2007), who analysed consistency of readability in IMA exams. The literature highlights the importance of readability, especially in exams that involve constructed-response questions (Arthur &

Everaert, 2012; Godev et al., 2002). This study addresses the lacuna in the literature by investigating the readability of SAICA's qualifying exams. Considering SAICA's reliance on constructed-response formats, its qualifying exams offer a suitable context for further investigation. We hope the findings in our study will stimulate research into the antecedents and consequences of the readability of professional accounting certification and licensure examinations elsewhere in the world.

### **Contributions**

This study contributes to the literature in three ways. First, it adds to the existing literature on the readability of texts by investigating the readability of high-stakes professional licensure examinations using a unique setting in which the majority of examinees are non-native speakers of English, the language used as the examination medium. Other professional accounting examinations in which most examinees are non-native English speakers (e.g. the ACCA) may benefit from this study's findings. Second, the present study provides additional insight into the academic and policy debate on examinees' overall dismal performance in SAICA's qualifying exams. By ascertaining their reading difficulty, this study highlights the potential prevalence of construct-irrelevant variance (CIV) in these exams. Thus, this study's findings shed light on a potential antecedent of examinee performance on exams. Third, the findings also provide input to the policy debate on ways to ameliorate the relative 'closure' of access to the chartered accountancy profession for Black African candidates in South Africa.

The rest of this paper is organised as follows. The section that follows presents a synthesis of the background literature and hypotheses. Section 3 outlines the study's empirical framework, and Section 4 discusses the findings. Section 5 concludes the paper.

## **Prior studies and hypotheses development**

### **Prior studies**

Recent statistics indicate a decline in examinees' performance on SAICA's qualifying exams (Ryan, 2019). Despite this trend, only a few studies have explored the antecedents of success on these exams. van Wyk (2011) identified age, marital status, frequency of exam attempts, and access to preparatory courses as key determinants of performance, in Part I of SAICA's qualifying exams.<sup>4</sup> In a related study, Kraus (2019) demonstrated that a blend of asynchronous online and in-person tuition in the Certificate in Theory of Accounting course enhances examinees' performance in Qualifying Exam – Part I (QE1). In another study, Kruger (2020) reported that both instructors and students perceive open-book exam format as beneficial to examinee performance, though examinees of Black African ancestry, relative to their White counterparts, found the format less advantageous (see also, Zamanian & Heydari, 2012). Other studies highlight the positive impact of financial and academic support, such as the Thuthuka project<sup>5</sup>, in improving the performance of Black African examinees in the qualifying exams (de Jager, 2014; Ontong et al., 2020).

Exams that are difficult to read can impede examinee comprehension, especially for those whose first language differs from language in which the exam is administered

(Xia et al., 2019). This challenge introduces construct-irrelevant variance (CIV), where factors unrelated to the intended construct – such as language proficiency – affect assessment outcomes, thereby undermining exam validity (Cronbach, 1980; Plake, 1988). CIV is regarded as a crucial threat to validity, particularly in constructed-response assessments that incorporate rich contextualised information (Zhai et al., 2021). In this regard, van Wyk (2011) found that English being a candidate's first language was a significant predictor of success in SAICA's qualifying exams, suggesting language proficiency may confound performance measurement. Similarly, based on a study of students' reading comprehension in a financial reporting course at a South African university, van Rensburg et al. (2014) recommended that implicit reading comprehension instruction be provided to students who speak English as a second language. Supporting this, Coetzee et al. (2016) document significant differences in reading comprehension, driven by prior academic performance, the alignment between the language of instruction and the student's first language, and enrolment in the Thuthuka programme. Despite such evidence on the role of language skills in exam performance, no study has explored the readability of SAICA's exams. This study addresses this gap by evaluating the readability of both the ITC and APC examinations.

### *Understanding the concept of readability*

Although readability is not a precisely defined construct (Loughran & McDonald, 2014), it is typically associated with comprehensibility, or the level of difficulty 'with which a reader can read and understand' a given text (Oakland & Lane, 2004, p. 244). Some scholars define readability in terms of document's suitability for its intended audience and context (Howes et al., 2014; Sand et al., 2012), while others emphasise the role of writing style (DuBay, 2004; Klare, 1963). Readability has also been linked with the structure, texture, and content density of a written text (Amiran & Jones, 1982). Conceptualisations anchored in the linguistic attributes of a text associate readability with the use of plain English attributes, such as the use of short sentences and active voice, and the avoidance of jargon (Bonsall et al., 2017). According to these conceptualisations, a written text with complex English attributes (e.g. passive voice, hidden verbs, superfluous words, legal and financial jargon, numerous defined terms, abstract words, and lengthy sentences, as well as unreadable design and layout) is difficult to comprehend (Bonsall et al., 2017).

Highlighting the importance of considering the targeted reader in determining readability, McLaughlin (1969, as cited in DuBay, 2004) conceptualised readability as 'the degree to which a given class of people find certain reading matter compelling and comprehensible' (p. 3). Aligned with this view, Davison and Kantor (1982) underscored the importance of 'background knowledge assumed in the reader' in assessing the readability of a text (p. 187). Combining conceptualisations anchored on linguistic characteristics of a text and those which emphasise the attributes of the targeted audience, Dale and Chall (1949) defined readability as 'the sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimum speed, and find it interesting' (p. 23). According to Tekfi (1987), readability

entails ‘ensuring that a given piece of writing reaches and affects its audience in the way the author intends’ (262).

### *Hypotheses development*

High-stakes professional licencing examinations, including those administered by SAICA, are often designed to assess advanced knowledge and skills through complex scenarios, problem-solving tasks, and in-depth analysis (Barac, 2012). This complexity can contribute to higher cognitive load for examinees (Sackett et al., 2001). Furthermore, as high-stakes professional licensure exams are summative in nature and aim to assess candidates’ competence to practice professionally, the cognitive demands are intentionally higher (Dawrant & Han, 2021). This increased cognitive load is deliberate, as these exams are designed to regulate admission to a profession by distinguishing between candidates with varying levels of readiness (Petchauer, 2015). However, such heightened cognitive load may inadvertently introduce construct-irrelevant variance (CIV), where extraneous factors, such as reading difficulty, rather than the professional knowledge and skills of the candidate, may influence assessment outcomes (Plake, 1988). This concern is particularly pronounced in constructed-response assessments, where examinees must interpret and analyse context-rich information to formulate responses (Zhai et al., 2021).

Globally, professional accounting certification exams may comprise multiple-choice questions (MCQs) and/or constructed-response questions, with the latter requiring examinees to produce written responses such as essays (Arthur & Everaert, 2012). SAICA has adopted the constructed-response format for both the ITC and APC exams, aligning with the goal of simulating ‘the engagements and assignments that chartered accountants encounter in professional practice or commerce and industry’ (Barac, 2012, p. 50). In this vein, SAICA asserts that the ITC exam’s aim is to assess an examinee’s ability ‘to apply concepts and principles underpinning the defined field of study to the problems arising within the relevant practical domain’ (Barac, 2012, p. 50). Thus, the ITC exam is designed to assess examinees’ core technical proficiency (Maroun, 2017). On the other hand, the APC exam seeks to assess an examinee’s professional competency, using a single multidisciplinary case study developed based on real-life scenarios (SAICA, 2020). Constructed-response questions, however, require distinct skills compared to multiple-choice questions and may be particularly challenging for candidates with weaker reading comprehension or writing skills (Funk & Dickson, 2011). These challenges may exacerbate CIV, as reading difficulty – an extraneous factor which might not be related to the intended constructs – could unduly influence exam performance.

Validation, in the context of exams, refers to processes of establishing their reliability and validity, ensuring that the exams offer accurate and valuable information (Cook & Hatala, 2016). It is not merely a checkbox to tick but a cornerstone of responsible administration of assessments (Baker & Linn, 2002). Plake (1988) recommends that validation of readability checks be included in the validation process of high-stakes professional licensure examinations, a factor that can mitigate CIV. In this regard, SAICA asserts that its exams undergo several validation processes prior to being administered to

students (SAICA, 2020), potentially mitigating the reading difficulty encountered by examinees.

Considering the documented heterogeneity in examinees' reading ability (Coetzee et al., 2016; van Rensburg et al., 2014), the complexity of the scenarios used to assess higher-level understanding and competencies (Barac, 2012), and the constructed-response format used across the exams, which tend to impose a higher cognitive load on examinees (Funk & Dickson, 2011), we argue that SAICA's qualifying exams are likely to be difficult to read. As a result, the potential for CIV to undermine the validity of these assessments remains a concern. Thus, we state our first hypothesis as follows:

*Hypothesis 1: Both the ITC and APC exams are difficult to read.*

As noted earlier, the foci of the two qualifying exams are different. While the ITC exam focuses on candidates' technical expertise, the APC evaluates their professional aptitudes (Maroun, 2017; SAICA, 2014). Consequently, the ITC exam requires candidates to perform technical calculations, whereas the APC exam tends to require them to provide advice to management, or a business, based on a multidisciplinary case study involving a real-life decision-making scenario (Barac, 2012, p. 50; SAICA, 2020). Owing to the need for step-by-step calculations and the application of logical reasoning, exams with questions involving quantitative calculations and problem-solving are likely to impose a higher cognitive load on candidates (Sweller, 1988). On the other hand, essay-type questions allow for more interpretive and expansive thinking and the ability to articulate and organise thoughts (Sweller, 1988). Thus, we contend that the ITC exams' numerical complexity (due to the numbers, formulas, and mathematical operations involved) is likely to impose a higher cognitive load on candidates than the typical plain text questions (occasionally with simple calculations) encountered in the APC exams.

As alluded to earlier, readability is also shaped by the target audience's background knowledge and comprehension skills (Davison & Kantor, 1982; also see McLaughlin (1969) as cited in DuBay (2004)). The ITC exam is typically taken within a year of a candidate completing university-level academic training (Maroun, 2017). In contrast, the APC requires candidates to have passed the ITC, and subsequently complete at least 20 months of professional training (SAICA, 2014). As time passes between the ITC and APC exams, a candidate is likely to gain additional knowledge and the ability to better comprehend written texts. Therefore, we expect that the APC exam will impose a lower cognitive load and be more readable than the ITC exam. Thus, we propose our second hypothesis as follows:

*Hypothesis 2: The ITC exam is more difficult to read than the APC exam.*

## Methods

### *Sample and data*

This study focused on ITC and APC exams administered between 2014 and 2022. We limited our sample to 2014 and beyond because SAICA only administered the APC exam from 2014 onwards. We gathered papers for both exams from SAICA's official website ([www.saica.org.za](http://www.saica.org.za)). The ITC is administered twice a year, once at the beginning

of the year, and again at mid-year. It comprises four papers covering the financial accounting, cost and financial management, auditing, and tax functions of accounting (Visser & Steenkamp, 2017). Furthermore, the exam for each paper has two parts: the 'reading' (or the 'scenario') part and the 'required' part. The 'scenario' part provides the prompt for the question that the candidate is expected to answer and the 'required' section of the exam outlines the constructed response that the candidate is expected to provide.<sup>6</sup> We first merged the 'scenario' and 'required' parts of each exam paper and determined the readability score for the merged document. We then considered the average readability score of the four papers to be the average readability score of the ITC exam for that year.

Unlike the ITC exam, the APC exam is conducted only once a year, at the end of the year, and consists of a single paper in which candidates are presented with a multidisciplinary case study based on real-life business issues. The only deviation from this pattern occurred in November 2021, when SAICA encountered technical difficulties during a regularly scheduled exam, necessitating the administration of a replacement exam in March 2022. For this instance, only the readability score of the November 2021 exam was considered in our analysis. The APC exam is presented to the candidates over two separate days: the pre-release case study presents the business problem 'scenario', while additional information provided on the exam day provides further details about the company and the 'required' part of the exam. Here, too, we combined the information from both days and determined a readability score for the combined document, which was then used as the APC exam's readability score for that year.

### *Measuring readability*

The literature presents a plethora of readability measures (Loughran & McDonald, 2014) which are grouped into two approaches: quantitative and qualitative (Howes et al., 2014). Readability measures employing quantitative approaches use formulas to quantify a document's textual features (Benjamin, 2012). These approaches measure readability by computing lexical density (Eggins, 2004; Gholami et al., 2012), ease of reading (Flesch, 1948; Flesch & Gould, 1949), difficulty of documents that present information in matrix form (Mosenthal & Kirsch, 1998), and the number of years of schooling required to read a text (Kincaid et al., 1975). In contrast, qualitative approaches measure readability in a systematic and exploratory fashion (Berelson, 1952; Krippendorff, 2018) entailing content analyses (Clerehan et al., 2005) to identify trends in the written communication of groups or institutions (Sproule, 2006). Both approaches have areas of similarity and overlap and have been used extensively in academic literature to measure the readability of written texts (Howes et al., 2014).

Research on the readability of written texts in accounting (see, for example, Adelberg & Razek, 1984; Courtis, 1986; Lewis et al., 1986; Pound, 1981; and Raabe et al., 1984) tends to use quantitative approaches to the measurement of readability. More specifically, the Gunning Fog Readability Index (GFRI), Flesch Reading Ease Index (FREI), and Flesch – Kincaid Readability Score (FKRS) are the most commonly used methods in the accounting literature (Lahmar & Piras, 2023; Luo et al., 2018; Phillips et al., 2007). The GFRI is computed based on the number of words in a sentence plus the percentage of words with three or more syllables (also known as 'complex words') in a sentence, multiplied

by 0.4 (Gunning, 1952; see also Loughran & McDonald, 2014). The resulting score provides an estimate of the number of years of education required to comprehend the text at first reading (Lahmar & Piras, 2023). The GFRI assigns a score to written text, with higher (lower) scores denoting less (more) readable text (Loughran & McDonald, 2014). Specifically, texts with a GFRI score of 5 are classified as ‘very easy to read’, 6 as ‘easy to read’, and 7 as ‘fairly easy to read’. A score of 9 indicates that the document uses ‘standard English’, 12 that it is ‘fairly difficult to read’, and scores from 14 to 16 indicate that the document is ‘difficult to read’. Finally, documents with scores greater than 16 are deemed ‘very difficult to read’ (Chapman et al., 2000).

The FREI is broadly similar to the GFRI, and its formula considers the average length of a sentence and the average number of syllables per word to determine the readability of a written text (Flesch, 1948). More specifically, it is computed as ‘206.835 minus (1.015 times number of words per sentence) minus (84.6 times number of syllables per word)’ (Cassell et al., 2019). Theoretically, FREI scores range from 0 to 100, with higher (lower) scores indicating lower (higher) levels of reading difficulty. Specifically, the FREI formula assigns the following descriptions to ranges of scores: 0–30, ‘very difficult to read’; 30–50, ‘difficult to read’; 50–60, ‘fairly difficult to read’; 60–70 indicates a document uses ‘plain English’; 70–80, ‘fairly easy to read’; 80–90, ‘easy to read’; and finally, 90–100 ‘very easy to read’ (Farr et al., 1951; Flesch, 1948).

The GFRI provides a straightforward metric for plain English attributes, as recommended by linguistic experts. Nonetheless, accounting exam texts are likely to have an extremely high proportion of words with three or more syllables, which the GFRI classifies as ‘complex words’. Despite this, these words (e.g. ‘depreciation’, ‘liability’) are well understood by most candidates (Lahmar & Piras, 2023; Loughran & McDonald, 2014; see also Kwolek, 1973). The FREI considers just the number of syllables, without classifying words as ‘complex’, which might help it to overcome GFRI’s shortcoming of potentially misclassifying ‘simple words’ with three or more syllables as ‘complex’. It is the most widely used, reliable, and tested metric for measuring the readability of a text (Lee & French, 2011). Because the FKRS was developed using the U.S.A. education system as a benchmark, its relevance to SAICA’s qualifying examinations is doubtful. Thus, we employed the GFRI and FREI in our study. All three indices can be calculated using Microsoft Word and are thus objective and easy to use (Phillips et al., 2007). Websites that calculate the GFRI of written texts are freely available (e.g. <http://gunning-fog-index.com>).

## Results

### Testing hypothesis 1

Table 1 presents the mean GFRI and FREI scores for the ITC and APC exams across the sample period, including the overall mean (median) scores, standard deviations, minimum and maximum values, and t-test statistics. Panel A of Table 1 shows that the mean (median) GFRI score of the ITC exam was 15.820 (16.520), suggesting that a typical ITC exam is ‘very difficult to read’. These results are comparable with GFRI scores of other professional examinations, such as the CMA (i.e. 17.000) and CPA (i.e. 16.000) examinations reported in Phillips et al. (2007). We further observed that the

**Table 1. Descriptive statistics:** The table presents the GFRI and FREI scores of the Initial Test of Competence (ITC) and Assessment of Professional Competence (APC) exams. It also provides summary descriptive statistics.

Year	Panel A: Initial Test of Competence Exam		Panel B: Assessment of Professional Competence Exam	
	GFRI scores	FREI scores	GFRI scores	FREI scores
2014	11.733	42.395	16.410	50.700
2015	16.838	34.300	14.900	44.100
2016	16.708	41.400	12.750	46.900
2017	14.093	40.675	13.700	44.100
2018	16.515	35.048	12.130	49.200
2019	14.808	37.043	13.770	47.900
2020	17.055	33.960	12.910	49.100
2021	15.940	32.100	13.540	47.300
2022	18.730	41.100	13.440	48.000
No. of Obs.	9.000	9.000	9.000	9.000
Mean	15.820	37.560	13.730	47.480
Median	16.520	37.040	13.540	47.900
Std. Dev.	2.030	3.880	1.270	2.230
Minimum	11.730	32.100	12.130	44.100
Maximum	18.730	42.400	16.410	50.700
Std. Error Mean	0.680	1.290	0.420	0.740
Degree of Freedom	8.000	8.000	8.000	8.000
Plain/Standard English	9.000	60.000	9.000	60.000
<i>t</i> -statistic	10.080	-17.360	11.180	-16.880
<i>p</i> -value	0.000	0.000	0.000	0.000

standard deviation of the GFRI score for the ITC exam was 2.030, with minimum and maximum scores of 11.730 and 18.730, respectively, suggesting variation in the readability of the exam across the sample period. The mean (median) FREI score for the ITC examination was 37.560 (37.040), further indicating that the ITC exam is generally ‘difficult to read’. Table 1 also shows the variability in the FREI score, with a standard deviation of 3.880 and scores ranging from 32.100 to 42.400, highlighting some measure of variation in the readability of the ITC exam over the sample period (see Table 1 (Panel A)).

To assess whether the ITC exam readability scores were due to random variability, we conducted single-sample *t*-tests. The results in Table 1 (Panel A) indicate that the sample mean GFRI score of the ITC exam (i.e. 15.820) was significantly higher ( $t = 10.080$   $p = 0.000$ ) than the upper threshold for ‘standard English’ (i.e. 9.000). Likewise, we observed that the sample mean FREI score of the ITC exam (i.e. 37.560) was significantly lower ( $t = -17.360$ ;  $p = 0.000$ ) than the lower threshold for ‘plain English’ (i.e. 60.000). Both results confirm that the observed average sample scores were not due to random variability.

Table 1 (Panel B) shows that the sample mean (median) GFRI score of the APC exam was 13.730 (13.540), indicating that the APC exam, on average, is ‘difficult to read’. These results are slightly lower than the GFRI scores reported for the CMA (i.e. 17.000) and CPA (i.e. 16.000) examinations in Phillips et al. (2007). Moreover, we observed that the standard deviation of the APC exam’s GFRI score was 1.270, with minimum and maximum scores of 12.130 and 16.410, respectively, suggesting variations in the readability of the APC exam across the sample period. Table 1 (Panel B) shows that the sample mean (median) FREI score of the APC examination was 47.480 (47.490), similarly suggesting that a typical APC exam is ‘difficult to read’. In addition, the results show variability in the FREI

**Table 2. Test of difference between sample mean GFRI and FREI scores of SAICA's qualifying exams:** The table presents results of Mann-Whitney U-test based on the GFRI and FREI scores of the Initial Test of Competence (ITC) and Assessment of Professional Competence (APC) exams.

		Grouping Variable	No of Obs.	Mean Rank	Sum of Ranks
Gunning Fog Reading Index		ITC	9.000	12.440	112.000
		APC	9.000	6.560	59.000
		Total	18.000		
Flesch Reading Ease Score		ITC	9.000	5.000	45.000
		APC	9.000	14.000	126.000
		Total	18.000		
Test Statistics <sup>a</sup>					
		Gunning Fog Reading Index	Flesch Reading Ease Score		
Mann-Whitney U		14.000	.000		
Wilcoxon W		59.000	45.000		
Z		-2.340	-3.578		
Asymp. Sig. (two-tailed)		.019	<0.001		
Exact Sig. [2*(one-tailed Sig.)]		.019 <sup>b</sup>	<0.001 <sup>b</sup>		

a. Grouping Variable: V1.

b. Not corrected for ties.

scores across the sample period, with a standard deviation of 2.230 and minimum and maximum scores of 44.100 and 50.700, respectively (see Table 1 [Panel B]).

To rule out the possibility that the APC exam readability scores were the result of random variability, we carried out single-sample t-tests. These showed that the APC exam's sample mean GFRI score (i.e. 13.730) was significantly higher ( $t = 11.180$   $p = 0.000$ ) than the 'standard English' threshold (i.e. 9.000). Similarly, t-test results showed that the sample mean FREI score of the APC exam (i.e. 47.480) was significantly lower ( $t = -16.880$   $p = 0.000$ ) than the 'plain English' threshold (i.e. 60.000) (see Table 1 (Panel B)).

### Testing hypothesis 2

The descriptive statistics in Section 4.1 indicate that the readability difficulty of a typical ITC exam (median GFRI/FREI score = 16.520/37.040) likely differs from that of a typical APC exam (median GFRI/FREI score = 13.540/47.900). To determine whether these differences were statistically significant, we conducted the Mann-Whitney U-test. The results, shown in Table 2, support a significant difference in the readability difficulty between the two exams.

The Mann-Whitney U-test, presented in Table 2, showed a statistically significant difference between the sample median GFRI score of the ITC exam (16.520) and the APC exam (13.540). Specifically, the mean rank, computed based on GFRI scores, of the ITC exam (12.44) was greater than that of the APC exam (6.56), with the difference being statistically significant ( $U = 14.000$ ;  $Z = -2.340$ ; and *Asymp. Sig. (2-tailed)* = 0.019). Similarly, we examined whether the sample median FREI score of the ITC exam (37.040) differed significantly from that of the APC exam (47.480). Once again, the Mann-Whitney U-test results showed that the mean rank, computed based on the FREI scores of the ITC exam (i.e. 5.00), was significantly lower than that of the APC exam (14.00) ( $U = 0.000$ ;  $Z = -3.578$ ; and *Asymp. Sig. (2-tailed)* < 0.001). These results suggest that the ITC exam, on average, imposes a higher cognitive and reading load than the APC exam.

**Table 3. Descriptive statistics based on the ‘required’ part of SAICA’s qualifying exams:** The table presents the GFRI and FREI scores of the ‘required’ part of SAIVA’s Initial Test of Competence (ITC) and Assessment of Professional Competence (APC) exams. It also provides summary descriptive statistics.

Year:	Panel A: Initial Test of Competence Exam		Panel B: Assessment of Professional Competence Exam	
	GFRI scores	FREI scores	GFRI scores	FREI scores
2014	19.370	40.700	16.900	41.200
2015	17.180	50.830	16.700	42.800
2016	17.460	50.580	13.880	48.800
2017	15.420	50.850	13.150	34.100
2018	17.290	50.850	15.530	31.000
2019	19.710	47.650	15.180	32.100
2020	16.440	52.250	11.210	39.300
2021	14.770	55.400	13.620	41.400
2022	17.260	46.380	16.900	32.500
No. of Obs.	9.000	9.000	9.000	9.000
Mean	17.210	49.500	14.790	38.130
Median	17.260	50.830	15.180	39.300
Std. Dev.	1.610	4.180	1.970	6.040
Minimum	14.770	40.700	11.210	31.000
Maximum	19.710	55.400	16.900	48.800
Std. Error Mean	0.540	1.390	0.660	2.010
Degree of Freedom	8.000	8.000	8.000	8.000
Plain/Standard English	9.000	60.000	9.000	60.000
<i>t</i> -statistic	15.299	-7.545	8.824	-10.853
<i>p</i> -value	0.000	0.000	0.000	0.000

## Checking robustness

### Examining the readability of the ‘required’ part of the exams

In our analyses thus far, we examined the readability of exams using merged documents that combine the ‘scenario’ and ‘required’ parts. In a typical SAICA exam, the ‘scenario’ section is likely to include numerical data such as numbers, tables, and charts, while the ‘required’ section primarily consists of plain texts. This difference in presentation might affect the cognitive load that each section places on an examinee (Davidson, 2005; Tufte, 2006). Considering this difference, it is plausible that exam readability might not be consistent across the ‘scenario’ and ‘required’ sections. This highlights the importance of checking whether our findings remain consistent when we analyse these two distinct parts of the exams separately. Table 3 presents the results based on our analyses of the readability of the ‘required’ sections of both exams.

The results in Table 3 (Panel A) pertain to the readability of the ‘required’ section of the ITC exam. The sample mean (median) GFRI score in this section of the ITC exam was approximately 17.210 (17.260), indicating that it was slightly less readable than the exam as a whole. There was also notable variability in readability over the sample period, with a standard deviation of 1.610, a minimum score of 14.770, and a maximum score of 19.710. Moreover, the average readability difficulty of the ‘required’ section, measured by sample mean GFRI score (17.210), was significantly higher ( $t = 15.299$ ;  $p = 0.000$ ) than the highest possible score for a document written in ‘plain English’ (9.00) (see Table 3 [Panel A]). This suggests that the ‘required’ section of the ITC exam is, on average, ‘very difficult to read’.

In a similar vein, the sample mean (median) FREI score for the ‘required’ section of the ITC exam was about 49.500 (50.830). Contrary to our GFRI-based observations, the FREI-based results suggest that this section might be slightly easier to read than the exam as a whole. However, variability was still present, with a standard deviation of 4.180, a minimum value of 40.700, and a maximum value of 55.400. Additionally, the t-test results showed that the sample mean FREI score for the ‘required’ section of the ITC exam (49.500) was significantly lower ( $t = -7.545$ ;  $p = 0.000$ ) than the minimum score for documents written in ‘plain English’ (60.000) (see Table 3 (Panel A)). Overall, based on the FREI scores, the ‘required’ section of a typical ITC exam appears to be ‘difficult to read’.

We also investigated the readability of the ‘required’ section of the APC exam. As Table 3 (Panel B) shows, the sample mean (median) GFRI score for this part was *circa* 14.790 (15.180), somewhat higher than the mean (median) GFRI score for the entire exam. In addition, the GFRI scores for the ‘required’ section of the APC exam exhibited variability, with a standard deviation of 1.970, a minimum value of 11.210, and a maximum score of 16.900. The t-test results demonstrated that the sample mean GFRI score for the ‘required’ section of the APC exam (14.790) was significantly higher ( $t = 8.824$ ;  $p = 0.000$ ) than the highest score possible for a document written in ‘plain English’ (9.000) (see Table 3 [Panel B]). These GFRI scores imply that the ‘required’ section of the APC exam is generally ‘difficult to read’.

Table 3 (Panel B) also provides summary statistics for the FREI scores. The results indicate that the sample mean (median) FREI score for the ‘required’ section of the APC exam was around 38.130 (39.300), suggesting that this section of the APC exam, on average, was less readable than the exam as a whole. We further observed standard deviations, and minimum and maximum values of 6.040, 31.000, and 48.800, respectively. The t-test results revealed that the sample mean FREI score for the ‘required’ section of the APC exam (38.130) was significantly lower ( $t = -10.853$ ;  $p = 0.000$ ) than the minimum possible FREI score for texts written using ‘plain English’ (60.000). Taken together, these results confirm that the ‘required’ section of a typical APC exam is ‘difficult to read’.

As in the main analyses, we examined whether there are statistically significant differences in readability between the ‘required’ sections of the two exams. Table 4 shows a significant difference between the median GFRI scores of the ‘required’ sections of the ITC (median GFRI/FREI = 17.260/50.830) and the APC exam (median GFRI/FREI = 15.180/39.300). Specifically, mean rank based on GFRI scores for the ‘required’ section of the ITC exam (12.670) was greater than that of the APC exam (6.330), with this difference being significant ( $U = 12.000$ ;  $Z = -2.518$ ; and *Asymp. Sig. (2-tailed)* = 0.012). Similarly, the mean rank based on FREI scores for the ‘required’ section of the ITC exam (13.330) was greater than that of the APC exam (5.670), with this difference being significant ( $U = 6.000$ ;  $Z = -3.048$ ; and *Asymp. Sig. (2-tailed)* = 0.002). These results suggest that the ‘required’ section of the ITC exam is more challenging to read than that of the APC exam. Hypothesis 2, that the ITC exam is more difficult to read than the APC exam, is therefore supported.

### Examining the readability of the ‘scenario’ part of the exams

In line with the argument in Section 4.3.1, we investigated the readability of the ‘scenario’ sections of SAICA’s qualifying exams. The results are shown in Table 5.

**Table 4. Test of difference between sample mean GFRI and FREI scores of the 'required' part of SAICA's qualifying exams:** The table presents results of Mann-Whitney U-test based on the GFRI and FREI scores of the 'required' part of SAICA's ITC and APC exams.

		Grouping Variable	No of Obs.	Mean Rank	Sum of Ranks
Gunning Fog Reading Index		ITC	9.000	12.670	114.000
		APC	9.000	6.330	57.000
		Total	18.000		
Flesch Reading Ease Score		ITC	9.000	13.330	120.000
		APC	9.000	5.670	51.000
		Total	18.000		
Test Statistics <sup>a</sup>					
		Gunning Fog Reading Index	Flesch Reading Ease Score		
Mann-Whitney U		12.000	6.000		
Wilcoxon W		57.000	51.000		
Z		-2.518	-3.048		
Asymp. Sig. (two-tailed)		.012	.002		
Exact Sig. [2*(one-tailed Sig.)]		.011 <sup>b</sup>	.001 <sup>b</sup>		

a. Grouping Variable: V1.

b. Not corrected for ties.

Table 5 (Panel A) presents results indicating that the mean (median) GFRI score for the 'scenario' section of the ITC exam was approximately 14.130 (14.160), slightly lower than the scores for the 'required' section and the entire exam. We observed variations in the GFRI scores for the 'scenario' section over the sample period, with a standard deviation of 1.160, a minimum of 12.300, and maximum of 15.480. The results of one sample t-test showed that the mean GFRI score for the 'scenario' section (14.130) was significantly higher ( $t = 13.27$ ;  $p = 0.000$ ) than the highest GFRI score for a document written

**Table 5. Descriptive statistics based on the 'scenario' part of SAICA's qualifying exams:** The table presents the GFRI and FREI scores of the 'scenario' part of SAICA's Initial Test of Competence (ITC) and Assessment of Professional Competence (APC) exams. It also provides summary descriptive statistics.

Year	Panel A: Initial Test of Competence Exam		Panel B: Assessment of Professional Competence Exam	
	GFRI scores	FREI scores	GFRI scores	FREI scores
2014	15.335	35.900	14.380	52.300
2015	14.933	37.300	15.080	44.000
2016	14.155	41.000	13.200	46.300
2017	15.010	41.050	13.670	45.100
2018	13.783	41.325	18.790	31.000
2019	12.645	40.000	13.690	49.100
2020	15.475	35.325	11.800	49.700
2021	12.298	41.675	14.540	47.600
2022	13.510	44.650	13.620	47.500
No. of Obs.	9.000	9.000	9.000	9.000
Mean	14.130	39.800	14.310	45.840
Median	14.160	41.000	13.690	47.500
Std. Dev.	1.160	3.040	1.920	6.100
Minimum	12.300	35.330	11.800	31.000
Maximum	15.480	44.650	18.790	52.300
Std. Error Mean	0.390	1.010	0.640	2.030
Degree of Freedom	8.000	8.000	8.000	8.000
Plain/Standard English	9.000	60.000	9.000	60.000
<i>t</i> -statistic	13.270	-19.940	8.290	-6.960
<i>p</i> -value	0.000	0.000	0.000	0.000

**Table 6. Test of difference between sample mean GFRI and FREI scores of the ‘scenario’ part of SAICA’s qualifying exams:** The table presents results of Mann-Whitney U-test based on the GFRI and FREI scores of the ‘scenario’ part of SAICA’s ITC and APC exams.

Grouping Variable		No of. Obs.	Mean Rank	Sum of Ranks	
Gunning Fox Reading Index	ITC	9	9.89	89.00	
	APC	9	9.11	82.00	
	Total	18			
Flesch Reading Ease Score	ITC	9	6.11	55.00	
	APC	9	12.89	116.00	
	Total	18			
Test Statistics <sup>a</sup>					
		Gunning Fog Reading Index	Flesch Reading Ease Score		
Mann-Whitney U		37.000	10.000		
Wilcoxon W		82.000	55.000		
Z		−.309	−2.693		
Asymp. Sig. (two-tailed)		.757	.007		
Exact Sig. [2*(one-tailed Sig.)]		.796 <sup>b</sup>	.006 <sup>b</sup>		

a. Grouping Variable: V1.

b. Not corrected for ties.

in ‘plain English’ (9.000). Similarly, the mean (median) FREI score for the ‘scenario’ section of ITC exam was approximately 39.800 (41.000). The FREI scores for the ‘scenario’ section of the ITC exam were heterogeneous, with a standard deviation, and minimum and maximum values of 3.040, 35.330, and 44.650, respectively. The t-test results showed that the mean FREI score for the ‘scenario’ section of the ITC exam (39.800) was significantly lower ( $t = -19.940$ ;  $p = 0.000$ ) than the lowest value of FREI score for a text written in ‘plain English’ (60.000). Overall, the results indicate that the ‘scenario’ section of a typical ITC exam is ‘difficult to read’, consistent with our observations in the main analyses.

We observe in Table 5 (Panel B) that the mean (median) GFRI score for the ‘scenario’ section of the APC exam was approximately 14.310 (13.690). The GFRI scores showed variability, with a standard deviation of 1.920, a minimum of 11.800, and a maximum of 18.790. The t-test results indicated that the mean GFRI score for this section (14.310) was significantly higher ( $t = 8.290$ ;  $p = 0.000$ ) than the highest score for a document written in ‘plain English’ (9.000). Likewise, the mean (median) FREI score for the ‘scenario’ section of the APC exam was approximately 45.840 (47.500), with a standard deviation of 6.100 and scores ranging from 31.000 to 52.300. The t-test results show that the mean FREI score (45.840) was significantly lower ( $t = -6.960$ ;  $p = 0.000$ ) than the minimum score for texts written in ‘plain English’ (60.000). These results confirm our earlier observation that the APC exam is ‘difficult to read’.

The results presented in Table 6 indicate a statistically insignificant difference between the median GFRI scores of the ‘scenario’ section of the ITC and the APC exams. Specifically, the mean rank based on GFRI scores for the ITC exam (9.890) was not significantly different from that of the APC exam (9.110) ( $U = 37.000$ ;  $Z = -0.309$ ; and *Asymp. Sig. (2-tailed)* = 0.757). However, the mean rank based on FREI scores for the ITC exam (5.000) was significantly greater ( $U = 10.000$ ;  $Z = -2.693$ ; and *Asymp. Sig. (2-tailed)* = 0.007) than that for the APC exam (14.000). The results suggest that while the reading difficulty of the ITC exam is higher than that of the APC exam based on FREI scores, there is no

significant difference based on GFREI scores. Thus, the results partially support the hypothesis that the ITC exam is more difficult to read than the APC exam.

## Discussion

Our findings confirm that both the ITC and APC examinations are difficult to read (Hypothesis 1), indicating that the exams' reading complexity may hinder examinees' comprehension. This suggests that SAICA's qualifying exams might be assessing not only the intended constructs, such as professional knowledge and skills, but also candidates' reading ability (Badgett, 2010). Thus, candidates lacking the 'background knowledge assumed' might not be able to adequately comprehend the questions in the exams (Davison & Kantor, 1982, p. 187). This, in turn, raises concerns about the potential presence of construct-irrelevant variance (CIV), as reading difficulty may influence assessment outcomes in ways unrelated to the competencies being tested (Cronbach, 1980; Plake, 1988). CIV is particularly problematic in constructed-response examinations, where candidates must engage with rich contextualised information and articulate well-structured responses (Zhai et al., 2021).

Our findings align with the argument that professional accounting examinations are designed to assess advanced knowledge and skills through complex scenarios, problem-solving tasks, and in-depth analysis (Barac, 2012) and thus are likely to impose a higher cognitive load on examinees (Sackett et al., 2001), which in turn may make them more challenging to read. Constructed-response questions, which are widely used in SAICA's exams, often require candidates to interpret rich contextualized information, formulate structured response, and apply professional judgement (Zhai et al., 2021). While the increased cognitive load might be necessary to distinguish between candidates of varying readiness levels for professional practice (Dawrant & Han, 2021; Petchauer, 2015), it may inadvertently introduce CIV by assessing candidates' reading comprehension skills rather than just their professional competencies. In this sense, the difficulty of SAICA's exams might reflect not only the technical nature of the assessments but also its reading difficulty, which could disadvantage candidate with lower reading comprehension.

The readability challenge is particularly relevant for Black African candidates whose first language is not English. Prior studies highlight the significant role of English being the first language in exam performance (e.g. van Wyk, 2011), showing that reading comprehension disparities are influenced by student first language (Coetzee et al., 2016). Given SAICA's reliance on constructed-response questions, which tend to exhibit higher levels of CIV (Zhai et al., 2021), readability difficulties may disproportionately disadvantage non-native English speakers. This reinforces the importance of initiatives such as the Thuthuka programme, which provide academic and financial support to candidates from historically disadvantaged backgrounds. Targeted interventions to improve the reading comprehension skills of candidates who are not native English speakers could enhance their ability to pass SAICA's qualifying exams, ensuring that language barriers do not unduly hinder access to the profession. Such interventions are particularly important considering that communication skills are identified as one of the professional skills under the International Accounting Education Standards (IAESB, 2019).

Plake (1988) recommends incorporating readability checks into the validation process for high-stakes professional licensure examinations to mitigate CIV. In line with this recommendation, SAICA's validation procedures may help address some readability challenges in its qualifying exams, though their effectiveness requires further investigation. Our findings indicate that the ITC exam's GREI scores closely resemble those of the CPA and CMA exams, suggesting its reading difficulty is consistent with international professional accounting examinations. However, the slightly lower GREI scores for the APC exam indicate that it is somewhat more readable. The validation processes SAICA employs before administering its exams (SAICA, 2020) may have contributed to this alignment in reading complexity with other professional examinations.

Our results support Hypothesis 2, that the ITC exam is more difficult to read than the APC exam. This finding aligns with Sweller's (1988) argument that exams requiring extensive quantitative calculations and problem-solving tasks – like the ITC exam – are likely to impose a higher cognitive load on candidates. Such exams demand step-by-step calculations and the application of logical reasoning, requiring significant mental effort. In contrast, essay-type exams, like those in the APC exam, typically allow for more interpretive and expansive thinking, enabling candidates to articulate and organise their responses more freely. The results further support the argument that the ITC exams' numerical complexity – stemming from the use of numbers, formulas, and mathematical operations – places a greater cognitive load on candidates than the primarily plain-text questions, which occasionally involve simple calculations, typically encountered in the APC exams.

## Conclusion

This study examined the readability of SAICA's ITC and APC examinations using the Gunning Fog Reading Index (GFRI) and Flesch Readability Ease Index (FREI). Consistent with prior research, our findings show that both exams are notably difficult to read (Hypothesis 1), reflecting the high cognitive demands typically associated with professional licensure exams (e.g. Dawrant & Han, 2021; Sackett et al., 2001). These exams use complex scenarios and problem-solving tasks to evaluate advanced knowledge and skills (Barac, 2012), contributing to their inherent difficulty (Petchauer, 2015). Furthermore, aligning with Sweller's (1988) argument that quantitative problem-solving imposes greater cognitive load than essay writing, the ITC exams were found to be more challenging to read than the APC exams.

Importantly, while reading difficulty may partly stem from the technical nature of accounting assessment, our findings raise concerns about the potential construct-irrelevant variance (CIV) in SAICA's qualifying exams (Cronbach, 1980; Plake, 1988). Poor readability may introduce unnecessary cognitive load, particularly for non-native English speakers, as reading difficulty of the exams, rather than the candidates' knowledge and skills, could influence assessment outcomes (Badgett, 2010). While professional accounting exams are often deliberately designed to impose a high cognitive load to assess readiness for practice, the potential introduction of unintended sources of variance through poor readability poses risks pertaining to fairness and validity (Plake, 1988), creating unintended barriers to entry into the profession. Thus, the findings underscore the need for professional bodies like SAICA to foster readability

of exams to ensure equitable assessment of outcomes for all examinees, regardless of language background.

Our study contributes to the existing literature on the readability of texts by investigating the readability of a high-stakes professional licensure examination using a unique setting in which a majority of the examinees are non-native speakers of the language used as the exam medium (English). In addition, other professional accounting examinations where a substantial portion of the examinees are not native English speakers (e.g. the ACCA) may benefit from these findings. The present study also provides additional insights which could potentially serve as inputs to the academic and policy debate on examinees' overall dismal performance in SAICA's qualifying exams. By demonstrating the reading difficulty of SAICA's qualifying exams, this study highlights the potential presence of construct-irrelevant variance in these exams. These findings shed light on a potential antecedent of examinee performance in professional accounting exams. Finally, the study findings provide inputs for the policy debate on the 'closure' of access to the highly exclusive chartered accountancy profession in South Africa.

Our findings highlight the need for SAICA to evaluate the readability of exam questions as a potential factor determining examinee performance. Given the diversity of candidates' linguistic backgrounds and the inherent complexity of high-stakes exams, SAICA may benefit from revisiting the design of exam questions to ensure they are accessible and comprehensible to a broader range of candidates. The study's findings highlight the need for SAICA to initiate conversations around its exam validation process, a crucial component in the process of ensuring exam readability (Plake, 1988). Additionally, preparatory programmes and resources may have to place greater emphasis on developing reading comprehension skills, particularly for candidates who speak English as a second language. This could involve integrating explicit reading comprehension instruction into the curriculum, as recommended by van Rensburg et al. (2014).

Our study has several limitations. First, although our findings are relevant to other professional licensure examinations within and outside the field of accountancy, and in different settings, our data were limited to SAICA's qualifying exams. Future studies should consider difficulties in exam readability not only outside the field of accountancy but also outside South Africa. Second, this study used the GFRI and FREI, which focus on textual characteristics. However, neither proxy takes inputs from the candidates. Future studies could focus either on examinees' attributes or combine the textual features of exams with data on examinees' attributes to provide additional insights into our understanding of the readability of high-stakes professional licensure exams. Third, although the present study explored the readability of SAICA's qualifying exams, it did not ascertain to what extent the low pass rate is attributable to readability challenges. Future studies should explore the relationship between exam readability and examinee performance.

## Notes

1. The South African Institute of Chartered Accountants (SAICA) is the leading professional accounting body in South Africa, with enormous influence on academe, professional licensing, and the accounting profession (Venter & de Villiers, 2013).

2. For instance, SAICA's report on the ITC exam (for the 2018–2023 period) reveal that about 55 percent of the candidate population are Black Africans, whose mother tongue is not necessarily English (SAICA, 2023a).
3. We use the terms 'candidates' and 'examinees' interchangeably.
4. Prior to the introduction of the ITC and APC exams, SAICA's qualifying exams were referred to as Qualifying Exam—Part I (QE 1) and Qualifying Exam—Part II (QE 2).
5. The Thuthuka Project is an initiative by SAICA aimed at increasing the number of black accounting students by providing financial and other support.
6. Each exam may have several questions within the 'required' part of an exam. We included all the questions in the 'required' section in the merged document.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Data availability statement

Data is available upon reasonable request to the authors.

## References

- Adelberg, A. H., & Razek, J. R. (1984). The cloze procedure: A methodology for determining the understandability of accounting textbooks. *Accounting Review*, 59(1), 109–122.
- Amiran, M. R., & Jones, B. F. (1982). Toward a new definition of readability. *Educational Psychologist*, 17(1), 13–30. <https://doi.org/10.1080/00461528209529241>
- Arthur, N., & Everaert, P. (2012). Gender and performance in accounting examinations: Exploring the impact of examination format. *Accounting Education*, 21(5), 471–487. <https://doi.org/10.1080/09639284.2011.650447>
- Badgett, B. A. (2010). *Toward the development of a model to estimate the readability of credentialing-examination materials* [Unpublished doctoral dissertation]. University of Nevada Las Vegas.
- Baker, E. L., & Linn, R. L. (2002). Validity issues for accountability systems. *Critical Issues in Educational Leadership Series*, 47.
- Barac, K. (2012). Learning approaches to the study of auditing followed by prospective South African chartered accountants. *Southern African Business Review*, 16(2), 47–68.
- Barac, K. (2015). Helping disadvantaged students: Findings from the thuthuka programme. *Accounting Education*, 24(2), 75–101. <https://doi.org/10.1080/09639284.2015.1010171>
- Bargate, K. (2012). The readability of managerial accounting and financial management textbooks. *Meditari Accountancy Research*, 20(1), 4–20. <https://doi.org/10.1108/10222521211234192>
- Benjamin, R. G. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review*, 24(1), 63–88. <https://doi.org/10.1007/s10648-011-9181-8>
- Berelson, B. (1952). *Content analysis in communication research*. The Free Press.
- Bonsall IV, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, 63(2-3), 329–357. <https://doi.org/10.1016/j.jacceco.2017.03.002>
- Businesstech. (2021). Chartered accountants are leaving South Africa – here's where they're going, and what they get paid. *Businesstech*. Retrieved December 12, from <https://businesstech.co.za/news/trending/544200/chartered-accountants-are-leaving-south-africa-heres-where-theyre-going-and-what-they-get-paid/>
- Cassell, C. A., Cunningham, L. M., & Liscic, L. L. (2019). The readability of company responses to SEC comment letters and SEC 10-K filing review outcomes. *Review of Accounting Studies*, 24(4), 1252–1276. <https://doi.org/10.1007/s11142-019-09507-x>

- Chapman, W. W., Aronsky, D., Fiszman, M., & Haug, P. J. (2000). *Contribution of a speech recognition system to a computerized pneumonia guideline in the emergency department. Proceedings of the AMIA symposium, 2000*. American Medical Informatics Association.
- Chiang, W.-C., Englebrecht, T. D., Phillips T. J., Jr, & Wang, Y. (2008). Readability of financial accounting principles textbooks. *The Accounting Educators' Journal*, 18. <https://www.ajejournal.com/ojs/index.php/aej/article/view/74>
- Clerehan, R., Buchbinder, R., & Moodie, J. (2005). A linguistic framework for assessing the quality of written patient information: Its use in assessing methotrexate information for rheumatoid arthritis. *Health Education Research*, 20(3), 334–344. <https://doi.org/10.1093/her/cyg123>
- Coetzee, S. A., Janse Van Rensburg, C., & Schmulian, A. (2016). Differences in students' Reading comprehension of international financial reporting standards: A South African case. *Accounting Education*, 25(4), 306–326. <https://doi.org/10.1080/09639284.2016.1191269>
- Cook, D. A., & Hatala, R. (2016). Validation of educational assessments: A primer for simulation and beyond, *Advances in Simulation* 1(1), 31. <https://doi.org/10.1186/s41077-016-0033-y>
- Courtis, J. K. (1986). An investigation into annual report readability and corporate risk-return relationships. *Accounting and Business Research*, 16(64), 285–294. <https://doi.org/10.1080/00014788.1986.9729329>
- Cronbach, L. J. (1980). Validity on parole: How can we go straight. *New Directions for Testing and Measurement*, 5(1), 99–108.
- Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, 26(1), 19–26.
- 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. <https://doi.org/10.1080/09639280410001695241>
- Davison, A., & Kantor, R. N. (1982). On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly*, 17(2), 187–209. <https://doi.org/10.2307/747483>
- Dawrant, A. C., & Han, C. (2021). Testing for professional qualification in conference interpreting. In M. Albl-Mikasa & E. Tiselius (Eds.), *The routledge handbook of conference interpreting (chapter 20)*. Routledge.
- de Jager, E. (2014). Thuthuka students' perceptions of factors influencing success. *Journal of Economic and Financial Sciences*, 7(1), 53–72. <https://doi.org/10.4102/jef.v7i1.130>
- DuBay, W. H. (2004). *The principles of readability*. Impact Information.
- Eggs, S. (2004). *Introduction to systemic functional linguistics*. Continuum.
- Espahbodi, A., Espahbodi, L., Espahbodi, R., Walker, R., & White, G. T. (2023). Determinants of CPA exam performance. *Journal of Accounting Education*, 64, 100859. <https://doi.org/10.1016/j.jaccedu.2023.100859>
- Farr, J. N., Jenkins, J. J., & Paterson, D. G. (1951). Simplification of flesch Reading ease formula. *Journal of Applied Psychology*, 35(5), 333–337. <https://doi.org/10.1037/h0062427>
- Flesch, R. F. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221. <https://doi.org/10.1037/h0057532>
- Flesch, R. F., & Gould, A. J. (1949). *The art of readable writing*. Harper & Row.
- Flory, S. M., Phillips T. J., Jr, & Tassin, M. F. (1992). Measuring readability: A comparison of accounting textbooks. *Journal of Accounting Education*, 10(1), 151–161. [https://doi.org/10.1016/0748-5751\(92\)90022-W](https://doi.org/10.1016/0748-5751(92)90022-W)
- Franklin, M. A., Lepak, G. M., & Myers, J. K. (2017). An examination of age and gender impact on CPA examination pass rates. *ASBBS Proceedings*, 24, 207.
- Funk, S. C., & Dickson, K. L. (2011). Multiple choice and short-answer exam performance in a college classroom. *Teaching of Psychology*, 38(4), 273–277. <https://doi.org/10.1177/0098628311421329>
- Gholami, J., Mosalli, Z., & Nikou, S. B. (2012). Lexical complexity and discourse markers in soft and hard science articles. *World Applied Sciences Journal*, 17(3), 368–374.

- Godev, C. B., Martínez-Gibson, E. A., & Toris, C. C. (2002). Foreign language Reading comprehension test: L1 versus L2 in open-ended questions. *Foreign Language Annals*, 35(2), 202–221. <https://doi.org/10.1111/j.1944-9720.2002.tb03155.x>
- Gunning, R. (1952). *The technique of clear writing*. McGraw Hill.
- Howes, L. M., Kirkbride, K. P., Kelty, S. F., Julian, R., & Kemp, N. (2014). The readability of expert reports for non-scientist report-users: Reports of forensic comparison of glass. *Forensic Science International*, 236, 54–66. <https://doi.org/10.1016/j.forsciint.2013.12.031>
- IAESB. (2019). Handbook of international education pronouncements. International Federation of Accountants. <https://Handbook-of-International-Education-Standards-2019-1.pdf>
- Kincaid, J., Fishburne, R., Rogers, R., & Chissom, B. (1975). *Derivation of new readability formula for navy enlisted personnel*. Navy Research Branch.
- Klare, G. R. (1963). *The measurement of readability*. Iowa State.
- Koh, M. Y., & Koh, H. C. (1999). The determinants of performance in an accountancy degree programme. *Accounting Education*, 8(1), 13–29. <https://doi.org/10.1080/096392899331017>
- Kraus, T. L. (2019). *Doing things differently: Transformation, innovation and student success in CTA and ITC* [Unpublished master's thesis]. University of Cape Town, Faculty of Commerce. <http://hdl.handle.net/11427/36859>
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage.
- Kruger, S. J. (2020). Accountancy students' and lecturers' perceptions of the effect of open-book assessments on writing examinations. *South African Journal of Higher Education*, 34(1), 158–175. <https://doi.org/10.20853/34-1-3376>
- Kwolek, W. F. (1973). A readability survey of technical and popular literature. *Journalism & Mass Communication Quarterly*, 50(2), 255–264.
- Lahmar, O., & Piras, L. (2023). Making sense and transparency in finance literature: Evidence from trends in readability. *Research in International Business and Finance*, 64, 101900. <https://doi.org/10.1016/j.ribaf.2023.101900>
- Lamb, J. H. (2010). Reading grade levels and mathematics assessment: An analysis of Texas mathematics assessment items and their Reading difficulty. *The Mathematics Educator*, 20(1), 22–34.
- Lee, S., & French, N. (2011). The readability of academic papers in the journal of property investment & finance. *Journal of Property Investment & Finance*, 29(6), 693–704. <https://doi.org/10.1108/14635781111171814>
- Lewis, N., Parker, L., Pound, G., & Sutcliffe, P. (1986). Accounting report readability: The use of readability techniques. *Accounting and Business Research*, 16(63), 199–213. <https://doi.org/10.1080/00014788.1986.9729318>
- Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4), 1643–1671. <https://doi.org/10.1111/jofi.12162>
- Lubbe, I., Peta Myers, L., & Van Rooyen, A. (2020). Introduction to special issue: Challenges for academics educating accounting professionals in South Africa. *South African Journal of Accounting Research*, 34(2), 91–95. <https://doi.org/10.1080/10291954.2020.1750544>
- Luo, J., Li, X., & Chen, H. (2018). Annual report readability and corporate agency costs. *China Journal of Accounting Research*, 11(3), 187–212. <https://doi.org/10.1016/j.cjar.2018.04.001>
- Maroun, W. (2017). Accounting for revenue using an accountability and business model framework: The case of the South African institute of chartered accountants' professional examinations. *South African Journal of Accounting Research*, 31(3), 240–254. <https://doi.org/10.1080/10291954.2016.1205248>
- McLaughlin, G. H. (1969). SMOG grading - a new readability formula. *Journal of Reading*, 22(8), 639–646.
- Mofokeng, L., & Mosala, M. (2021). Counting the cost: The dismal pass and admission rates of black chartered accountancy candidates. *Daily Maverick*. Retrieved May 23, from <https://www.dailymaverick.co.za/opinionista/2021-05-03-counting-the-cost-the-dismal-pass-and-admission-rates-of-black-chartered-accountancy-candidates/>

- Mosenthal, P. B., & Kirsch, I. S. (1998). A new measure for assessing document complexity: The PMOSE/IKIRSCH document readability formula. *Journal of Adolescent & Adult Literacy*, 41(8), 638–657.
- Oakland, T., & Lane, H. B. (2004). Language, Reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing*, 4(3), 239–252. [https://doi.org/10.1207/s15327574ijt0403\\_3](https://doi.org/10.1207/s15327574ijt0403_3)
- Ontong, J., De Waal, T., & Wentzel, W. (2020). How accounting students within the thuthuka bursary fund perceive academic support offered at one South African university. *South African Journal of Higher Education*, 34(1), 197–212.
- Petchauer, E. (2015). Passing as white: Race, shame, and success in teacher licensure testing events for black preservice teachers. *Race Ethnicity and Education*, 18(6), 834–857. <https://doi.org/10.1080/13613324.2013.792796>
- Phillips, T. J., Daily, C. M., & Luehlfing, M. S. (2007). A note on the readability of professional materials for management accountants. In J. Y. Lee & M. J. Epstein (Eds.), *Advances in management accounting* (pp. 311–318). Emerald Group Publishing Limited.
- Plake, B. S. (1988). Application of readability indices to multiple-choice items on certification/licensure examinations. *Educational and Psychological Measurement*, 48(2), 543–551. <https://doi.org/10.1177/0013164488482031>
- Plucinski, K. J. (2011). Readability of cost accounting textbooks. *Business Education Innovation Journal*, 3(1), 15.
- Plucinski, K. J., & Hall, L. A. (2012). A comparison of the readability of advanced accounting textbooks. *Business Education Innovation Journal*, 4(1), 45.
- Pound, G. (1981). A note on audit report readability. *Accounting & Finance*, 21(1), 45–55. <https://doi.org/10.1111/j.1467-629X.1981.tb00028.x>
- Raabe, W. A., Stevens, K. C., & Stevens, W. P. (1984). Tax textbook readability: An application of the cloze method. *Journal of the American Taxation Association*, 6(1), 66.
- Ramalepe, P. (2022). Why South Africa is facing a talent shortage in the auditing industry. *News24*. Retrieved June 29, from <https://www.news24.com/news24/tech-and-trends/news/why-south-africa-is-facing-a-talent-shortage-in-the-auditing-industry-20230629#:~:text=Why%20South%20Africa%20is%20facing%20a%20talent%20shortage%20in%20the%20auditing%20industry,-29%20Jun%202023&text=The%20number%20of%20chartered%20accountant%20candidates%20dropped%20by%204%25%20in%202022.&text=Accounting%20and%20audit%20firms%20are,industry%2C%20according%20to%20an%20expert>
- Razek, J. R., Hosch, G. A., & Pearl, D. (1982). Readability of accounting textbooks. *The Journal of Business Education*, 58(1), 23–26.
- Rezaee, A. A., & Norouzi, M. H. (2011). Readability formulas and cohesive markers in Reading comprehension. *Theory & Practice in Language Studies*, 1(8), 1005–1010. <https://doi.org/10.4304/tpls.1.8.1005-1010>
- Rodrigues, L. L., Pinho, C., Bugarim, M. C., Craig, R., & Machado, D. (2018). Factors affecting success in the professional entry exam for accountants in Brazil. *Accounting Education*, 27(1), 48–71. <https://doi.org/10.1080/09639284.2017.1361851>
- Roos, S. (2009). Factors affecting southern African students' success in CIMA examinations. *Meditari: Research Journal of the School of Accounting Sciences*, 17(1), 49–67.
- Ryan, C. (2019, March 22). Startling rise in chartered accountancy failure rates. *MoneyWeb*. <https://www.moneyweb.co.za/in-depth/investigations/startling-rise-in-chartered-accountancy-failure-rates/>
- Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist*, 56(4), 302–318. <https://doi.org/10.1037/0003-066X.56.4.302>
- Sand, K., Eik-Nes, N. L., & Loge, J. H. (2012). Readability of informed consent documents (1987–2007) for clinical trials: A linguistic analysis. *Journal of Empirical Research on Human Research Ethics*, 7(4), 67–78. <https://doi.org/10.1525/jer.2012.7.4.67>

- Shealy, R. T., & Stout, W. F. (2012). An item response theory model for test bias and differential test functioning. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 197–239). Routledge.
- South African Institute of Chartered Accountants (SAICA). (2014). Competency framework: Detailed guidance for the academic programme competencies of a CA(SA) at the point of the Initial Test of Competence (ITC) (assessment of core technical knowledge). <https://saicawebprstorage.blob.core.windows.net/uploads/Competency-Framework-ITC.pdf>
- South African Institute of Chartered Accountants (SAICA). (2019). Transformation initiatives are paying off says accountancy body. [https://saicawebprstorage.blob.core.windows.net/uploads/Examination-docs/Press\\_release\\_ITC\\_Jan\\_2019b.pdf](https://saicawebprstorage.blob.core.windows.net/uploads/Examination-docs/Press_release_ITC_Jan_2019b.pdf): SAICA
- South African Institute of Chartered Accountants (SAICA). (2020). Guidance for professional programmes competencies of a CA(SA) at the point of the Assessment of Professional Competence. <https://saicawebprstorage.blob.core.windows.net/uploads/Competency-Framework-APC.pdf>
- South African Institute of Chartered Accountants (SAICA). (2023a). South African chartered accountants lead in global trustworthiness. <https://www.saica.org.za/news/south-african-chartered-accountants-lead-in-global-trustworthiness>
- South African Institute of Chartered Accountants (SAICA). (2023b). Initial Test of Competence exam results. <https://saicawebprstorage.blob.core.windows.net/uploads/ITC-June-2023-Statistics-Website.pdf> South African Institute of Chartered Accountants
- Sproule, W. (2006). Content analysis. In M. Walter (Ed.), *Social research methods: An Australian perspective* (pp. 113–134).
- Stephaniewix. (2017). South Africa loses the top spot in global rankings for auditing standards. *International Accounting Magazine*. <https://www.internationalaccountingbulletin.com/news/newssouth-africa-loses-the-top-spot-in-global-rankings-for-auditing-standards-5947804/>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. [https://doi.org/10.1207/s15516709cog1202\\_4](https://doi.org/10.1207/s15516709cog1202_4)
- Tekfi, C. (1987). Readability formulas: An overview. *Journal of Documentation*, 43(3), 261–273. <https://doi.org/10.1108/eb026811>
- Traugh, H. M., Powers, O. S., & Adedokun, A. J. (1987). Readability of accounting principles texts. *Journal of Education for Business*, 62(4), 159–162. <https://doi.org/10.1080/08832323.1987.10772798>
- Tufte, E. R. (2006). *The visual display of quantitative information*. Graphics Press.
- van Rensburg, C. J., Coetzee, S. A., & Schmulian, A. (2014). South African financial reporting students' Reading comprehension of the IASB conceptual framework. *Journal of Accounting Education*, 32(4), 1–15. <https://doi.org/10.1016/j.jaccedu.2014.10.003>
- van Wyk, E. (2011). A note: The SAICA part I qualifying examinations: Factors that may influence candidates' success. *South African Journal of Accounting Research*, 25(1), 145–174. <https://doi.org/10.1080/10291954.2011.11435157>
- Venter, E., & de Villiers, C. (2013). The accounting profession's influence on academe: South African evidence. *Accounting, Auditing & Accountability Journal*, 26(8), 1246–1278. <https://doi.org/10.1108/AAAJ-06-2012-01027>
- Visser, A., & Steenkamp, G. (2017). Explicit Reading time in chartered accountancy examinations: Perceptions of students and lecturers. *South African Journal of Accounting Research*, 31(1), 19–34. <https://doi.org/10.1080/10291954.2015.1105552>
- Walkington, C., Clinton, V., & Shivraj, P. (2018). How readability factors are differentially associated with performance for students of different backgrounds when solving mathematics word problems. *American Educational Research Journal*, 55(2), 362–414. <https://doi.org/10.3102/0002831217737028>
- Weil, S., & Wegner, T. (1997). Increasing the number of black chartered accountants in South Africa: An empirical review of educational issues. *Accounting Education*, 6(4), 307–323. <https://doi.org/10.1080/096392897331389>
- Woo, A., Wendt, A., & Liu, W. (2009). Readability of licensure examinations. *Clear Exam Review*, 20(1), 21–23.

- World Competitiveness Center (WCC). (2023). World Competitiveness Ranking: A comprehensive annual report and worldwide reference point on the competitiveness of countries. <https://www.imd.org/centers/wcc/world-competitiveness-center/rankings/world-competitiveness-ranking/2023/>
- Xia, M., Kochmar, E., & Briscoe, T. (2019). Text readability assessment for second language learners. arXiv e-prints, arXiv: 1906.07580. <https://doi.org/10.48550/arXiv.1906.07580>
- Zamanian, M., & Heydari, P. (2012). Readability of texts: State of the art. *Theory & Practice in Language Studies*, 2(1), 43–53. <https://doi.org/10.4304/tpls.2.1.43-53>
- Zhai, X., Haudek, K. C., Wilson, C., & Stuhlsatz, M. (2021). A framework of construct-irrelevant variance for contextualized constructed response assessment. *Frontiers in Education*, 6, 751283. <https://doi.org/10.3389/educ.2021.751283>