

A Content Analysis of the Modernization of Regulation S-K Items 101, 103, and 105

**Orry Swift
Lamar University**

**Ricardo Colon
Lamar University**

In 2020, the U.S. Securities and Exchange Commission (SEC) adopted amendments to modernize the disclosure requirements of Regulation S-K Items 101, 103, and 105. These amendments aim to enhance the readability of disclosure documents and provide investors with more relevant information. This study conducts a content analysis of the modernized disclosure items to assess the extent and nature of the changes implemented. A sample of annual reports filed before and after the amendments was examined to determine whether the SEC meets its stated goals of improving the readability of disclosure documents and discouraging repetition and immaterial information. We conclude that the stated SEC goals for modernization were not met in terms of textual characteristics. These findings provide insights into the effectiveness of the SEC's efforts to update disclosure rules and improve investor's informational value.

Keywords: content analysis, readability, disclosures, modernization, textual characteristics, Regulation S-K

INTRODUCTION

Regulation S-K is one of the central regulatory frameworks governing the disclosure requirements for public companies in the United States. It outlines the use of qualitative descriptions of financial reports, whereas Regulation S-X focuses on quantitative descriptions. Items 101, 103, and 105 of Regulation S-K pertain to a description of a company's business, legal proceedings, and risk factors, respectively (SEC, 2020). Over time, these disclosure items have become outdated, with redundant requirements and lack of alignment with investors' contemporary information needs.

In response, the SEC proposed and adopted amendments to modernize these key disclosure items in 2020. The objective was to revise companies' disclosure obligations to provide more relevant and concise information to investors while reducing compliance burdens for companies. The following is a short summary of the amendments that are analyzed in this study:

Item 101 Revision:

In filings made after a registrant's initial filing, registrants may provide an update on the general development of the business rather than a full discussion. The update must disclose all material developments that have occurred since the registrant's most recent filing and contain a full discussion

of the general development of its business, and incorporate by reference that prior discussion (SEC, 2020).

Item 103 Revision:

The required information may be provided by hyperlink or cross-reference to legal proceedings disclosure located elsewhere within the same document, such as in Management Discussion & Analysis (MD&A) or a note to the financial statements to avoid duplicative disclosure (SEC, 2020).

Item 105 Revision:

A summary risk factor disclosure of no more than two pages is required if the risk factor section exceeds 15 pages (SEC, 2020).

While the amendments to Regulation S-K aim to assist small reporting firms to maintain compliance, many provisions also affect large firms. We analyze the provisions within these amendments that may materially impact text disclosures within the 10-K reporting framework. Section II provides a background on the textual analysis of financial statement disclosures in the existing accounting literature. Section III describes the content analysis measures from prior research (i.e., fog, length, boilerplate, and tone) to develop the hypotheses and methodology. Section IV discusses the results and their potential implications. In Section V, we conclude that the stated SEC goals for modernization were not met in terms of textual characteristics.

PRIOR LITERATURE

Over the past two decades, accounting content analysis has primarily focused on annual reports in the public sector and examined one of three aspects of financial disclosures: the amount of disclosure, tone, or transparency (readability) of reports (Li, 2010). The amount of disclosure addresses the physical length of an annual report, and researchers often use it as a proxy for complexity or readability of the report (Peterson, 2012; Lee, 2012; Leuz & Schrand, 2009; Miller, 2010). Research on the tone of financial disclosures in accounting can employ one of two methods: a rules-based or “dictionary” approach that classifies text into different categories based on specific rules, or a statistical approach that examines the correlations between the text and document type (Jurafsky and Martin, 2014; Mitchell, 2006). Studies employing readability measures offer a more refined approach than the amount of disclosure, often measuring the reading level based on content complexity.

This study adopted a dictionary approach and used the Linguistic Inquiry and Word Count (LIWC) software to analyze the tone, language metrics, and cognitive processes of annual financial reports. Using this approach, the aim was to differentiate between the use of these three dimensions of textual content before and after modernization to determine how the provisions within modernization meet the stated SEC goals for readability, repetition, and materiality. To accomplish this task, a content analysis of financial reports from a subset of large accelerated filers (S&P 100) from 2018 to 2022 was performed.

Regulation S-K serves as the central repository for non-financial statement disclosure requirements under U.S. federal securities laws. Items 101, 103, and 105 are critical components that detail business descriptions, legal proceedings, and risk factors, respectively. Modernization efforts are part of the SEC’s broader initiative to streamline and enhance its disclosure effectiveness. Two articles have directly addressed this modernization, and both have stated similar conclusions: the amendments have left the items in an “underwhelming state” with various solutions prescribed (Lee, 2020; Frey, 2021). Lee (2020) critiqued modernization efforts, arguing that they failed to address fundamental issues in disclosure practices, and Frey (2021) examined updates to Items 101 and 105, emphasizing the need for simplicity and clarity in corporate reporting.

The length and readability of corporate disclosures have been longstanding concerns among investors and other parties who rely on the information presented within the disclosures. Leuz and Schrand (2009) found that excessive disclosure lengths often hinders effectiveness. They examined the link between disclosure and the cost of capital for firms that experienced an exogenous capital shock after the Enron scandal. They found that firms responded to this shock by expanding the number of pages in their 10-K filings and providing additional interim disclosures, such as conference calls and 8-K filings. They also noted that this observation was more pronounced

for firms with greater financing needs and positive capital shock costs. Similarly, Li (2008) introduced the concept of "fog," where complex language obscures the clarity of disclosures. These issues persist despite efforts to simplify and streamline reporting requirements. Fog was employed as a readability proxy in Li (2008) and is commonly used in content analysis literature to assess reading levels.

Boilerplate language, characterized by generic and repetitive content, has been the focus of several studies. Dyer et al. (2017) highlighted the prevalence of boilerplate disclosures, emphasizing their negative impact on the informativeness of financial reports. Hope et al. (2016) further noted that disclosure specificity is crucial for investor decision-making, yet many firms continue to rely on standardized templates. Tetlock (2007) examined the tone of financial disclosures and found that the language used can significantly influence investor perceptions.

HYPOTHESIS DEVELOPMENT AND METHODOLOGY

The effectiveness of the changes was analyzed to determine whether the specific goals in the last two lines of the SEC's summary were met within the period after the change. This entailed performing content analysis on the relevant sections of 10-Ks for both two years before and two years after the changes to Regulation S-K. The aim was to improve disclosures for investors (three measures of tone), simplify compliance for registrants (word count), improve readability (fog), and discourage repetition (boilerplate disclosures).

This study provides a content analysis of the amended Items 101, 103, and 105 to examine key changes and their potential impacts on public company disclosures. The analysis evaluates how the modernized rules aim to enhance disclosure quality, reduce redundancy, and facilitate more principles-based disclosures tailored to each registrant's specific circumstances. In accordance with the stated SEC goals and the nature of each revision, three hypotheses were posited:

H₁: *Regulation S-K Item 101 revision decreases FOG, LENGTH, and impacts tone (NEGATIVE, POSITIVE, UNCERTAIN).*

H₂: *Regulation S-K Item 103 revision decreases FOG, BPLAVG and BPLMAX*

H₃: *Regulation S-K Item 105 revision decreases LENGTH, FOG, and impacts tone (NEGATIVE, POSITIVE, UNCERTAIN)*

The Gunning-Fog index (FOG) measures the reading level text as:

$FOG = (\text{words per sentence} + \text{percentage of complex words}) * 0.4$, where complex words are words with at least three syllables.

The Lingua:EN:Fathom Module in Perl was employed to measure both the FOG and LENGTH (Lang and Stice-Lawrence, 2015; Boritz et al., 2016; Bushee et al., 2018). This module analyzes English text in a string or text file and computes the FOG index by counting the number of words, sentences, syllables (Lingua:EN:Syllable), and blank and non-blank text. To measure LENGTH, the natural logarithmic transformation of the word count from Lingua:EN:Fathom was used to pare down large figures.

According to the Loughran-McDonald English Business Dictionary (March 2019 version) several common measures of tone exist in the literature. These include measures of negative, positive, and uncertain tones (Feldman et al. 2010; Henry, 2008; Kothari et al. 2009; Smith, 2017). We restrict the measures of tone to this dictionary to help identify tone usage in business writing and to avoid including proper nouns or words from foreign languages. The number of NEGATIVE, POSITIVE, and UNCERTAIN tones was measured using three accompanying word dictionaries from Loughran and McDonald (2017) to capture the tone of the disclosure. Word lists were converted into text files and then extracted into columns for import into LIWC for use as custom dictionaries. The negative and uncertain word dictionaries approximated the level of negative content discussed in each section of the 10-K. The positive dictionary captured any usage of tone, in addition to the content in the

negative and uncertain dictionaries. These dictionaries were used in the Linguistic Inquiry and Word Count software (LIWC) to analyze tone counts. LIWC reports the word counts for any given measure of tone in a dictionary for a given text file. Finally, the 10-K sections were converted into text files and passed through LIWC for analysis after the cleaning of non-recognizable characters.

Boilerplate disclosure is a common construct in content analysis literature (McMullin, 2016). The measure of boilerplate disclosure, BPLAVG, is unique in that it does not compare text to a “corpus,” or comprehensive text, to determine whether a disclosure is a boilerplate. Plagiarism software, WCopyfind, was used on the 10-Ks following prior literature in content analysis to compare texts (Cazier and Pfeiffer, 2015; McMullin, 2016; Beatty et al. 2019) and evaluate boilerplate disclosures. WCopyFind compares two documents side-by-side using certain textual parameters such as “shortest phrase to match” and “minimum % of matching words” to measure the aggregate percentage of plagiarism between the two documents. These results produce a matrix for each firm’s disclosure, in which each year of the four sample years has three unique results. The plagiarism percentages for each firm year were averaged across the four years, and this figure was assigned to BPLAVG, or the average boilerplate figure when comparing a single firm year to the other three firm years. BPLMAX is the maximum boilerplate disclosure within the matrix.

Given the hypotheses and explanations of the content analysis variables, the following three models were employed for testing:

$$SK_101 = \alpha + \beta_1(LENGTH_t) + \beta_2(FOG_t) + \beta_3(NEGATIVE_t) + \beta_4(POSITIVE_t) + \beta_5(UNCERTAIN_t) + \varepsilon \quad (1)$$

$$SK_103 = \alpha + \beta_1(LENGTH_t) + \beta_2(FOG_t) + \beta_3(BPLAVG_t) + \beta_4(BPLMAX_t) + \varepsilon \quad (2)$$

$$SK_105 = \alpha + \beta_1(LENGTH_t) + \beta_2(FOG_t) + \beta_3(NEGATIVE_t) + \beta_4(POSITIVE_t) + \beta_5(UNCERTAIN_t) + \varepsilon \quad (3)$$

where SK_101, SK_103, and SK_105 are dummy variables that denote the periods before and after the modernization of the SK Regulations. Each model analyzes the effect of change on the textual characteristics discussed in that section.

The S&P 100 was chosen as the sample for analysis. This was done for several reasons: (1) data are readily available via SEC EDGAR; (2) large accelerated filers have the most stringent disclosure requirements under Regulation S-K; (3) these firms tend to be timelier than smaller firms regarding the timing of their 10-K uploads to the database, which is relevant for the recency of the data up to 2022; and (4) while specific provisions of modernization targeted small firms, all provisions affected large firms. Among the S&P 100, 75 of the 100 firms were chosen for the final 4-year sample owing to missing data and industry exclusions (N=300). Additionally, owing to the effective date of modernization (November 9, 2020) and variance in fiscal year-ends, the sample period was either 2018-2021 or 2019-2022 to ensure that two years of data were available both before and after modernization.

RESULTS

Table 1 presents the descriptive statistics for the S&P 100 firms during the sample period 2018-2022. Each construct was divided into relevant sections in the 10-K: business, legal, and risk factor sections. Business and risk factor sections were found to be much longer on average than legal sections, with average word counts of 7,017 and 9,485 words for the business and risk factor sections, respectively, and only 210 words for the legal sections¹. The Fog was at a high school reading level for the median firm-year. The median of 17.73 can be interpreted as a high school senior reading level, as this figure directly translates to the year of instruction, while the risk factor section was found to be at a college reading level (median of 20.19). Boilerplate figures, BPLAVG and BPLMAX, were particularly high across the sample for the legal section, indicating that it was heavily plagiarized from year to year. Both the median and median boilerplate figures exceeded 75% of the repeated text from the 10-Ks within the sample period for each firm.

TABLE 1
DESCRIPTIVE STATISTICS (N=300)

Variable	Mean	Std Dev	Min	Median	Max
LENGTH – BUS (ln)	8.68	0.04	6.64	8.76	9.95
LENGTH – LEGAL (ln)	4.24	0.07	2.48	3.82	7.91
LENGTH – RISK (ln)	9.02	0.03	7.03	9.06	10.18
FOG – BUS	17.85	0.10	7.21	17.73	24.00
FOG – RISK	20.08	1.41	16.22	20.19	23.88
NEGATIVE – BUS	0.32	0.19	0.01	0.29	0.92
POSITIVE - BUS	1.95	0.04	0.78	1.90	4.10
UNCERTAIN - BUS	42.48	10.05	21.64	41.50	74.79
NEGATIVE – RISK	1.16	0.02	0.29	1.12	2.15
POSITIVE - RISK	1.75	0.36	0.85	1.72	2.96
UNCERTAIN – RISK	25.57	0.36	5.91	26.19	46.52
BPLAVG - LEGAL	0.77	0.02	0.00	0.91	1.00
BPLMAX - LEGAL	0.81	0.33	0.00	0.96	1.00

Tables 2, 3, and 4 present the results of the tests of the three models for H₁, H₂, and H₃. The test of H₁ in Table 2 indicates that the revision of Regulation S-K item 101 had no significant relationship with the FOG or any measure of tone. While a significant relationship with LENGTH existed, it increased after the effective date of the regulation, moving in the opposite direction of expectations and countering the SEC’s stated readability goals. This was due to the “may” clause within the Item 101 revision, which gives registrants the option of providing an update rather than a full discussion. Large accelerated filers are likely to view a full discussion as necessary for their investors.

TABLE 2
MULTIVARIATE RESULTS (N=300, SK_101)

Variable	Parameter Estimate	Standard Error	t Stat	P-value
Intercept	-2.9506	1.5685	-1.8812	0.0609
LENGTH – BUS	0.0949**	0.0458	2.0733	0.0390
FOG – BUS	0.0133	0.0173	0.7655	0.4446
Negative – BUS	2.0888	1.4336	1.4571	0.1462
Positive – BUS	-1.2769	1.2588	-1.0144	0.3112
Uncertainty - BUS	0.0993	0.0821	1.2096	0.2274

***, **, and * indicate significance at the .01, .05, and .10 levels, respectively.

TABLE 3
MULTIVARIATE RESULTS (N=300, SK_103)

Variable	Parameter Estimate	Standard Error	t Stat	P-value
Intercept	0.6434	0.2095	3.0713	0.0023
SPECIFICITY – LEGAL	-0.0001	0.0023	-0.0498	0.9603
BPLAVG – LEGAL	-0.2270	0.4443	-0.5110	0.6097
BPLMAX – LEGAL	0.1555	0.4440	0.3501	0.7265
LENGTH - LEGAL	-0.0219	0.0376	-0.5821	0.5609

***, **, and * indicate significance at the .01, .05, and .10 levels, respectively.

TABLE 4
MULTIVARIATE RESULTS (N=300, SK_105)

Variable	Parameter Estimate	Standard Error	t Stat	P-value
Intercept	-2.6082	0.6988	-3.7325	0.0002
LENGTH - RISK	0.1803***	0.0519	3.4740	0.0006
FOG - RISK	0.0345	0.0196	1.7656	0.0785
Negative - RISK	0.6568***	0.1333	4.9262	0.0000
Positive - RISK	0.0026	0.0026	0.9937	0.3212
Uncertainty - RISK	0.0009	0.0059	0.1542	0.8775

***, **, and * indicate significance at the .01, .05, and .10 levels respectively.

Table 3 presents the results of the effect of the Item 103 revision on the legal section of the 10-Ks, which allows registrants to hyperlink their legal disclosure if it appears elsewhere in the annual report. The SEC stated both in its summary goals and directly in the revision that the goal was to “avoid duplicative disclosure” (SEC, 2020). The test of H₃ revealed no significant change in boilerplate text (BPLAVG or BPLMAX) and no significant change after the effective date of the regulation for FOG or LENGTH. This indicates that registrants chose not to hyperlink their legal discussion and instead provided a full discussion of material contingencies and effects of litigation in both the opening legal section and in the MD&A.

Finally, the effects of Item 105 revisions on textual characteristics are presented in Table 4. The SEC has stated that the risk factor section should be paired down to two pages if it exceeds 15 pages (SEC, 2020). Significant relationships were found only for LENGTH and NEGATIVE, suggesting that the word count of the risk factor section actually increased for large accelerated filers, and the negative tone increased after the effective date of the regulation. The intent of the SEC was to decrease the length of this section; however, it had the opposite effect. Manual inspection of the documents revealed that many registrants paired their sections to slightly fewer than 15 pages so that they did not fall under this revision, once again countering the SEC’s stated readability goals.

CONCLUSION

The modernization of Regulation S-K Items 101, 103, and 105 did not fully realized its objectives among large accelerated filers. Although these amendments aimed to foster clarity, conciseness, and relevance, current practices indicate persistent issues of readability, length, and boilerplate content. Future regulatory efforts must address these shortcomings to enhance the effectiveness of corporate disclosures.

This study had a limited scope in that the sample included only large accelerated filers in the S&P 100, while the SEC had a much broader target with the modernization of Regulation S-K items, which included smaller reporting companies and non-accelerated and accelerated filers. The impact of modernization is likely to be more

meaningful for these filers and their investors as reduced disclosure eases the burden on firms and may increase investors' readability.

ENDNOTE

- ¹ Natural logarithms of word counts are included in table 1. The counts listed in the manuscript are not pared.

REFERENCES

- Albring, S.M., Robinson, D., & Robinson, M. (2014). Audit committee financial expertise, corporate governance, and the voluntary switch from auditor-provided to non-auditor-provided tax services. *Advances in Accounting*, 30(1), 81–96. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0882611014000031>
- Beatty, A., Cheng, L., & Zhang, H. (2019). Are risk factor disclosures still relevant? Evidence from market reactions to risk factor disclosures before and after the financial crisis. *Contemporary Accounting Research*, 36(2), 805–838.
- Boritz, J.E., Hayes, L., & Timoshenko, L.M. (2016). Determinants of the readability of SOX 404 reports. *Journal of Emerging Technologies in Accounting*, 13(2), 145–168.
- Bushee, B.J., Gow, I.D., & Taylor, D.J. (2018). Linguistic complexity in firm disclosures: Obfuscation or information? *Journal of Accounting Research*, 56(1), 85–121.
- Cazier, R.A., & Pfeiffer, R.J. (2015). Why are 10-K filings so long? *Accounting Horizons*, 30(1), 1–21.
- Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, 64(2–3), 221–245. Retrieved from <https://www.sciencedirect.com/science/article/abs/pii/S0165410117300282>
- Feldman, R., Govindaraj, J., Livnat, J., & Segal, B. (2010). Management's tone change, post earnings announcement drift, and accruals. *Review of Accounting Studies*, 15(4), 915–953.
- Frey, J. (2021). Striving for Simplicity: Updates to Regulation S-K Items 101 and 105. *Journal of Securities Regulation*, 45(2), 56–78.
- Henry, E. (2008). Are investors influenced by how earnings press releases are written? *Journal of Business Communication*, 45, 363–407.
- Hope, O.K., Hu, D., & Lu, H. (2016). The Benefits of Specific Risk-Factor Disclosures. *Review of Accounting Studies*, 21(4), 1005–1045. Retrieved from <https://link.springer.com/article/10.1007/s11142-016-9375-8>
- Jurafsky, D., & Martin, J.H. (2014). *Speech and language processing* (Vol. 3). London: Pearson.
- Kothari, S.P., Li, X., & Short, J.E. (2009). The effect of disclosures by management, analysts, and financial press on the equity cost of capital: A study using content analysis. *The Accounting Review*, 84, 1639–1670.
- Lee, Y.J. (2012). The effect of quarterly report readability on information efficiency of stock prices. *Contemporary Accounting Research*, 29(4), 1137–1170.
- Lee, J. (2020). Modernizing Regulation S-K: Ignoring the Elephant in the Room. *Securities Regulation Journal*, 48(3), 129–150.
- Leuz, C., & Schrand, C. (2009). Disclosure and the Cost of Capital: Evidence from Firms' Responses to the Enron Shock. *Journal of Accounting Research*, 47(2), 241–273. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-679X.2009.00325.x>
- Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2–3), 110–135.
- Li, F. (2008). Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics*, 45(2–3), 221–247. Retrieved from <https://www.sciencedirect.com/science/article/abs/pii/S0165410108000222>
- Li, F. (2010). The information content of forward-looking statements in corporate filings – A naïve Bayesian machine learning approach. *Journal of Accounting Research*, 48, 1049–1102.

- McMullin, J. (2016). *Can I borrow your footnotes? Footnote boilerplate's learning externality*.
- Mitchell, T.M. (2006). *The discipline of machine learning* (Vol. 3). Carnegie Mellon University, School of Computer Science, Machine Learning Department.
- Miller, B.P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, 85(6), 2107–2143.
- Peterson, K. (2012). Accounting complexity, misreporting, and the consequences of misreporting. *Review of Accounting Studies*, 17(1), 72–95.
- Smith, K. (2017). *Tell me more: A content analysis of expanded auditor reporting in the United Kingdom*. <http://dx.doi.org/10.2139/ssrn.2821399>
- Tetlock, P.C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance*, 62(3), 1139–1168. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2007.01232.x>
- U.S. Securities and Exchange Commission. (2020). *Modernization of Regulation S-K Items 101, 103, and 105*. Retrieved from <https://www.sec.gov/news/press-release/2020-192>
- SEC Release Nos. 33-10825; 34-89670; File No. S7-11-19. Retrieved from <https://www.sec.gov/rules/final/2020/33-10825.pdf>