

Readability and Understandability of Notes to Financial Statements

Samantha Valentim Telles¹ 
Bruno Meirelles Salotti¹ 

Abstract

Purpose – Our research examined the impact of readability on the understandability of notes to the financial statement and investigated the influence of a glossary on understandability.

Theoretical framework – Previous studies have treated readability and understandability as synonymous, but understandability is a broader concept influenced by contextual complexity, syntax, and individual characteristics.

Design/methodology/approach – We conducted an experiment involving 111 participants in Brazil with a background in accounting. We manipulated components of the Flesch readability metric and assessed their impact on understandability, as measured by the Meaning Identification Test.

Findings – Enhancing the readability of financial statement notes did not improve their understandability. However, participants' overall comprehension abilities and personal characteristics were found to significantly influence understandability. Additionally, the inclusion of a glossary of technical terms had a positive impact on understandability only for users with limited prior knowledge.

Practical & social implications of research – Researchers should consider alternative metrics to assess understandability. Standard setters and regulators should reconsider the effectiveness of presenting a glossary. Investors can advocate for improved text quality measures and the presentation of high quality information.

Originality/value – Most readability studies focus on reports written in English, and they are mainly applied to Anglophone countries. Our study extends this area of research in Brazil to reports written in Portuguese. Additionally, it highlights the need to explore the relationship between readability and understandability, as traditional readability formulas may not fully capture the comprehensibility of a text.

Keywords: Readability, understandability, notes to the financial statements.

1. Universidade de São Paulo, Departamento de Contabilidade e Atuária, São Paulo, SP, Brasil

How to cite:

Telles, S. V., & Salotti, B. M. (2024). Readability and Understandability of Notes to Financial Statements. *Revista Brasileira de Gestão de Negócios*, 26(1), e20230127. <https://doi.org/10.7819/rbgn.v26i01.4251>

Received on:
June/07/2023
Approved on:
Mar/06/2024

Responsible Editor:
Prof. Dr. Ivam Ricardo Peleias

Reviewers:
César Silva; José Dias Filho

Evaluation process:
Double Blind Review

This article is open data



Revista Brasileira de Gestão de Negócios

<https://doi.org/10.7819/rbgn.v26i01.4251>

I Introduction

Accounting narratives are essential for conveying financial information to shareholders through annual reports (Jones, 1997). However, challenges in accounting communication arise due to semantic differences, the incorporation of everyday language with specific meanings, and the use of terms from different languages (Dias & Nakagawa, 2001; Lopes et al., 2009; Evans et al., 2015).

Readability and understandability are often considered interchangeable, but they are not necessarily synonymous (Jones, 1997). Readability refers to the complexity of the text itself and is a passive, text-centered measure that assesses the difficulty of reading a passage. On the other hand, understandability refers to the reader's ability to grasp the intended meaning. It is a broader concept that encompasses the interaction between the text and the reader and reflects the reader's capacity to derive knowledge from the text. Factors that influence understandability include the complexity of the context, educational background, prior experience, syntactic difficulty, and individual characteristics such as background knowledge, interest, and general reading ability (Smith & Taffler, 1992; Jones, 1997; Jones & Smith, 2014).

Therefore, the primary objective of this research was to examine the impact of readability on the understandability of notes to the financial statements.

In 2014, the Brazilian Accounting Standards Committee (CPC - Comitê de Pronunciamentos Contábeis (2014)) issued a guideline, the Brazilian Accounting Standards Guideline number 7 (*Orientação do Comitê de Pronunciamentos Contábeis - OCPC 07*). This guideline outlines the fundamental requirements for the preparation and disclosure of accounting-financial reports, particularly regarding the notes. One of the stipulations is the emphasis on understandability. OCPC 07 suggests that technical language should be used only when absolutely necessary and, in this case, encourages companies to provide a comprehensive and concise glossary of such terms alongside the financial statements. Therefore, assuming that a glossary is needed, the secondary purpose of this research was to investigate how the inclusion of a glossary might affect the understandability of the notes.

Typically, annual reports consist of approximately 80% textual narrative, with the remaining portion consisting of numerical data and other quantitative information (Lo et al., 2017). A study conducted by KPMG (2011) found that the notes are the most significant source of

disclosure complexity. Glaum et al. (2013) discovered a significant association between the quality of the notes and analyst forecast errors, suggesting that the improper presentation of the notes may affect the relevance of the information. In the preface of the Securities and Exchange Commission's (SEC) "A Plain English Handbook," Warren Buffet highlighted the struggle to understand public companies' financial documents and emphasized the lack of technical knowledge among investors (Securities and Exchange Commission, 1998).

EY (Ernst & Young) monitors comments on public company filings from the SEC staff and has observed that many comments call for better disclosure to improve the understandability of financial statement information (Ernst & Young, 2016).

Although readability studies have been conducted since the 1950s, most have focused on reports written in English, primarily examining Anglophone countries such as the United States, the United Kingdom, Australia, and Canada (Moreno & Casasola, 2015; Gosselin et al., 2021). The importance of studying information in Portuguese stems from the fact that it is a widely spoken language worldwide, with over 265 million speakers, and is the most widely spoken language in the Southern Hemisphere (United Nations Educational, Scientific and Cultural Organization, 2022).

Recently, readability studies have emerged in other markets using different readability metrics (Ajina et al., 2016; Gomes et al., 2018; Bacha & Ajina, 2020; Hassan et al., 2019; Santos et al., 2019; Borges & Rech, 2019; Monteiro et al., 2020; Silva et al., 2020; Telles & Salotti, 2020; Soepriyanto et al., 2021; Sun et al., 2022; Cadarin & Theiss, 2022; Mendes & Lucena, 2022; Souza & Borba, 2022; Sena et al., 2023). However, these studies do not examine the proximity of readability to understandability. Evidence of this disconnection can be seen in the fact that a text written in reverse order would yield the same readability score as one written in normal order, despite being significantly more difficult to comprehend (Schriver, 1989). Therefore, focusing solely on readability metrics can have potentially negative impacts on a text (Dreyer, 1984; McNamara et al., 2002; Graesser et al., 2004).

We conducted an experiment involving 111 participants in Brazil, manipulating readability components and assessing their impact on understandability. Our findings suggest that enhancing readability does not necessarily increase the understandability of notes to the financial

statements. Nevertheless, personal comprehension skills and characteristics play an important role. Additionally, the inclusion of a glossary affects understandability, but it might only be beneficial for users with limited prior knowledge of financial statement notes, as it could potentially overload others with excessive information, thus reinforcing the OCPC 07 requirements for a concise glossary when needed.

2 Theoretical framework

Readability formulas are quantitative and objective tools used to assess the difficulty of a text without requiring an estimate of the reader's abilities (Rush, 1985). These formulas, as described by Klare (1974-1975), aim to provide

indices that predict the likely difficulty of a text based on language variables, making them predictive devices.

In the accounting literature, various readability measures have been employed, as documented in Table 1. For a detailed discussion of readability, we recommend the studies conducted by Loughran and McDonald (2016) and Gosselin et al. (2021), as they provide comprehensive literature reviews on this subject.

In each of the previous studies, readability was assessed using different metrics that employed different methods to measure word length and applied different weights to the components of readability.

Previous studies have indicated poor readability in accounting texts from Brazilian firms (Gomes et al.,

Table 1
Previous studies on readability and understandability

Authors	Researched subject	Readability index used	Sample	Conclusions
Soper & Dolphin (1964)	Improved readability of accounting reports.	Flesch	1) 25 firms from 1948 and 1961. 2) 42 people from four different financial literacy groups.	1) A statistically significant decrease in the readability of accounting information. 2) There is a relationship between understandability and readability.
Worthington (1977)	The relationship between the readability index and the educational level of shareholders.	Dale-Chall	96 companies from the 1974 Fortune list.	36% of shareholders would have problems understanding the notes to the financial statements of 88% of companies.
Lewis et al. (1986)	The similarities in the internal structure of five different readability indexes.	Dale-Chall Flesch Fog Kwolek Lix	9 firms from 1977 to 1980.	Using any of the indexes studied, financial reports to employees, which are supposed to be easily read, presented a score that classified them as "difficult to read" and showed no improvement over time.
Smith & Taffler (1992)	The differences between readability and understandability.	Flesch Lix	Experiment with 18 subjects.	For more sophisticated users, such as accountants, Flesch and understandability scores were much closer to understandability than those of a less sophisticated group, such as undergraduate students.
Rennekamp (2012)	Readability and investors' reaction.	No specific readability index used.	Experiment with 234 subjects.	More readable press releases elicited stronger reactions from small investors, indicating that increased readability increased investors' perceptions of reliability.
Lo et al. (2017)	Whether managers use complex MD&A disclosures to hide information from investors.	Fog	26,967 firm-years observations from 2000 to 2012.	Managers use financial disclosure to manipulate investors' understanding of firm value.
Telles & Salotti (2020)	Comparisons of readability and intelligibility metrics for notes in English and in Portuguese.	Flesch index Intelligibility metrics	176 firm-year observations from firms in Brazil from 2012 to 2015.	Readability and intelligibility metrics do not measure the same aspects of a text. Additionally, native language texts were found to be more readable and intelligible than non-native language texts.
Sun et al. (2022)	The impact of IFRS and GRI adoption on the readability and conciseness of corporate reports.	Fog Number of pages	Top 100 Chinese A-share listed companies.	IFRS adoption had a positive impact on readability and conciseness, while GRI adoption had a negative impact on both variables.

2018; Borges & Rech, 2019; Silva et al, 2020; Cadorin & Theiss, 2022; Mendes & Lucena, 2022; Souza & Borba, 2022; Sena et al., 2023). The lack of a relationship between readability and information quality, as found by Sena et al. (2023), or a firm's disclosure strategy that uses narrative complexity to maintain social legitimacy, as found by Hassan et al. (2019), could explain this.

Striving for better readability scores without maintaining coherence and cohesion can have a negative impact on understandability (Graesser et al., 2004). Readability can also be used as an obfuscation tool (Gosselin, Maux, & Smaili, 2021), with companies concealing unfavorable information through complex texts (Li, 2008; Lo et al., 2017; Noh, 2021). This may also be related to the findings of Silva et al (2020), who found that more complete information about economic facts tends to be less readable.

There are conflicting results regarding the relationship between readability and understandability (Razik, 1969; Dreyer, 1984; Telles & Salotti, 2020). Intelligibility metrics may provide a more comprehensive and accurate assessment of understandability than readability measures (Telles & Salotti, 2020), suggesting a possible inverse relationship between readability and understandability. Souza and Borba (2022) also found a low association between commonly used readability metrics and a modified metric designed to include additional aspects of text complexity.

Although the results of previous research are conflicting, our first research hypothesis is as follows:

H1: Improving readability affects the understandability of notes to the financial statements.

According to Rush (1985), science and social studies employ specialized vocabulary that artificially increases the number of difficult words and affects readability scores. In accounting, certain words have different meanings compared to other fields (Dias & Nakagawa, 2001; Lopes et al., 2009). Words such as "income" or "cost" may not be considered complex by readability scores, but they can lead to differences in understanding between the preparer and the user due to their technical nature. On the other hand, as noted by Loughran and McDonald (2014), words such as "company," "operations," and "management" may be considered complex by readability measures, but they are unlikely to confuse users.

In addition to syntactic, semantic, and textual skills, a reader must have a specific ability related to the socio-historical reality reflected in the text (Leffa, 1996). As Dias and Nakagawa (2001) pointed out, effective communication of accounting data requires that both the preparers and the users of such information assign the same meaning to the symbols used. Therefore, a basic understanding of accounting is necessary for analyzing financial statements. The user is expected to have a certain level of familiarity with accounting standards and terminology (Hendriksen & Breda, 1999; Iudícibus, 2009).

To minimize differences in references, it is possible to use clear and simplified terminology and avoid overly technical language in financial statements (Worthington, 1977; Iudícibus, 2009). From this perspective, the problem with accounting communication lies in the divergent references between the preparer and the user regarding the terms and statements presented.

Based on these arguments, and considering the OCPC 07 recommendation to provide a concise glossary when needed, our second hypothesis is as follows:

H2: Increasing users' familiarity with accounting vocabulary through a glossary has a positive impact on the understandability of financial statement notes.

3 Methodology

Most of the studies on this topic have relied on archival research, with only a few exceptions such as Smith and Taffler (1992), Rennekamp (2012), and Besuglov and Crasset (2021), who conducted experiments. In our research, we developed an experimental instrument and administered it to undergraduate and graduate students enrolled in business-related courses. We used Qualtrics®, an internet-based tool, to collect the data and conducted the experiment online. Our study received ethical clearance from an ethics committee and was categorized as "exempt" due to the absence of ethical concerns.

Elliott et al. (2007) found that MBA students can serve as suitable proxies for non-professional investors, justifying our choice of participants. Although our experimental design required participants to have some level of accounting knowledge, it did not require an advanced level of expertise. Therefore, the students were considered to be comparable to potential investors in terms of their suitability for the study.

Prior to the main experiment, we conducted a pre-test involving four Ph.D. students from the Faculty of Economics, Administration, Accounting, and Actuarial Sciences at the University of São Paulo (FEA-USP). We also sought the assistance of a professor experienced in conducting experiments at the University of Texas at El Paso (UTEP) to review and evaluate our research instrument. Their feedback was invaluable in improving certain aspects of the instrument and rectifying any issues identified.

This experiment was conducted from July to October 2017. E-mail invitations were sent to coordinators of different private and public higher education institutions in Brazil, who then, forwarded them to their students.

To assess readability, we focused on two components of the Flesch index (Equation 1): sentence length and number of syllables. This metric, which is one of the oldest formulas for measuring readability, is widely regarded as highly accurate (Gosselin et al., 2021). For our study, we used an adapted version of this metric specifically designed for the Portuguese language, as developed by Martins et al. (1996) and used in previous studies (e.g. Telles & Salotti, 2020; Sena et al., 2023):

$$\begin{aligned} \text{Adapted FLESCH index} = & \\ - 84.6 (\text{Number of syllables/Number of words}) & \quad (1) \\ - 1.015 (\text{Number of words/Number of sentences}) & \\ + 248.835 & \end{aligned}$$

Longer words and sentences imply a lower Flesch index and lower readability.

To measure understandability, we used the Meaning Identification Test (MIT), which employs paraphrases that retain the meaning of the original sentence, as well as paraphrases that change the meaning. Jones and Smith (2014) examined several measures for assessing understandability in accounting texts and concluded that the MIT comprehension test yielded superior results compared to other measures in terms of measuring passage comprehension. Given the theoretical framework of our study, it seems reasonable to consider the MIT technique as an appropriate measure since it aligns both the writer's intentions and the user's comprehension.

We selected a real company, BRF (2015), as the basis for our experiment. BRF is a Brazilian multinational company engaged in the raising, production, and slaughtering of poultry and pork for the processing, production, and sale of fresh meat, processed products, and more. Our decision to choose a real company was motivated by

the desire to provide participants with information that closely resembles real financial information. Additionally, BRF was chosen because its primary business activity is not overly difficult to understand, which helps to avoid biasing the results.

We simplified BRF's financial statements to prevent participant fatigue (Smith & Taffler, 1992) and to minimize potential bias in our findings.

Using an online dictionary of synonyms (Sinônimos, 2016), we replaced words with synonyms that had either more or fewer syllables, depending on the treatment. This was done to create as much contrast as possible between the treatments in terms of readability.

To manipulate sentence length, we converted some closing punctuation marks to commas in the longer sentence treatment and vice versa in the shorter sentence treatment. In some cases, we had to add or remove words from the text while still adhering to the rules of proper writing. However, the number of words added or removed was not significant enough to affect the complexity of the text.

As a result, we had four different treatments for the notes, as outlined in Table 2, Panel A: (i) words with fewer syllables and shorter sentences; (ii) words with fewer syllables and longer sentences; (iii) words with more syllables and shorter sentences; and (iv) words with more syllables and longer sentences. In all four treatments, the meaning and information conveyed by the sentences remained the same.

Table 2, Panel B, provides descriptive statistics for all components of the Flesch index, as well as the overall Flesch index score.

Treatment 1 received the most readable notes (with a Flesch index score of 36.01), indicating that they would likely be more understandable. Conversely, Treatment 4 received the least readable notes (with a Flesch index score of 8.86), suggesting less understandability. Treatments 2 and 3 were employed to examine which component of the Flesch metric is more influential in achieving understandability.

To construct the MIT questions, we made word substitutions with synonyms across all treatments. For the treatments with fewer syllables per word, we used synonyms from the treatments with more words, and vice versa. In cases where we aimed to create a paraphrase with a change in meaning, we modified one or more words in the previously generated paraphrase using antonyms or words with different meanings. These changes were

Table 2
Experimental treatments

		PANEL A		
	Fewer syllables	More syllables		
Fewer words	1	3		
More words	2	4		
		PANEL B		
	Treatments			
	(1)	(2)	(3)	(4)
Flesch index	36.01	27.12	17.35	8.86
Mean syllables per word	2.34	2.32	2.56	2.54
Mean words per sentence	14.24	24.79	14.37	25.02
Number of sentences	213	125	211	124
Number of words	3,033	3,099	3,033	3,102
Number of syllables	7,112.01	7,200.01	7,776.00	7,868.00

implemented to ensure that any differences observed were attributable to readability rather than variations in the instrument itself. If participants were able to identify whether the sentences were paraphrases or not, this indicated higher understandability.

3.1 Experimental protocol

Participants were randomly assigned to one of four treatments without prior knowledge of their specific assignment.

Our experiment consisted of two main rounds, and participants were not given a pre-determined time limit to complete each round. In the first round, participants received a simplified set of financial reports (including the balance sheet, income statement, statement of changes in shareholders' equity, cash flow statement, and value-added statement) with the notes to the financial statements presented with one of the treatments described above. In this round, sentences with word replacements were presented, and participants were required to indicate whether these replacements changed the meaning of the original sentence in the notes (Supplementary Data 1. Survey Questions and Possible Answers for Treatments). This phase of the experiment corresponds to a between-subjects analysis, as participants were exposed to different treatments.

In the second round, participants received the same set of reports and notes, following the same treatment assigned in the first round, along with a glossary containing all the technical words used in the reports. Therefore, they were exposed to a single treatment in both rounds, making this a within-subjects analysis. They were asked to answer the same MIT questions as in the previous round. Participants had access to their previous answers

and were given the option to either keep their answers or change them using the glossary.

We created the glossary based on the one provided by the CPC - Comitê de Pronunciamentos Contábeis (2009) (for small and medium enterprises (SME or *Pequenas e Médias Empresas* – PME) and it was useful to test for information redundancy (Supplementary Data 2. Glossary of Accounting Terms Used (in Portuguese)).

In addition, we asked participants if they made any changes to their answers and the reason behind the changes (Supplementary Data 3. Phase 2 Extra Questions). This was done to ensure that modifications were due to participants' understanding and use of the glossary rather than missing information in the first round. We tested whether the reason affected the results when it was not the glossary, but this was not statistically significant.

Jones (1997) suggests that different readers may exhibit different levels of understanding due to their individual reading abilities. The theoretical framework employed by Marchant et al. (1988) suggests that the comprehension process involves an interaction between the linguistic message, the context, and the reader's prior knowledge. To account for these factors, we took measures to control for them.

After participants completed both rounds, we administered a new MIT test that included questions on general subjects (Supplementary Data 4. Phase 3 Pairs of Sentences). This allowed us to determine whether any difficulties participants encountered in understanding the accounting text were specific to this domain or stemmed from personal difficulties. In this round, participants were presented with two pairs of sentences and were asked to determine whether or not they had the same meaning.

To prevent participants from focusing solely on the paraphrased sentences, we presented larger texts with additional context. Additionally, we used a gift card as an incentive to encourage participants to pay close attention to the task and to mitigate any potential issues related to attention or engagement.

4 Results

4.1 Between-subjects experiment: ANCOVA

As presented in Table 3, the study included a total of 112 graduate and undergraduate students as participants, although the distribution of participants across the treatments was not balanced. Treatment 1 had 28 responses, Treatment 2 had 25, Treatment 3 had 29, and Treatment 4 had 30 (Supplementary Data 5. Database).

Participants were required to answer a total of 19 questions, but none of the treatments resulted in any participant answering all questions correctly. On average, participants answered nearly 70% of the questions correctly across all treatments, which is similar to the findings of Jones and Smith (2014), who reported a correct response rate of 73.9%. If participants had answered all questions at random, the average score would be 50% (Jones & Smith, 2014), equivalent to 9.5 points. Only seven respondents scored below the random expectation, four from Treatment 1 and one from each of the other treatments.

Based on a Kruskal-Wallis test for non-parametric data, we were unable to reject the hypothesis that the treatments had indistinguishable means (sigma of 0.193) at the 5% significance level. This suggests that there were no significant differences in understandability between the treatments. To further investigate these results and identify any potential factors that may have influenced the outcomes, we conducted an analysis of covariance (ANCOVA), following the approach employed by Besuglov and Crasselt (2021).

In our study, we used the following covariates:

“Score - Phase 3”: To control for overall reading ability.

The coherence of a text depends on the reader’s skills and knowledge (Graesser et al., 2004). By including this covariate, we aimed to account for individual differences in reading ability.

“Time studying accounting” and “Accounting major”:

To investigate the influence of participants’ background on understandability. “Accounting major” indicates whether the participant’s major field of study was accounting, suggesting that they had at least some knowledge of the subject. “Time studying accounting” represents the amount of time participants had studied accounting, which could reflect a stronger background in the subject for those with more experience.

“Undergraduate” and “Public school”: To control for factors such as educational level and type of school. Smith and Taffler (1992) found that education and experience are significant factors in explaining understandability.

Overall, the distribution of participants across the covariates was similar across treatments. Participants were recruited from different regions of Brazil (except the North) and the Federal District. On average, participants were male, majored in accounting, attended a public university, and were enrolled in an undergraduate course.

We tested for the ANCOVA assumptions, including normality, equal sample size, absence of multicollinearity, homogeneity of variances, linearity, and homogeneity of regression slopes. The model passed these assumptions using a squared version of the dependent variable in our analysis.

According to Tabachnick and Fidell (2013), univariate outliers in the dependent variables and covariates are allowed in ANCOVA analysis. However, multivariate outliers among the dependent variables and covariates can introduce heterogeneity in the regression and may lead to either rejection of the ANCOVA or inappropriate adjustment of the dependent variable. To assess multivariate outliers, we used the Mahalanobis D^2 measure as suggested by Hair et al. (2010). Using SPSS® to calculate the D^2 , we

Table 3
Descriptive Statistics: Dependent Variable (Understandability in Phase 1)

	All treatments	Treatment 1	Treatment 2	Treatment 3	Treatment 4
N	112	28	25	29	30
Mean	13.77	13.04	13.57	14.48	13.87
Median	14	14	13	15	14

identified one observation as a multiple outlier belonging to Treatment 3. As a result, we excluded this observation from the analysis, leaving Treatment 3 with 28 observations. Therefore, the total number of participants is now 111. This exclusion did not affect the previously established conclusions about normality.

In Model 1, we included the four covariates that passed all the tests without incorporating any interactions between them. Model 2 introduced the covariate “time studying accounting.” Model 3 excluded the variable “Score - Phase 3.” Model 4 incorporated interactions between dichotomous variables to explore the combined characteristics of the participants and their impact on the scores obtained in Phase 1. This resulted in four new covariates, bringing the total number of covariates to eight, which adheres to the limit suggested by Hair et al. (2010).

The high values of the type III sum of squares can be attributed to the fact that the dependent variable is already squared. The R^2 value indicates the proportion of the observed data that can be explained by the model. The adjusted R^2 , on the other hand, is used to compare the results of each model as it penalizes the inclusion of new variables.

In Model 2, when the variable “time studying accounting” was included, the adjusted R^2 decreased compared to Model 1. However, the interpretation of the results remained unchanged. It is worth noting that we also ran the test without the variable “accounting major,”

since it could be redundant, potentially explaining the lack of significance in both variables. However, all the results remained consistent, with the main difference being an adjusted R^2 of 0.131, only slightly higher than that of Model 2.

The results of Model 3 showed that removing the variable “Score - Phase 3” not only did not yield different results, but also significantly decreased the adjusted R^2 , indicating that this model is not as effective as the previous ones.

In Model 4, the adjusted R^2 increased to 0.204, suggesting a stronger model even with the inclusion of four new variables. Once again, the variable “readability” was not significant, indicating no difference between treatments. Additionally, “undergraduate” and most of the interactions were significant at the 5% level. Hence, differences in educational level, courses, and type of school appear to have a more significant impact in explaining understandability than readability levels for accounting information in Portuguese.

In all models where it was present, the variable that consistently showed significance in explaining understandability (Squared Score – Phase 1) was the proxy for previous ability in understanding (Score - Phase 3), at a 1% significance level.

Table 4 shows that the variable representing the treatment for readability was not significant in any

Table 4
ANCOVA test

Source	Type III Sum of Squares			
	Model 1	Model 2	Model 3	Model 4
Corrected model	85305.684 (a)***	85853.615 (a)***	37191.144 (a)	130638.672 (a)***
Intercept	43547.91***	44087.598***	476689.84***	47202.374***
Readability	6032.398	5527.285	17463.485	2446.188
Score – Phase 3	48114.54***	48519.801***		42504.035***
Accounting major	1108.455	261.898	0.531	7316.43
Undergraduate	3834.286	4367.269	11048.699*	21424.554***
Public school	6586.746	6936.382	9199.924	8766.122*
Time studying accounting		547.931		
Accounting major * Undergraduate				4939.925
Accounting major * Public school				13502.096**
Undergraduate in process * Public school				36985.801***
Account. major * Undergraduate * Public school				17513.401**
Error	374887.09	374339.160	423001.63	329554.1
Total	4682918	4682918.000	4682918	4682918
Corrected total	460192.78	460192.775	460192.78	460192.78
R^2	0.185	0.187	0.081	0.284
Adjusted R^2	0.130	0.123	0.028	0.204

* p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01.

of the models studied. This suggests that there is no significant difference between the treatments, indicating that different levels of readability do not have an impact on the understandability of accounting information presented in Portuguese.

Complementary tests were also performed on the models. The lack of fit test showed that the covariance models were well fitted, with non-significant sigma values. Additionally, the Levene test indicated that the error variances were homogeneous across treatments. Both tests confirmed that all the models were good fits. The paired comparison between treatments revealed no significant differences between any pair of treatments, further supporting the previous tests. However, the paired comparison test for Model 3 indicated a potential difference between treatments 1 (fewer syllables per word and fewer words per sentence) and 3 (more syllables per word and fewer words per sentence) with a sigma of 0.058. This suggests that there may be a marginal difference between these treatments. Nonetheless, Model 3 is not as good as the others based on its lower R^2 and F statistic of the corrected model, which indicates that all the variables together are equal to zero.

Additional tests were conducted to assess the robustness of the results. For example, including interactions between “Score - Phase 3” and all three dummy variables from Model 1 did not change the results. The significance of the “Score - Phase 3” variable became marginally significant at the 5% level (sigma of 0.058). None of the interactions were significant and the adjusted R^2 decreased to 0.111.

We tested the models with only one of the dummy variables and “Score - Phase 3” as covariates, but the conclusions did not change. We also tested versions of Model 1 without one of the dummy variables, but again, all tests showed similar results. Finally, we tested whether removing one of the dummy covariates and including an interaction between the remaining two could change the results. We found that when “accounting major” is excluded and an interaction between “undergraduate” and “public school” is included, the adjusted R^2 becomes 0.154 and “undergraduate” becomes significant at the 5% significance level (sigma of 0.042). However, this conclusion is not significantly different from the one obtained in Model 4.

Additionally, we tested a new variable provided by Qualtrics® that indicates the time taken by participants to answer the survey. The average time to complete the

task was 53 minutes. However, this variable did not show any significance or change in the results. We also included the variable “age” in several models, but it did not alter the results, even when we included interactions with other variables such as “undergraduate” and “accounting major.” Furthermore, gender was not significant in any of the tests conducted.

Although we tested many other models, the conclusion remains the same. Overall, there was no significant difference between the treatments. Therefore, while the cohesion of the text may have changed, since it is a text-related characteristic, coherence, which is related to the reader’s mental representation, was not affected. As Dreyer (1984) suggested, longer words or sentences may be correlated with reading difficulty, but they may not be the cause. In some cases, longer words or sentences are easier to understand.

The results of our research are consistent with the findings of Telles and Salotti (2020) and Souza and Borba (2022). The former found no evidence that readability and intelligibility (their proxy for understandability) metrics measure the same aspects of a text, while the latter found a low association between the original Flesch index and a modified version that included logical operators, tokens, and other text attributes that can be argued to be closer to understandability than the original Flesch.

However, this contradicts the findings of Soper and Dolphin (1964). The discrepancy in our results could be due to the language of the tests, as theirs was conducted in English, or to the cultural differences between the firms studied, as noted by Noh (2021), which can affect readability. Additionally, it could be influenced by the instrument used, which involved two different firms with different levels of readability and required participants to answer true or false questions. Further research using different instruments is needed to draw more definitive conclusions.

Furthermore, the variable “Score - Phase 3” was found to be significant in explaining the dependent variable. This finding is consistent with Jones (1997), suggesting that prior text comprehension skills are more influential than readability in explaining the understandability of accounting information. Additionally, other reader characteristics such as educational level, courses, and type of school attended can also play a significant role in explaining understandability, as expected (Smith & Taffler, 1992; Jones, 1997).

Therefore, we reject H1: Improving readability does not affect understandability.

4.2 Within-subjects experiment: means test and mediation analysis

In this analysis, we examined the participants' scores before and after receiving the glossary, using the Wilcoxon signed-rank test. The phase 1 and 2 scores differed significantly at the 5% significance level (two-tailed sigma of 0.03).

When we analyzed each treatment separately, the scores were not significant at the 5% significance level (Treatment 1: p-value of 0.0521; Treatment 2: 0.6987; Treatment 3: 0.1496; and Treatment 4: 0.1154).

To determine if the glossary was the reason for the score changes, we performed a mediation analysis following Hayes' (2013) approach. In phase 2, we presented the participants with their answers from phase 1 and allowed them to choose whether to keep the same answer or change it. This means that the scores from phase 1 strongly influenced the scores in phase 2. We included the question "How much did you use the glossary?" to

ensure that the glossary had an impact on the dependent variable. This mediator variable was presented as a slider question where participants had to indicate their level of usage of the glossary on a scale from 0 to 100. Thus, the predictor is the score from phase 1, the criterion is the score from phase 2, and the level of glossary usage reported by the participants serves as the mediator, positioned between the other two variables.

Table 5 presents the descriptive statistics of the dependent variable (Score - Phase 2) and the mediator variable.

No participant answered all questions correctly in phase 2. The statistics for "Score - Phase 2" show little deviation from the data and a distribution close to normal. The median and minimum of the mediator have the same value, indicating that most participants did not use the glossary.

The results of the mediation analysis are shown in Figure 1.

There are certain conditions that need to be met to establish mediation (Baron & Kenny, 1986):

Table 5
Descriptive statistics of mediation analysis: Dependent variable of phase 2 and mediator

Variable	Mean	Median	S.D.	Minimum	Maximum
Score - Phase 2	13.98	14	2.405	8	18
Percentage use of glossary	10.13	0	18.44	0	91

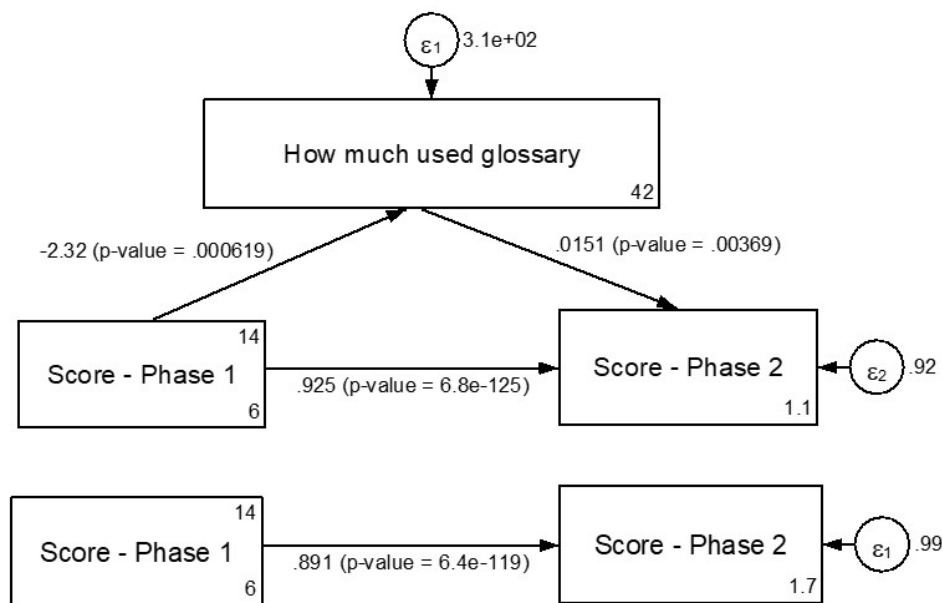


Figure 1. Mediation results: Score of phases 1 and 2 with and without the mediation of "percentage use of glossary"

1. The mediator is affected by the independent variable with a p-value of 0.001 (adjusted R^2 of 0.087 and F-statistic of 11.506). The negative beta (-2.316) indicates a negative relationship, as expected. This means that participants who understood the notes well in the first phase used the glossary less in the second phase.

2. The dependent variable is affected by the independent variable with a p-value of 0 and a high correlation between the scores in both phases (adjusted R^2 of 0.827 and F-statistic of 527.875). Again, the beta (0.891) is consistent with the expected positive relationship. This indicates that participants had similar scores from one phase to the other.

3. The dependent variable is affected by the mediator in the presence of the independent variable and a high correlation between the scores in both phases (adjusted R^2 of 0.838 and F-statistic of 285.478). A p-value of 0.004 confirms this condition, and the sign of the beta (0.015) follows the expected relationship.

4. The coefficient that relates only the independent variable to the dependent variable should have a greater absolute value than the coefficient of the same relationship in the model with the mediator. However, since the beta of the regression with the mediator (0.925) is greater than 0.891, this condition was not met.

The total effect of 0.891 consists of the positive direct effect of 0.925 and the negative indirect effect of 0.034 (-2.316 multiplied by 0.015). We calculated the standard error of the indirect effect using Sobel's (1982) method, which resulted in -2.1348. This value is less than the critical value (± 1.96), indicating a significant indirect effect at the 5% significance level. These results are consistent with the Wilcoxon test and suggest that the glossary acts as a partial mediator influencing the outcome.

Based on these findings, it can be inferred that the presence of the glossary has an impact on the understandability of the accounting information, as expected according to OCPC 07. For individuals with limited or partial knowledge

of accounting terminology, the glossary can be useful in bridging the gap in their understanding. This aligns with Leffa's (1996) suggestions that a larger and more diverse audience requires a more redundant message, whereas a smaller, more specialized, and homogeneous audience requires a less redundant message.

However, when the user of accounting information already has prior knowledge, the glossary becomes redundant and provides more information than necessary. Thus, it is not used at all. These findings support the suggestion of OCPC 07 that the glossary should contain only industry-specific or company-specific terms.

In conclusion, we cannot reject H2: Increasing users' familiarity of accounting vocabulary through a glossary may indeed enhance the understandability of the notes.

4.3 Other analysis

The scores varied across the different treatments but showed significant similarities. Questions 5, 10, and 19 (related to basis of preparation, intangible assets, and new standards, respectively) had the lowest scores, while questions 6, 12, 15, and 16 (related to basis of preparation, intangible assets, and equity, respectively) had the highest scores. This suggests that the type of note does not significantly affect understanding.

The correlation between the Flesch index of each question and the scores from both phases was 0.2342 and remained consistent when separated by phase or treatment. These results support previous findings of a limited association between readability and understandability.

The sample was divided into groups based on glossary usage ("How much used glossary > 0") and tested for statistical differences between the phases within each subgroup. Surprisingly, those who did not use the glossary had significantly different scores in phases 1 and 2, while those who used it had indistinguishable scores. These unexpected results may be attributed to the reduced sample size, as the sample was almost halved. Additionally, the heterogeneous levels of glossary usage may have contributed to the contradictory findings.

A Wilcoxon test indicated that, except for Treatment 1, the scores in phase 1 and phase 2 were not significantly different.

Mediation tests were conducted for each treatment. Treatment 1 showed a significant indirect effect of the glossary. For the other treatments, the indirect effect of the glossary was not significant, which is consistent

with previous findings. However, the small sample size in each subgroup may have influenced these results and contributed to the divergence from the findings obtained from the full sample.

Excluding observations with zero glossary usage and including participants who used the glossary yielded unchanged evidence in the mediation test. Including “accounting major” as a covariate did not alter the results either. Employing bootstrapping techniques also yielded consistent results.

5 Concluding remarks

The main purpose of this study was to examine the impact of readability on the understandability of financial statement notes. We conducted an experiment involving a sample of 111 graduate and undergraduate students from business-related courses in Brazil. By manipulating components of readability, we investigated how participants’ comprehension was affected. We also controlled for relevant participant characteristics to ensure that the observed effects were due to the manipulations.

Our findings indicate that general text comprehension skills and individual characteristics play a more important role than readability in explaining the understandability of accounting information in Portuguese. These results suggest that readability measures alone are insufficient for evaluating the understandability of financial statement notes.

Furthermore, our research revealed that improving familiarity with technical terms through a glossary influenced understandability. However, the evidence also suggests that the glossary may be used primarily by individuals with limited prior knowledge of financial statement notes. For those who are already familiar with the subject, the glossary may simply contribute to information overload.

It is important to note that our study has certain limitations. The findings may not be generalizable to other languages and even outside of Brazil due to the specific vocabulary of each country. Additionally, the results may not apply to other financial information.

The implications of our findings are relevant for various stakeholders. For firms, the results show that glossaries that provide users with already known information are not useful, emphasizing the need for glossaries that contain sector- and company-specific terms. Researchers could explore alternative metrics to assess understandability. Standard setters and regulators could provide clearer guidelines on glossary development,

especially considering that, depending on the information, it can lead to information overload with limited user utility. Finally, investors can advocate for more easily understandable information to facilitate informed decision-making.

References

- Ajina, A., Laouiti, M., & Msolli, B. (2016). Guiding through the Fog: Does annual report readability reveal earnings management? *Research in International Business and Finance*, 38, 509-516. <http://dx.doi.org/10.1016/j.ribaf.2016.07.021>.
- Bacha, S., & Ajina, A. (2020). CSR performance and annual report readability: Evidence from France. *Corporate Governance (Bradford)*, 20(2), 201-215. <http://dx.doi.org/10.1108/CG-02-2019-0060>.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182. <http://dx.doi.org/10.1037/0022-3514.51.6.1173>. PMID:3806354.
- Besuglov, E., & Crasselt, N. (2021). The effect of readability and language choice in management accounting reports on risk-taking: An experimental study. *Journal of Business Economics*, 91(1), 5-33. <http://dx.doi.org/10.1007/s11573-020-00980-4>.
- Borges, G., & Rech, I. (2019). Determinantes da legibilidade das notas explicativas de companhias brasileiras. *Revista de Gestão, Finanças e Contabilidade*, 9(3), 31-51. <http://dx.doi.org/10.18028/rgfc.v9i3.7522>.
- BRF. (2015). *Relatórios Financeiros: Demonstrações Financeiras Padronizadas*. https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/empresas-listadas.htm?codigo=16292
- Cadorin, J., & Theiss, V. (2022). Gerenciamento de impressão das narrativas contábeis da Vale S.A. em detrimento do rompimento da barragem de Brumadinho. *Revista Mineira de Contabilidade*, 23(2), 32-42. <http://dx.doi.org/10.51320/rmc.v23i2.1344>.

- Comitê de Pronunciamentos Contábeis. (2009, 04 de dezembro). *Pronunciamento Técnico PME - Contabilidade para Pequenas e Médias Empresas*. <http://www.cpc.org.br>
- Comitê de Pronunciamentos Contábeis. (2014, 26 setembro). *Orientação Técnica OCPC 07 - Evidenciação na Divulgação dos Relatórios Contábil-Financeiros de Propósito Geral*. <http://www.cpc.org.br/CPC/Documentos-Emitidos/Orientacoes/Orientacao?Id=95>
- Dias Fo., J. M., & Nakagawa, M. (2001). Análise do processo da comunicação contábil: Uma contribuição para a solução de problemas Semânticos, Utilizando Conceitos da Teoria da Comunicação. *Revista Contabilidade & Finanças*, 15(26), 42-57. <http://dx.doi.org/10.1590/S1519-70772001000200003>.
- Dreyer, L. G. (1984). Readability and responsibility. *Journal of Reading*, 27(4), 334-338. <http://www.jstor.org/stable/40029347>
- Elliott, W., Hodge, F., Kennedy, J., & Pronk, M. (2007). Are M.B.A. Students a Good Proxy for Nonprofessional Investors? *The Accounting Review*, 82(1), 139-168. <http://dx.doi.org/10.2308/accr.2007.82.1.139>.
- Ernst & Young. (2016). *SEC comments and trends: An analysis of current reporting issues*. US Professional Practice Group.
- Evans, L., Baskerville, R., & Nara, K. (2015). Colliding Worlds: Issues relating to language translation in accounting and some lesson from other disciplines. *Abacus*, 51(1), 1-36. <http://dx.doi.org/10.1111/abac.12040>.
- Glaum, M., Baetge, J., Grothe, A., & Oberdörster, T. (2013). Introduction of International Accounting Standards, disclosure quality and accuracy of analysts' earnings forecasts. *European Accounting Review*, 22(1), 79-116. <http://dx.doi.org/10.1080/09638180.2011.558301>.
- Gomes, M., Ferreira, R., & Martins, V. (2018). O impacto da OCPC 07 sobre o tamanho e a legibilidade das notas explicativas de companhias brasileiras. *Revista Universo Contábil*, 14(2), 162-184. <http://dx.doi.org/10.4270/ruc.2018216>.
- Gosselin, A.-M., Maux, J., & Smaili, N. (2021). Readability of accounting disclosures: A comprehensive review and research agenda. *Perspectives Comptable*, 20(4), 543-581. <http://dx.doi.org/10.1111/1911-3838.12275>.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202. <http://dx.doi.org/10.3758/BF03195564>. PMID:15354684.
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis*. Prentice Hall.
- Hassan, M., Abbas, B., & Garas, S. (2019). Readability, governance and performance: A test of the obfuscation hypothesis in Qatari listed firms. *Corporate Governance (Bradford)*, 19(2), 270-298. <http://dx.doi.org/10.1108/CG-05-2018-0182>.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis*. The Guilford Press.
- Hendriksen, E. S., & Breda, M. E. (1999). *Teoria da contabilidade* (5ª ed.). Atlas.
- Iudícibus, S. (2009). *Teoria da contabilidade* (9ª ed.). Atlas.
- Jones, M. J. (1997). Methodological themes: Critical appraisal of the cloze procedure's use in the accounting domain. *Accounting, Auditing & Accountability Journal*, 10(1), 105-128. <http://dx.doi.org/10.1108/09513579710158739>.
- Jones, M., & Smith, M. (2014). Traditional and alternative methods of measuring the understandability of accounting narratives. *Accounting, Auditing & Accountability Journal*, 27(1), 183-208. <http://dx.doi.org/10.1108/AAAJ-04-2013-1314>.
- Klare, G. R. (1974-1975). Assessing readability. *Reading Research Quarterly*, 10(1), 62-102. <http://dx.doi.org/10.2307/747086>.
- KPMG. (2011). *Disclosure overload and complexity: Hidden in plain sight*. Ferf - Financial executives research foundation.
- Leffa, V. J. (1996). *Aspectos da leitura: Uma perspectiva psicolinguística*. Sagra DC Luzatto.
- Lewis, N. R., Parker, L. D., Pound, G. D., & Sutcliffe, P. (1986). Accounting Report Readability: The Use of

- Readability Techniques. *Accounting and Business Research*, 16(63), 199-213. <http://dx.doi.org/10.1080/00014788.1986.9729318>.
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2), 221-247. <http://dx.doi.org/10.1016/j.jacceco.2008.02.003>.
- Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting and Economics*, 63(1), 1-25. <http://dx.doi.org/10.1016/j.jacceco.2016.09.002>.
- Lopes, J. E., Ribeiro, R., & Cavalcante, P. (2009, Julho/Dezembro). A Informação Contábil à Luz da Teoria da Comunicação. *Revista Ambiente Contábil*, 1(2), pp. 22-37. <https://periodicos.ufrn.br/ambiente/article/view/1335>
- Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4), 1643-1671. <http://dx.doi.org/10.1111/jofi.12162>.
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187-1230. <http://dx.doi.org/10.1111/1475-679X.12123>.
- Marchant 3rd., H. G., Royer, J. M., & Greene, B. A. (1988). Superior reliability and validity for a new form of the sentence verification technique for measuring comprehension. *Educational and Psychological Measurement*, 48(3), 827-834. <http://dx.doi.org/10.1177/0013164488483032>.
- Martins, T. B., Ghiraldelo, C. M., Nunes, M. G., & Oliveira Jr., O. N. (1996). *Readability Formulas Applied to Textbooks in Brazilian Portuguese* (Série Computação, 28). ICMSC-USP. Notas do ICMSC-USP.
- McNamara, D. S., Louwerse, M. M., & Graesser, A. C. (2002). *Coh-Metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension*. Project funded by the University of Memphis.
- Mendes, M. S., & Lucena, W. G. L. (2022). Legibilidade nos releases de resultados das Companhias Brasileiras Abertas. *Revista Contabilidade, Gestão e Governança*, 25(3), 416-434. <http://dx.doi.org/10.51341/cgg.v25i3.2918>.
- Monteiro, J., Rengel, R., Sousa, A., & Borba, J. (2020). Não basta ler, é preciso compreender: Um enfoque na legibilidade do resultado e do EBITDA. *Revista Universo Contábil*, 16(3), 31-49. <http://dx.doi.org/10.4270/ruc2020314>.
- Moreno, A., & Casasola, A. (2015). A readability evolution of narratives in annual reports: A longitudinal study of two Spanish companies. *Journal of Business and Technical Communication*, 29, 1-34. <http://dx.doi.org/10.1177/1050651915620233>.
- Noh, M. (2021). Culture and annual report readability. *International Journal of Accounting & Information & Management*, 29(4), 583-602. <http://dx.doi.org/10.1108/IJAIM-05-2021-0094>.
- Razik, T. A. (1969). A study of American newspaper readability. *Journal of Communication*, 19(4), 317-324. <http://dx.doi.org/10.1111/j.1460-2466.1969.tb00854.x>.
- Rennekamp, K. (2012). Processing fluency and investors' reactions to disclosure readability. *Journal of Accounting Research*, 50(5), 1319-1354. <http://dx.doi.org/10.1111/j.1475-679X.2012.00460.x>.
- Rush, R. T. (1985). Assessing readability: Formulas and alternatives. *The Reading Teacher*, 39(3), 274-283. <https://www.jstor.org/stable/20199072>
- Santos, E., Calixto, L., & Bispo, M. (2019). Impact of the Brazilian Accounting Pronouncements Committee's New Guideline (OCPC 07) on the downsizing of companies' Notes. *Revista Contabilidade & Finanças*, 30(79), 58-72. <http://dx.doi.org/10.1590/1808-057x201806650>.
- Schrifer, K. A. (1989). Evaluating text quality: The continuum from text-focused to reader-focused methods. *IEEE Transactions on Professional Communication*, 32(4), 238-255. <http://dx.doi.org/10.1109/47.44536>.
- Securities and Exchange Commission. (1998). *A plain english handbook: How to create clear SEC disclosure documents*. Office of Investor Education and Assistance. <https://www.sec.gov/pdf/handbook.pdf>
- Sena, T., Santana Jr., J., & Freitas, S. (2023). Gerenciamento de resultados e legibilidade das notas explicativas: Manipulações nas evidenciações contábeis. *Revista de*

Educação e Pesquisa em Contabilidade, 17(1), 84-109. <http://dx.doi.org/10.17524/repec.v17i1.3214>.

Silva, R., Castro, L., Domingos, S., & Ponte, V. (2020). Nível de Disclosure e forma de apresentação das notas explicativas: Um estudo à Luz da OCPC 07. *Revista Contabilidade Vista & Revista*, 31(3), 74-96. <http://dx.doi.org/10.22561/cvr.v31i2.5354>.

Sinônimos. (2016). *Dicionário online de sinônimos*. <https://www.sinonimos.com.br/>

Smith, M., & Taffler, R. (1992). Readability and understandability: Different measures of the textual complexity of accounting narrative. *Accounting, Auditing & Accountability Journal*, 5(4), 84-98. <http://dx.doi.org/10.1108/09513579210019549>.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, 290-312. <http://dx.doi.org/10.2307/270723>.

Soepriyanto, G., Tjokroaminoto, S., & Zudana, A. (2021). Annual report readability and accounting irregularities: Evidence from public listed companies in Indonesia. *Journal of Financial Reporting and Accounting*, 19(5), 793-818. <http://dx.doi.org/10.1108/JFRA-01-2020-0006>.

Soper, F., & Dolphin Jr., R. (1964). Readability and corporate annual reports. *The Accounting Review*, 39(2), 358-362. <https://www.jstor.org/stable/243526>

Souza, J., & Borba, J. (2022). Readability como medida de complexidade textual: Determinantes e evidências em empresas brasileiras. *Revista Contabilidade & Finanças*, 33(88), 112-129. <http://dx.doi.org/10.1590/1808-057x202114180>.

Sun, Y., Wang, J., & Huang, K. (2022). Does IFRS and GRI adoption impact the understandability of corporate reports by Chinese listed companies? *Accounting and Finance*, 62(2), 2879-2904. <http://dx.doi.org/10.1111/acfi.12908>.

Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics*. Pearson.

Telles, S., & Salotti, B. (2020). Intelligibility VS Readability: Understandability measures of financial information. *Revista Universo Contábil*, 16(2), 110. <http://dx.doi.org/10.4270/ruc.2020209>.

United Nations Educational, Scientific and Cultural Organization. (2022). *World portuguese language day*. <https://www.unesco.org/en/days/portuguese-language-day>

Worthington, J. S. (1977). The readability of footnotes to financial statements and how to improve them. *Journal of Reading*, 20(6), 469-478. <https://www.jstor.org/stable/40010895>

Supplementary Material

Supplementary material accompanies this paper: <https://doi.org/10.7910/DVN/X55H6J>

Supplementary Data 1. Survey Questions and Possible Answers for Treatments

Supplementary Data 2. Glossary of Accounting Terms Used (in Portuguese)

Supplementary Data 3. Phase 2 Extra Questions

Supplementary Data 4. Phase 3 Pairs of Sentences

Supplementary Data 5. Database

Financial support:

This work was supported by a PhD scholarship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Open science:

Telles, Samantha Valentim; Salotti, Bruno Meirelles, 2024, "Supplementary Data - Readability and Understandability of Notes to Financial Statements", <https://doi.org/10.7910/DVN/X55H6J>, Harvard Dataverse, V1

Conflicts of interest:

The authors have no conflicts of interest to declare.

Copyrights:

RBGN owns the copyrights of this published content.

Plagiarism analysis:

RBGN performs plagiarism analysis on all its articles at the time of submission and after approval of the manuscript using the iThenticate tool.

Authors:

Samantha Valentim Telles, PhD, Universidade de São Paulo, São Paulo, Brasil.

E-mail: samantha.telles@usp.br

Bruno Meirelles Salotti, PhD, Universidade de São Paulo, São Paulo, Brasil.

E-mail: brunofea@usp.br

Authors' contributions:

1st author: Definition of research problem; development of hypotheses or research questions (empirical studies); definition of methodological procedures; data collection; literature review; statistical analysis; analysis and interpretation of data; manuscript writing.

2nd author: Definition of research problem; development of hypotheses or research questions (empirical studies); definition of methodological procedures; critical revision of the manuscript; manuscript writing.