# THE EFFECT OF REPUTATION ON READABILITY OF ANNUAL REPORTS FOR S&P100 FIRMS – A COMPARISON AMONG MEASURES

*Xuan Huang, College of Business, California State University, Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840, 562-985-2756, xuan.huang@csulb.edu*
*Sudha Krishnan, College of Business, California State University, Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840, 562-985-5757, Sudha.krishnan@csulb.edu*
*Ping Lin, College of Business, California State University, Long Beach, 1250 Bellflower Blvd., Long Beach, CA 90840, 562-985-4560, ping.lin@csulb.edu*

## ABSTRACT

In this paper, we compared three main measures of readability: Fog index, Flesch score, and Log file size. We computed these measures for the annual financial reports of S&P100 firms from 2016 to 2020 and reported trends in readability measures such as increasing average words per sentence and file size. We also used the reputation score provided by Fortune's "America's Most Admired Companies List" and investigated the relationship between firm reputation and annual report readability. Based on Log file size measure of readability, the more reputable firms also present more readable reports.

**Keywords:** Readability, Readability measures, Reputation, Annual reports

## INTRODUCTION

The business environment and practice have become exceedingly complicated in the past three decades. For example, new financial instruments, complicated corporate structures and transactions have come into existence. Accordingly, the financial statements that reflect the operating results of these companies have become more difficult to comprehend. For one aspect, the language used in annual reports and other disclosure documents has become more complex and potentially less readable. These documents are carefully vetted by lawyers and treated like liability documents. Commissioner Hunt mentioned in his speech on February 6, 1997 at the first Mergers & Acquisitions conference in Florida "… many of us have lost sight of the fact that the disclosure documents that are filed with the SEC every year are not only liability documents -- but are intended to be one of the primary ways that the corporate community communicates with investors." [1]

A key requirement of financial reporting is to provide information clearly and concisely. Both Concepts Statement 2 and the *Framework* (1989) included **understandability**, "a qualitative characteristic that enables users to comprehend the information and therefore make it useful for making decisions". One way to improve the understandability of financial statements is through enhancing document readability. Recent academic research has provided a few means of measuring the level of readability. This article aims to introduce these measures and examine the trend of readability in 10-K filings with a S&P 100 firms.

## HISTORY OF PLAIN ENGLISH MOVEMENT IN SEC

In the 1990s, the SEC under the leadership of Arthur Levitt attempted to change the corporate culture of America by promoting plain English in SEC filings. In July 1997, the SEC employees had a contest to find the "worst piece of gobbledygook" reported to the SEC and translate to plain English. [2] In 1998, the SEC issued Release 33-7497 requiring issuers to write sections of the prospectuses in plain English [3] along with a handbook on how to create clear SEC disclosure documents. [4] After a decade of implementation, the SEC commissioned four focus groups in 2008 to evaluate whether the plain English handbook and the attempt to change corporate culture were effective. However, majority of the focus group participants did not use the annual reports and proxy documents for decision making as they reported that the materials were too lengthy and complex. Many of them also felt that the lengthy documents and verbosity are often used to hide something. [5]

Readability and writing in plain English again became the emphasis when the SEC required a Compensation Disclosure & Analysis (CD&A) to be included to the proxy document filed by corporates. In 2006, the SEC issued Release no. 33-8732A amending disclosure requirements regarding executive compensation, related party transactions, director independence and other corporate governance issues [25]. They added rules 13a-20 and 15d-20 under the Exchange Act that required companies to disclose the information in clear plain English. [6]

In 2007, the Division of Corporate Finance at the SEC reviewed 350 companies on their CD&As and concluded that it could be better organized to make it more readable for both the individual and institutional investors. For example, one of their suggestions was to increase the font size in the tables and footnotes, wherever practical, to make them more readable. [7]

The same year, Chris Cox, the SEC Chairman created Committee on Improvements of Financial Reporting (CIFiR) to "provide specific recommendations on how unnecessary complexity in the financial reporting system can be reduced and how the system can be made more useful to investors." In their 2008 report, the CIFiR strongly suggested that the SEC require an executive summary in plain English for both the 10-K and 10-Q that would provide a roadmap to a more detailed discussion later in the document. This would make the documents more understandable to help investors navigate the SEC filing. [8]

All of these actions suggest that the SEC is aware that with complex markets, the communications with the investor is also getting complicated. Since 1998, they have extended plain English usage to more corporate filings. In 2010, the White House passed the Plain Writing Act 2010 [9] requiring all federal communications to be written in plain English [10]. With the federal authorities and SEC implementing plain English, the issue arises as to how management can evaluate and ensure that their communications with the shareholders is readable.

## MEASURE DESCRIPTIONS

Measures of readability were developed in the 1950s and were traditionally used in linguistic research. It was only until recently that such measures are applied to evaluate business communications. Academic research has linked earning levels to readability. Annual reports of firms with lower earnings are found to be less readable; while firms with easy to read reports are likely to have more persistent positive earnings [22]. Research has also suggested that companies with better written financial documents have more significant increase in trading activity around filing dates as small investors find it easier to process the information [24]. Similarly, Lawrence [21] finds that individual investors tend to invest more in companies with easier to read and shorter 10-K. In many of these studies, the four most

commonly used measures of readability are Fog Index, Flesch Score, Plain-English Measure and Length of Document.

**Fog Index**

The first measure of plain English is Fog index - developed by Gunning [17] and is defined as:
*Fog index* = (Average number of words per sentence+ Percentage of complex words) x 0.4
According to [17] Fog index represents the number of formal education years a reader of average intelligence would need to read the text once and understand its meaning. The index is affected by two parameters: average number of words in sentences and the percentage of complex words among the entire text. Complex words are defined as words with three syllables or more. In general, a Fog index of 18 means the text is unreadable; 14–18 means it is difficult to read; 12–14 means the level of readability is ideal for average readers; 10–12 means the text is acceptable to read; and 8–10 means the level is easy and meant for children. Based on a recent study on readability, the mean Fog index of over 55,000 annual reports between 1993 and 2003 is 19.3 [22]. Such a high Fog index indicates average annual reports are almost impossible for a reader with average intelligence to comprehend after one reading.

To further investigate whether either one of two components of Fog index can represent readability, we also use the following measures.
*AWS* = Average number of words per sentence
*Complex%* = Percentage of complex words

**Flesch Score**

A second but also related readability measure is Flesch Reading Ease Score (Flesch score), defined as:
*Flesch* = 206.835 – (1.015 x average number of words per sentence) – (84.6 x average number of syllables per word)

The Flesch score ranges between 0 – 100. The higher the score, the easier it is to read the document. Scores below 60 are considered difficult. A score below 30 is considered confusing and requires at least a college degree to understand the material. The Flesch Score is also widely used across disciplines [14]. It is so popular that the score calculator is included the Microsoft Office software. To improve the score, it is suggested that the writer use shorter sentences and words.

**File Size**

A recent academic study has proposed 10-K file size as another measure of readability [23]. The authors argue that since business routinely uses complex words with three or more syllables, such company, operation and management, the Fog and Flesch index can't capture the readability of financial statements. Instead, they test the file size of "complete submission text file" on SEC's EDGAR website and find that firms with large file size have higher earnings surprises and analyst dispersion, indicating the document is less readable.
*LogSize* = Log (File size of "Complete submission text file" on SEC's EDGAR website)

## REPUTATION

Reputation refers to the external parties' "judgment of a corporation based on assessments of the financial, social, and environmental impacts attributed to the corporation over time" [12]. Reputation effect indicates the impact of reputation on firm behavior and how it is treated by other parties. Analytical research on reputation utilizes game theory to explain the lower agency cost enjoyed by reputable players [16] [18]. For example, a firm that is always making on time repayments should gain access to lower cost of debt [15]. Empirically, Anigner et al. [11] finds a negative correlation between company reputation and cost of debt financing. Additional research on reputation has linked it to better investor perception, lower cost of equity and higher financial reporting quality, all in the direction predicted by agency theory.

Specifically, Cao et al. (2012) [13] posits that high company reputation can result in higher financial reporting quality through at least the following ways. First, reputable firms emphasize accountability and creditability, therefore employees hired, especially accountants and internal auditors, are likely to have higher ethics. Second, with such honest and transparent culture, accounting department is more likely to cooperate with external auditors and board of directors are more likely to invest in technologies that enhance reporting quality. With a sample of over 8000 Fortune 1000 company-year observations, they find that companies with higher reputations are less likely to misstate their annual financial statements.

In this paper, we argue that readability is one construct of financial reporting quality. An annual report that is long and difficult to understand is not going to be perceived as of high quality by investors. Reputable companies who care about maintaining their reputation would emphasize readable public financial reports. Instead of hiding the true economic state of a company behind lengthy and complex language, reputable companies use the financial statements to convey their financial position honestly without any ulterior motives. Therefore, consistent with above arguments in [13], more reputable companies would likely generate more readable reports through selective hiring, continuous investment in reporting system and fostering ethical corporate culture.

**H1: The readability measure of annual financial statements will be higher for companies with higher reputation.**

## SAMPLE DESCRIPTION

To analyze different measures of readability, we choose the Standard and Poor's 100 firms in 2021. These firms are a sub-set of the S&P 500 containing 100 U.S. companies with an unadjusted market capitalization of $13.1 billion or greater, investable weight factor[1] of at least 0.1 and positive earnings in last quarter and last four quarters cumulatively. In general, these firms represent large-cap companies with stable earnings. For reputation scores, we use Fortune's "America's Most Admired Companies List" from 2016 to 2020.

We compiled the annual reports of S&P 100 companies for the years 2012 to 2020. Companies' 10-K files were downloaded from SEC Edgar website in XBRL format. We then compared the various

[1] Investible Weight Factor (IWF) = Total free holding available for investors/Total number of shares

readability measures to test which company reports were the most readable and which measures were the most consistent. The sample consists of 737 firm-year observations from 2012 to 2020.

## RESULTS

There are five measures of readability: Fog, Flesch, Average words per sentence, Percent complex words, and File size. Table 1 reports descriptive statistics for readability measures. Each year, we obtain the mean, standard deviation, $25^{th}$, $50^{th}$, and $75^{th}$ percentile of the variables in our sample. We then report the annual average of the cross-sectional statistics for the variables.

**TABLE 1 DESCRIPTIVE STATISTICS**

| | Mean | STD | 25% | 50% | 75% |
|---|---|---|---|---|---|
| Fog | 31.833 | 3.640 | 30.168 | 32.163 | 34.047 |
| Flesch | 6.873 | 11.403 | 3.764 | 8.073 | 11.791 |
| Average words per sentence | 19.162 | 3.608 | 17.362 | 19.425 | 21.059 |
| Percent complex words | 0.317 | 0.017 | 0.308 | 0.318 | 0.327 |
| File Size | 1.117 | 1.133 | 0.780 | 1.251 | 1.732 |

Note: This table reports descriptive statistics of readability measures. Fog index is equal to (Average number of words per sentence+ Percentage of complex words) x 0.4. Flesch score is equal to 206.835 – (1.015 x average number of words per sentence) – (84.6 x average number of syllables per word). Average words per sentence is the number of words in the 10-K divided by a count of sentence terminations. Percent complex words is the percentage of 10-K words with more than two syllables. File Size is the natural log of the text document file size in megabytes.

Table 2 presents the mean of readability measures by year of 10-K filings. Fog index has been stable and increased slightly in recent years. Flesch was constant but decreased a lot in the recent two years. Average words per sentence have been steadily increasing for these years. Percent complex words was relatively stable. File size kept increasing during the sample years. Overall, based on the change of mean readability measures, 10-K filings are less readable during the sample years. Table 3 shows mean of each readability measure by industry. We follow Fama and French 12 industry classifications. Across different industries, Fog index, Average words per sentence, and Percent complex words do not have many variations. While Flesch and File Size are different across industries. Industries with lower Flesch index tend to have larger File Size, consistent with the readability level. The lower Flesch, less readable and the larger File Size, less readable of reports.

To examine the relation between companies' reputation and readability of 10-K filings, we conduct correlation analysis. Table 4 reports the Spearman correlations among key variables including reputation and readability measures. The correlation between Reputation and Fog index is significant positive at 0.09. Reputation is negatively related to Flesch, Percent complex words, and File size respectively, at significance level of 10%. Reputation is positively correlated with Average words per sentence. Among readability measures, Fog and Flesch are significantly positive related. File Size is positively related to Fog, Average words per sentence, Percent complex words, but negatively related to Flesch. Overall, the correlation among key readability measures is consistent with the difficulty of readability levels. Loughran and McDonald [23] argue that traditional readability measures like the Fog Index are poorly

specified to evaluate financial documents, and they propose that the 10-K file size provides a better proxy for readability than traditional measures. Based on the correlation between reputation and File Size, our hypothesis is supported: the more reputable firms tend to report more readable 10-K reports.

**TABLE 2 MEAN OF READABILITY MEASURES BY YEAR**

| Year | Fog | Flesch | Average words per sentence | Percent complex words | File Size |
|---|---|---|---|---|---|
| 2012 | 31.125 | 9.927 | 18.516 | 0.315 | 0.618 |
| 2013 | 31.355 | 9.571 | 18.597 | 0.319 | 0.805 |
| 2014 | 31.362 | 8.901 | 18.563 | 0.320 | 0.951 |
| 2015 | 31.155 | 9.343 | 18.336 | 0.320 | 1.136 |
| 2016 | 31.207 | 9.657 | 18.399 | 0.320 | 1.114 |
| 2017 | 31.412 | 10.073 | 18.668 | 0.319 | 1.086 |
| 2018 | 31.889 | 7.132 | 19.035 | 0.321 | 1.277 |
| 2019 | 33.442 | -0.354 | 20.847 | 0.315 | 1.529 |
| 2020 | 33.092 | -0.348 | 20.998 | 0.302 | 1.359 |

Note: This table presents the mean of readability measures by year of 10-K filings. Fog index is equal to (Average number of words per sentence+ Percentage of complex words) x 0.4. Flesch score is equal to 206.835 – (1.015 x average number of words per sentence) – (84.6 x average number of syllables per word). Average words per sentence is the number of words in the 10-K divided by a count of sentence terminations. Percent complex words is the percentage of 10-K words with more than two syllables. File Size is the natural log of the text document file size in megabytes.

**TABLE 3 MEAN OF READABILITY MEASURES BY INDUSTRY**

| | Fog | Flesch | Average words per sentence | Percent complex words | File Size |
|---|---|---|---|---|---|
| *Consumer Nondurables* | 31.454 | 6.580 | 18.771 | 0.317 | 1.386 |
| *Consumer Durables* | 33.478 | 7.301 | 21.029 | 0.311 | 1.199 |
| *Manufacturing* | 31.355 | 11.675 | 19.016 | 0.309 | 0.599 |
| *Energy* | 30.723 | 6.468 | 18.170 | 0.314 | 1.651 |
| *Chemicals and Allied Products* | 34.893 | 1.924 | 21.482 | 0.335 | 1.728 |
| *Business Equipment* | 33.194 | 4.988 | 20.333 | 0.321 | 0.856 |
| *Telecommunication* | 34.214 | 7.377 | 21.484 | 0.318 | 0.978 |
| *Utilities* | 33.162 | 1.651 | 20.211 | 0.324 | 1.727 |
| *Wholesales and Retails* | 30.865 | 10.353 | 18.208 | 0.316 | 0.862 |
| *Healthcare* | 30.827 | 7.061 | 18.110 | 0.318 | 1.151 |
| *Finance* | 31.822 | 4.895 | 19.238 | 0.315 | 1.092 |
| *Other* | 31.884 | 7.561 | 19.238 | 0.316 | 1.081 |

Note: This table shows mean of each readability measure by industry. Industries are classified following Fama and French 12 industry classifications. Fog index is equal to (Average number of words per

sentence+ Percentage of complex words) x 0.4. Flesch score is equal to 206.835 – (1.015 x average number of words per sentence) – (84.6 x average number of syllables per word). Average words per sentence is the number of words in the 10-K divided by a count of sentence terminations. Percent complex words is the percentage of 10-K words with more than two syllables. File Size is the natural log of the text document file size in megabytes.

**TABLE 4 SPEARMAN CORRELATION**

| | Reputation | Fog | Flesch | Average words per sentence | Percent complex words |
|---|---|---|---|---|---|
| Fog | **0.090** | | | | |
| Flesch | **-0.046** | **-0.613** | | | |
| Average words per sentence | **0.095** | **0.982** | **-0.553** | | |
| Percent complex words | **-0.055** | **0.206** | **-0.388** | 0.044 | |
| File Size | **-0.046** | **0.194** | **-0.423** | **0.194** | **0.054** |

Note: This table reports Spearman correlation among variables. Fog index is equal to (Average number of words per sentence+ Percentage of complex words) x 0.4. Flesch score is equal to 206.835 – (1.015 x average number of words per sentence) – (84.6 x average number of syllables per word). Average words per sentence is the number of words in the 10-K divided by a count of sentence terminations. Percent complex words is the percentage of 10-K words with more than two syllables. File Size is the natural log of the text document file size in megabytes. $p < 0.10$ is bolded.

## CONCLUSIONS

As the financial markets grow more complicated over time, it is essential to have readable written communication with the investors. We observe that the quality of financial disclosure has been an important issue with the SEC and the Federal government, as they have amended existing rules and issued new disclosure requirements to enhance readability of financial reports [25]. We summarize five different measures of readability from academic research and study the veracity of such disclosures of the S&P100 firms from 2012 to 2020. We analyze whether companies that have a higher reputation tend to report more readable 10-K reports using various readability measures. We find that File Size – considered one of the better measures of readability and reputation are correlated indicating that the more reputable firms tend to report more readable 10-K reports.

**Reference:**

[1] http://www.sec.gov/news/speech/speecharchive/1997/spch151.txt
[2] SEC News Digest, Issue 97-136 July 16, 1997, Chair Levitt issues challenge to SEC staff to heighten awareness of need for plain English.
[3] http://www.sec.gov/rules/final/33-7497.txt
[4] http://www.sec.gov/pdf/handbook.pdf
[5] Focus Groups about Plain English documents, Final report, May 2008
[6] http://www.sec.gov/rules/final/2006/33-8732a.pdf
[7] http://www.sec.gov/divisions/corpfin/guidance/execcompdisc losure.htm

[8] http://www.sec.gov/about/offices/oca/acifr/acifr-finalreport.pdf
[9] http://braley.house.gov/press-release/obama-signs-braley%E2%80%99s-plain-language-act
[10] http://www.plainlanguage.gov/plLaw/law/index.cfm
[11] Anginer, D., Warburton, A. J., & Yildizhan, C., 2019, Corporate reputation and cost of debt. Unpublished paper, The World Bank, Syracuse University, and University of Michigan.
[12] Barnett, M. L., Jermier, J. M., & Lafferty, B. A., 2006, Corporate reputation: The definitional landscape. *Corporate Reputation Review*, 9, 26–38
[13] Cao, Y., Myers, L. A., & Omer, T. C., 2012, Does company reputation matter for financial reporting quality? Evidence from restatements. *Contemporary Accounting Research*, 29, 956–990.
[14] Dubay, 2007, Unlocking Language. Charleston, South Carolina, BookSurge Publishing.
[15] Eaton, J., & Gersovitz, M., 1981, Debt with potential repudiation: Theoretical and empirical analysis. *Review of Economic Studies*, 48, 289–309.
[16] Fama, E. F., 1980, Agency problems and the theory of the firm. The Journal of Political Economy 88(2): 288–307.
[17] Gunning, R., 1952, The Technique of Clear Writing, New York, McGraw-Hill.
[18] Holmstrom, B., 1982, Moral hazards in teams. *Bell Journal of Economics* 13(2): 324–40.
[19] Huang, X., Krishnan, S., and Lin, P., 2016, Readability of Compensation Discussion and Analysis: The Impact of the Plain English Initiative. *International Journal of Services and Standards,* 11(1): 20-42.
[20] Laksmana, I., Tietz, W. and Yang, Y., 2012, Compensation discussion and analysis (CD&A): Readability and management obfuscation, *Journal of Accounting and Public Policy*, 31(2), 185-203.
[21] Lawrence, A., 2013, Individual investors and financial disclosure. *Journal of Accounting & Economics,* 56(1), 130-147.
[22] Li, F., 2008, Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and Economics* 45, 221-247.
[23] Loughran, T., and B. McDonald., 2014, Measuring readability in financial disclosures. *Journal of Finance* 69: 1643-1671.
[24] Miller, B., 2010, The effects of reporting complexity on small and large investor trading, *The Accounting Review* 85, 2107-2143.
[25] Robinson, J. R., Xue, Y., and Yu, Y., 2011, Determinants of disclosure noncompliance and the effect of the SEC review: Evidence from the 2006 mandated compensation disclosure regulations, *The Accounting Review* 86, 1415-1444.