



ELSEVIER

Contents lists available at ScienceDirect

Journal of Accounting and Economics

journal homepage: www.elsevier.com/locate/jaeA plain English measure of financial reporting readability[☆]

Samuel B. Bonsall IV^{a,b}, Andrew J. Leone^{c,*}, Brian P. Miller^d,
Kristina Rennekamp^e

^a The Ohio State University, United States^b The Pennsylvania State University, United States^c University of Miami, United States^d Indiana University, United States^e Cornell University, United States

ARTICLE INFO

Article history:

Received 5 February 2015

Received in revised form

8 March 2017

Accepted 21 March 2017

Available online 5 April 2017

Keywords:

Plain English reporting

Textual analysis

Regulation

Mandatory disclosure

ABSTRACT

We propose a new measure of readability, the Bog Index, which captures the plain English attributes of disclosure (e.g., active voice, fewer hidden verbs, etc.). We validate this measure using a series of controlled experiments and an archival-based regulatory intervention to prospectus filing readability. We also demonstrate the importance of understanding the underlying drivers of quantity-based measures of readability. In particular, we caution researchers that a vast amount of the variation in Form 10-K file size over time is driven by the inclusion of content unrelated to the underlying text in the 10-K (e.g., HTML, XML, PDFs).

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

A growing body of research in accounting and finance examines whether, and to what extent, qualitative attributes of corporate communication (e.g., tone, readability) affect the decision-making of investors and information intermediaries. Although relatively clear guidance exists for measuring linguistic attributes of financial disclosure such as tone (e.g., Loughran and McDonald, 2011; Henry and Leone, 2016), the most appropriate measure of readability is less clear. Researchers have primarily selected from a limited set of existing readability measures that are based on either writing clarity (e.g., the Gunning Fog Index) or disclosure quantity (e.g., file size of the filing). We extend this literature by introducing a new measure of readability, the Bog Index, which is designed to capture the plain English attributes of disclosure. We then use controlled experiments, a regulatory intervention, and archival-based capital market tests to validate and compare this new measure with existing readability measures.¹

[☆] We are grateful for discussions and helpful comments from Rob Bloomfield, Kimball Chapman, Khrystyna Bochkay, Roman Chychyla, Wayne Guay (editor), Patrick Hopkins, Michelle Hutchens, Robert Libby, Jeff McMullin, Sundaresh Ramnath, Lorien Stice-Lawrence, K.R. Subramanyam, Eric Weisbrod, an anonymous reviewer, and workshop participants at the 11th Annual Conference on Empirical Legal Studies, Baruch College, Chinese University of Hong Kong, Hong Kong Polytechnic University, The Pennsylvania State University, Singapore Management University, and the University of Buffalo. We also express gratitude to the StyleWriter software engineers for programming assistance and Bradley Carrico, Mike Durney, Adam Hadley, Eric Holzman, Sharadh Ramesh, and Barrett Wheeler for research assistance. Brian Miller gratefully acknowledges financial support from the PwC Faculty Fellowship. The parsing code and Bog Index data are available for download on Brian Miller's website at <https://kelley.iu.edu/bpm/index.html>.

* Corresponding author.

E-mail address: a.leone@miami.edu (A.J. Leone).¹ The Bog Index data are available for download on Brian Miller's website at <https://kelley.iu.edu/bpm/index.html>.

Although the exact definition varies, there is general agreement that “readability” refers to the ease with which a reader can process and comprehend written text. In terms of financial disclosure readability, the Securities and Exchange Commission (SEC) provides some very specific guidance in recommending that managers employ plain English attributes by avoiding writing constructs like passive voice, weak or hidden verbs, superfluous words, legal and financial jargon, numerous defined terms, abstract words, unnecessary details, lengthy sentences, and unreadable design and layout in their financial disclosures (SEC, 1998b). Lower frequencies or the absence of these constructs are consistent with clear writing that language experts have asserted is critical to users’ understanding of written materials in a variety of contexts including the military, healthcare, and law (DuBay, 2004). Despite the strong support for plain English from linguists and regulators, few studies employ plain English measures as proxies for readability.

We begin by evaluating how well alternative measures of readability capture the plain English attributes recommended by linguistic experts and the SEC. Most readability studies use the Gunning (1952) Fog Index as a proxy for readability when testing the causes and consequences of financial reporting readability.² Given this measure is based on average sentence length and the proportion of words with three or more syllables, the measure does in theory capture two broad plain English attributes. However, recent research by Loughran and McDonald (2014a) raises valid criticisms that the complex word component of the Fog Index treats all words with three or more syllables as “complex” even though the meaning of many of these multisyllabic words (e.g., Company) would be well understood by even the least sophisticated investors. As such, only the sentence length portion of the Fog Index appears to map directly into the SEC’s plain English guidelines.

In addition to measuring the writing clarity component of readability using the Fog Index, many other studies rely on the quantity of textual disclosure to measure readability. Although historically these quantity-based measures have focused on the number of words contained in a financial filing, more recently Loughran and McDonald (2014a, p. 1644) advocate the use of “file size of the 10-K as an easily calculated proxy for document readability.” It is important to note that both of these quantity-based measures are limited as measures of plain English readability because they only capture a single plain English attribute: superfluous words. Further, the SEC cautions that quantity of disclosure measures may have some shortcomings in capturing writing clarity. Specifically, the SEC notes that there may be a trade-off between writing clarity and the quantity of disclosure: “writing a disclosure in plain English can sometimes increase the length of particular sections...” (SEC, 1998a). In summary, both the Fog Index and quantity-based measures of readability capture some attributes of plain English disclosures, but they are not comprehensive.

Given the limitations of these measures in capturing plain English attributes, we propose an alternative multi-faceted measure of disclosure clarity, the Bog Index, which is based on plain English writing principles and captures the spirit of almost all of the SEC’s guidelines regarding clear communication with investors. The Bog Index is derived from a commercial software program, *StyleWriter*, which captures attributes specifically mentioned in the SEC Plain English Handbook including sentence length, passive voice, weak verbs, overused words, complex words, and jargon (SEC, 1998b). Among these many features of the Bog Index, one unique aspect of the measure stems from the way in which word complexity is determined. Rather than assuming all multi-syllabic words are complex, as is done in computing the Fog Index, word complexity is instead determined by word familiarity based on a proprietary list of over 200,000 words. Thus, the Bog Index measure of writing clarity overcomes the major criticism of the Fog Index related to capturing word complexity based on syllable counts alone.

Prior research has employed a subset of the plain English attributes used in the Bog Index. For instance, Miller (2010) uses output from earlier and much more limited versions of the same software used to generate the Bog Index to create a somewhat ad hoc measure of plain English reporting.³ More recently, Loughran and McDonald (2014b) creates a measure of plain English writing (*LM PE Index*), based on a small subset of the plain English attributes highlighted by the SEC. As discussed in more detail in subsequent sections, compared to these earlier measures, the Bog Index provides a much more comprehensive set of factors and is calculated using a pre-programmed algorithm that eliminates researcher discretion related to how to calculate the measure.

One of the challenges that prior researchers have faced in validating readability measures is that archival-based capital market tests, which are used for proxy validation, are joint tests that the proxy captures the construct of interest and that the construct is related to the outcome. For instance, in our setting it is possible that the Bog Index is a valid proxy for readability, but that readability does not affect a capital market outcome such as stock return volatility. As such, we use several complementary approaches to validate the Bog Index as a proxy for readability including an experiment, a quasi-experiment, and an archival examination of a readability-related regulatory intervention.

The experiment and the quasi-experiment enable us to examine the effects of the Bog Index on surrogates for investors while holding constant other potentially influential variables (Libby et al., 2002). The evidence shows that participants who receive a more readable disclosure, as measured by the Bog Index, rate the disclosure to be significantly easier to read than participants who receive a less readable disclosure. This suggests that the Bog Index captures financial statement readers’ internal evaluations of readability.

² Numerous studies have examined the impact of financial reporting complexity on retail investors (Lawrence, 2013; Rennekamp, 2012; Miller, 2010), sell-side equity analysts (Bozanic and Thevenot, 2015; Lehavy et al., 2011), rating agencies (Bonsall and Miller, 2017), and the voluntary disclosure behavior of firms (Guay et al., 2016).

³ This pre-cursor plain English measure of readability created from components of the earlier versions of *StyleWriter* is far less comprehensive than the components calculated by the more recent software package. As such, in this study, we focus on the more comprehensive standardized Bog Index measure as it subsumes this prior measure.

Download English Version:

<https://daneshyari.com/en/article/5086522>

Download Persian Version:

<https://daneshyari.com/article/5086522>

[Daneshyari.com](https://daneshyari.com)