READABILITY OF ANNUAL REPORTS ON THE VIENNA STOCK EXCHANGE: A TEST OF MANAGEMENT OBFUSCATION HYPOTHESIS

Stellner, B.

Bernhard Stellner / Mendel University in Brno, Faculty of Business and Economics, Department of Management, Zemědělská 1, 61300 Brno, Czech Republic. Email: bernhard.stellner@hotmail.com

Abstract

This research examines the relation between annual report readability and company performance in a German-speaking country, Austria. The incomplete revelation hypothesis, management obfuscation hypothesis and agency theory assume that firms with lower performance strategically use readability in their disclosures to obfuscate negative results. For investors, reading, analysing, and interpreting data becomes a costly affair; this weakens the negative effect of such data on a firm's reputation and share price. We use LIX and Flesch formulas to measure the readability of letters to the shareholders and/or interviews with the board in annual reports. The sample consists of 37 companies that are listed on the Prime Market of the Vienna Stock Exchange and their data from the year 2009 to 2020. Company performance is measured by the change in turnover, profit, and share price. The analysed sections mostly show high to very high levels of difficulty. During the observation period, readability levels do not change significantly. We find that the annual reports of firms with lower performance are not harder to read and, therefore, cannot confirm the management obfuscation hypothesis. A significant influence of change in profit/loss on readability is minutely observed. Possible reasons for this observation could be characteristics of the German language, statistical outliers, the long observation period, more professional investor relations offices, and changing communication channels between companies and stakeholders. The last point, changing communication channels, also puts the obfuscation hypothesis and its application to readability up for discussion again.

Implications for Central European audience: Our study shows that also Central European countries are confronted with low levels of readability in annual reports. Nevertheless, we cannot see a clear tendency towards obfuscation in corporate disclosures.

Keywords: agency theory; management obfuscation hypothesis; LIX; annual report readability; Vienna Stock Exchange

JEL Classification: M12, M51, M54

Introduction

Companies listed on a stock exchange have a strong desire to communicate well and in their interests with stakeholders. The primary means of written communication is the annual report which typically consists of a narrative and quantitative information (Baker & Kare, 1992).

Depending on the legislation, some parts follow a clear structure and content, while other parts, such as the letter to the shareholder or interview with the board members, maybe arranged freely or even left out. Nevertheless, both are well-established, non-audited instruments that offer the possibility to choose, emphasise and obfuscate reported information. Stakeholders seek to be informed on history, status, and future expectations after reading these sections as they also see it as a type of summary. Ultimately, (potential) shareholders will also make their investment decisions based on the contents of the disclosure (Courtis, 1995). The traditional efficient markets hypothesis (EMH) claims that share prices fully reflect all available information (Bloomfield, 2002; Fama, 1970). This also means data that is obfuscated or difficult to obtain from annual reports is always entirely considered and priced in. On the other hand, based on the incomplete revelation hypothesis (IRH) that acknowledges the cost of extracting information for trading decisions, the management obfuscation hypothesis claims that management has an incentive to obfuscate bad news to its stakeholders. Reading, analysing and interpreting data becomes a costly affair for investors and, therefore, weakens the negative effect of such data on a firm's reputation and share (Bloomfield, 2002; Courtis, 1998; Li, 2008). Adjusting the readability, a trigger for understandability, of the letter to the shareholders and/or interview with the board is a potential way to spread good and obfuscate bad news, results or outlooks (Baker & Kare, 1992; Courtis, 1995; Li, 2008; Rutherford, 2003; Thoms et al., 2020). If the management obfuscation hypothesis held true, lower performance (outlook) of a company would result in lower readability measures. Previous results on this hypothesis are mixed and unclear. There is a broad consensus that annual reports are generally difficult to read, yet the relation between performance measures and readability was observed to be significant only in some studies (Adelberg, 1979; Baker & Kare, 1992; Hrasky et al., 2009; Jones, 1988; Li, 2008; Subramanian et al., 1993; Thoms et al., 2020), while others find no evidence towards obfuscation (Courtis, 1986, 1995; Rutherford, 2003; Smith et al., 2006).

1 Theoretical background

1.1 Management obfuscation hypothesis

In the construct of a company, the cooperation of managers and owners (shareholders) with a separation of ownership and control is frequently modelled with the help of the principal-agent theory. Shareholders as principals aim to delegate management tasks to their agent, the manager. In this agency relationship, not all interests are aligned toward a common goal (Jensen & Meckling, 1976). Managers are assumed to act selfish, have short-term thinking, and avoid personal risk. Therefore, they take actions that primarily lead to benefits for themselves, but that could also countervail their principals' interests (Eisenhardt, 1989). Given their deeper involvement within the company, managers have an advantage with respect to information in terms of time and detail compared to shareholders (Jensen & Meckling, 1976).

Building on the information asymmetry between shareholder and manager, the management obfuscation hypothesis describes a setting within the principal-agent relation, where the manager underlines the good news or results and obfuscates the bad news or performance. In readability research, the obfuscation hypothesis typically assumes that good results are presented in a more readable manner (Courtis, 1998). Hence, the manager practices impression management and "obscures the intended message, or confuses, distracts or perplexes readers, leaving them bewildered or muddled" by using jargon, irrelevant information, complexity, and

length (Courtis, 2004). As a result, negative information becomes more difficult to process for investors and undesired consequences on reputation or share price are delayed or impeded (Bloomfield, 2002). Speaking in terms of the IRH, managers could attempt to control the processing costs of information according to their needs (Bloomfield, 2002; Li, 2008). Other scholars argue with the signalling theory: managers tend to highlight good performance by making it easily accessible and readable (Rutherford, 2003). We consider the management obfuscation hypothesis and signalling theory congruent in their practical effect on readability.

1.2 Readability

The readability of a text describes how difficult it is for a reader to understand the text (Oakland & Lane, 2004). It drives the realisation of the value of information in a text (Li, 2008). Although there is a lack of common ground on the question of whether readability equals understandability (Jones, 1988) or not (Smith & Taffler, 1992), we may say that readability at least triggers the degree to which a text is understood. Oakland and Lane (2004) consider text difficulty to be driven by both the reader (fluency, background knowledge, language, motivation, and engagement) and the text factors (syntax, vocabulary, idea density, cognitive load). In measuring and processing readability with conventional formulas, the background of the reader is usually not considered. There is a variety of different formulas and methods that predict the readability of a text with a single indicator by counting language variables (Subramanian et al., 1993). Typically, a combination of word and sentence length is processed. A popular example for English texts is the Flesch index, which considers the average number of syllables per word and the average number of words per sentence.

Flesch score = 206,835 − 84,6 x word length in syllables − 1,015 x sentence length (1)

A score close to 0 indicates very difficult texts, while 100 indicates a comic style level (Flesch, 1948).

For German texts, the LIX (Läsbarhetsindex or Lesbarkeitsindex) gained popularity. This score consists of the sum of average sentence length and the share of words with more than six letters.

LIX score = sum of words / number of sentences + number of long words / sum of words x 100 (2)

Lower results indicate better readability, with professional literature usually scoring above 60 (Lenhard & Lenhard, 2011). Table 1 shows typical LIX readability levels and text types for the respective scores.

Table 1 | LIX Readability Scale

LIX score	Readability level	Text type
over 60	very difficult	professional
50 to 60	difficult	nonfiction
40 to 50	medium	fiction
30 to 40	easy	youth
under 30	very easy	children

Source: (Lenhard & Lenhard, 2011)

The following exemplary sentences are taken from the 2016 annual report of the Frauenthal Holding AG 2016, where the rletter to the shareholders has a LIX score of 68.10: *“Dank einer*

guten Marktentwicklung im Bereich Automotive steigt das operative Konzernergebnis EBITDA1 um MEUR 1,5 auf MEUR 33,4, und dies trotz erheblicher einmaliger Aufwendungen im Handelskonzern in Zusammenhang mit der Umstellung der IT- und Logistiksysteme. Mit der Schaffung einer gemeinsamen Infrastruktur im Frauenthal Handelskonzern für Einkauf, Lager, Logistik und Administration können erhebliche Synergien realisiert werden. Der Gedanke der „one company" bei gleichzeitigem selbständigen Auftritt der Marken SHT, ÖAG und Kontinentale am Markt wird durch die Etablierung der Organisationseinheit Frauenthal Service, die den Vertriebsmarken alle Logistik- und Back-office Leistungen zur Verfügung stellt, ermöglicht."

In general, one should consider that "*No formula will guarantee that you write well. Nonsense written simply is still nonsense*." (Gunning, 1969). Readability formulas have numerous limitations such as considering jargon, concept difficulty and density, logic, coherence and any type of formatting or visual support for the reader. Nevertheless, the formulas convince through their automation capability, simplicity, and independence from the readers (Bayerlein & Davidson, 2011; Bruce et al., 1981; Crossley et al., 2019; Redish, 2000). We believe that for readability studies, the absolute result of one single text is of less interest than the influence of time or other factors on the score of a bigger sample of texts (Thoms et al., 2020). Readability formulas also form the primary tool of readability research of written company disclosures (Adelberg, 1979; Courtis, 2004; Jones, 1988; Li, 2008; Rutherford, 2003; Smith & Taffler, 1992; Subramanian et al., 1993; Thoms et al., 2020). Alternative procedures for assessing readability include usability tests with representative readers (Redish, 2000), asking readers to assess text where every nth word is deleted[1] (Smith & Taffler, 1992; Taylor, 1953) and other, more modern and multi-dimensional computational linguistics analysis tools such as the Coh-Metrix (Chang & Stone, 2019; Crossley et al., 2019). They all have in common that they are more costly to process. Machine learning could help create more precise readability formulas in the future (Crossley et al., 2017).

The United States Securities and Exchange Commission (SEC) tackled the topic of low readability by introducing its Plain English initiative with the aim of making corporate disclosures more user-friendly and readable. The Plain Writing Act of 2010 has put an additional overall emphasis on creating documents that the public can understand and use (Habib & Hasan, 2018). German-speaking countries still lack ideas to make corporate communications more readable. Initiatives toward better understandability rather stem from non-institutional sources such as scholars (Keller, 2007) or magazines (Palan, 2011).

1.3 Previous research on the Influence of Performance on Readability

Early studies on the readability of annual reports (elements) primarily focused on the description of scores and their evolvement over time. They noted readability scores that indicate high levels of difficulty, which hinders the ability of an average reader to fully understand the texts (Barnett & Leoffler, 1979; Pashalian & Crissy, 1950; Soper & Dolphin, 1964). While this notion prevailed in later studies, a second research stream considered the components of company performance, measured with a variety of different indicators such as earnings per share (Adelberg, 1979), ROA/ROI/ROE/ROCE (Baker & Kare, 1992; Courtis, 1986; Hossain & Siddiquee, 2008; Jones,

[1] *Cloze Procedure*

Volume **11** | Issue **5** | 2022
https://doi.org/10.18267/j.cebr.307

1988; Rutherford, 2003), profit indicators (Li, 2008; Smith & Taffler, 1992), etc. It seems that the disagreement on how investors can evaluate the performance of a company is also reflected in academia. In terms of readability measurement, the Flesch index seems to remain the medium of choice, albeit quite frequently complemented by other scores such as Fog and LIX (Courtis, 1995; Li, 2008; Smith & Taffler, 1992). So far, research has almost entirely focused on English disclosures. There is also common agreement amongst scholars that the letter to the shareholders and/or management discussion and analysis (MD&A) is the most suitable section of reports to be analysed. Other sections either follow restrictive templates or are re-used from year to year without changes. Analysing whole reports without extracting passages is less expedient, although more efficient. The relation between performance and readability has, thus, ever since resulted in rather ambiguous outcomes: while some authors observe a tendency towards obfuscation, resulting in a positive relationship between company performance and disclosure readability (Adelberg, 1979; Baker & Kare, 1992; Hossain & Siddiquee, 2008; Hrasky et al., 2009; Jones, 1988), others find no such relation (Courtis, 1986, 1995; Rutherford, 2003; Smith et al., 2006). We agree with Li, who criticises the small sample sizes of well below 100 reports of studies from the 1970s to the 1990s. In 2008, Li, using a sample size of 55,719 firm years, found that lower earnings and earnings persistence results in lower readability. Hassan et al. (2019) found similar results in the case of Qatari listed firms, where annual report readability correlates positively with profitability, although readability was interpreted as an independent variable in this study. Recent Iranian (Hesarzadeh et al., 2020) and Brazilian (de Souza et al., 2019) studies led to similar results and confirmed the management obfuscation hypothesis. A more recent study from Germany analysed 30 DAX companies and found negatively connotated passages in the chairmen's addresses to be less readable. With the help of tone analysis, the authors raise the notion that this does not happen by intentional obfuscation but is rather a side product, as companies need to explain circumstances. Nevertheless, the existence of a direct relation between company and performance was not examined (Thoms et al., 2020).

2 Research gap, objectives and methodology

2.1 Research gap and objectives

We find past results unconvincing in their conclusiveness on the acceptance of the management obfuscation hypothesis. Most of the previous studies only focus on limited time spans. However, we aim to cover a time range of more than a decade. Moreover, non-English disclosures are virtually unexplored with regard to the readability of annual reports and their drivers (Hassan et al., 2019). Our paper aims to test the management obfuscation hypothesis in a German-speaking country, Austria, and gain a general understanding of the readability standards of company disclosures. We specifically focus on the firms in the Prime Market of the Vienna Stock Exchange (ATX prime).

2.2 Hypotheses

As discussed in subchapter 1.3, previous studies have observed high levels of textual difficulty in narrative company disclosures in the English language, irrespective of obfuscation tendencies. Attempting to compare general readability between English and German seems delicate, albeit German is said to be the more complex language of the two. High levels of reading difficulty were recently observed in the German market (Thoms et al., 2020). Therefore, we see no

argument supporting the expectation that narrative company disclosures in the Austrian market should be more readable.

Hypothesis 1: *ATX prime companies publish letters to the shareholders that are difficult or very difficult to read.*

Motivated by the tendency toward the validity of the management obfuscation hypothesis in prior research (Hassan et al., 2019; Li, 2008; Hrasky et al., 2009), this research tests the obfuscation hypothesis in a non-English environment. As mentioned before, a series of authors have found evidence for a positive relationship between company performance and readability and vice versa. We find no evidence to support the notion that this relation would not be applicable to the Austrian market.

Hypothesis 2: *ATX prime companies with higher performance publish letters to the shareholders that are easier to read.*

We also attempt to analyse the development of readability over time. With the absence of initiatives such as the Plain Writing Act or a public discussion on readability, understandability, and accessibility of company disclosures in German-speaking countries, we see no reason why the readability of narrative company disclosures would have improved. The same reasoning should also apply to the length of letters to the shareholder and interviews with the management board.

Hypothesis 3*: The readability of narrative company disclosures of ATX prime companies has not changed significantly between 2009 and 2020.*

Hypothesis 4: The length of the examined sections of narrative company disclosures of ATX prime companies has not changed significantly between 2009 and 2020.

2.3 Sample

The present study uses a sample of 37 Austrian companies listed on the Vienna Stock Exchange. We selected those companies that were part of the ATX Prime Market in September 2021 and retrieved their annual data from the year 2009 to 2020, which resulted in a theoretical number of 444 observations. Several reasons (listing, availability, change of reporting period and others) resulted in a total of 417 usable firm-year observations of readability. Altogether around 11,000 sentences and 212,000 words were analysed. Firstly, we focused on the letter to the shareholders. Interviews with the management board were used alternatively as we observed that companies tend to use either or. Both sources were only used if sources stemmed from the management board and not the advisory board or other experts. Unlike other scholars (Courtis, 1995), we did not process a randomly selected 100-word passage but used the entire text. As the Vienna Stock Exchange does not provide a central register with company disclosures or key performance indicators, the annual reports were downloaded from the companies' homepages or other archives manually. Older publications were requested directly from the companies' investor relations offices. The PDF files were screened individually, unlocked in some circumstances, and relevant passages were manually extracted into word files. Salutations, images, diagrams, signatures, and other additional elements were left out. As headlines can ease readability, they were not excluded. This procedure ensures that only those passages that are relevant for a standard reader are extracted. We are convinced that these steps increase the

Volume **11** | Issue **5** | 2022
https://doi.org/10.18267/j.cebr.307

quality of the textual data included in the study, even though direct processing of the complete PDF files would be more efficient and effective at first glance.

Company data were primarily collected with the business research database Orbis. Data extraction was carried out on October 14th, 2021. Orbis' limitation – data time lag of up to three years on performance indicators – was compensated for by manually screening annual reports and other online sources such as www.finanzen.net or www.onvista.de for missing data.

As a general principle, business years that did not end on December 31st were considered the year with the most affected months.

2.4 Variables

The main indicator for readability used in this study was the LIX (Lesbarkeitsindex) score as we dealt with German disclosures. We calculated the LIX and text length score using the online tool of www.psychometrica.de (Lenhard & Lenhard, 2011). Moreover, we calculated a Flesch score for all text excerpts with the online tool www.leichtlesbar.ch. Both online tools work efficiently and have scientific backgrounds. Due to the high correlation of LIX and Flesch scores, it was easy to visually detect data entry errors of score pairs.

Multiple performance-related indicators were gathered. We collected year-end share prices in EUR from the end of 2008 to 2020 and calculated a share price change factor (share price t divided by share price t-1) to account for market performance in period t. A similar logic was applied for turnover in EUR (turnover t divided by turnover t-1), whereby the lack of turnover data for the year 2009 restricted the application of this logic for that year. Thus, the first turnover change factor is available for 2010. For for-profit/loss, we applied a different method. As negative values would distort a simple division, we applied a binary system instead (1 = profit t is bigger than profit t-1, and vice versa). A decrease in annual loss would mean good performance in this respect.

We considered the company sizes measured by the number of employees as a control variable. Based on the initial thoughts of other scholars (Ajina et al., 2016; Hassan et al., 2019) and the broader concept of agency theory (Jensen & Meckling, 1976), we considered shareholder influence as the control variable. Companies with higher independence from their shareholders are assumed to write in a more readable manner (Ajina et al., 2016; Oliveira et al., 2006). The Orbis database offers an indicator by the Bureau van Dijk, the BvD Independence Indicator, which attempts to measure whether a company is independent of its owners based on the number of shareholders and their percentual holding (Horobet et al., 2019). We transferred the original format (ratings from A+ for high independence to D for low independence) into a numerical one ranging from 0.5 to 4. Time was considered as a dummy variable ranging from 1 (2009) to 12 (2020) in the regression model.

2.5 Methods

Our core data of readability scores and performance (change) indicators is available for 37 companies and up to 12 business years. We conducted both longitudinal and cross-sectional analyses using the IBM SPSS Statistics 28. For Hypothesis 1, we used simple descriptive statistics such as mean and standard deviation to interpret the absolute readability scores. Hypothesis 2 was tested with a linear regression model where readability was the dependent

variable, company performance was the independent variable, and company size and shareholder independence were the control variables. Data pairs of the same companies were used for all available years. For further analysis, we also calculated medians and applied Pearson bivariate correlation matrices. Hypothesis 3 and 4 used a series of multiple t-tests to show the differences in readability and disclosure length over time.

3 Results

The 37 companies investigated in this study had an average of 13,400 employees, an annual turnover of 3,359 MEUR and an annual profit of 218 MEUR during the period under review (see Table 2). The LIX scores (M = 58.14) and Flesch scores (M = 16.98) had a high negative correlation ($r = (413) = -.82$, $p < .001$). Due to this high congruence, we carried out the remaining analyses with LIX only.

Table 2 | Descriptive Statistics of the Sample

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Number of employees	37	40	72,904	13,403	16,591
Turnover	422	2.6	43,040	3,359	5,419
Profit/Loss	425	-3,403	3,067	218	500
LIX 2009	31	50.00	75.90	60.08	4.58
LIX 2010	35	50.90	72.50	59.38	3.88
LIX 2011	35	42.90	70.20	57.93	4.72
LIX 2012	36	45.70	77.70	58.25	5.73
LIX 2013	36	44.20	72.60	58.21	5.64
LIX 2014	36	47.60	64.70	58.27	3.87
LIX 2015	37	49.20	73.00	58.30	4.94
LIX 2016	36	47.50	71.30	57.65	5.68
LIX 2017	37	48.20	72.10	57.99	5.30
LIX 2018	37	40.80	70.00	56.67	6.02
LIX 2019	38	47.60	71.10	58.16	4.93
LIX 2020	36	46.10	68.40	57.61	5.86

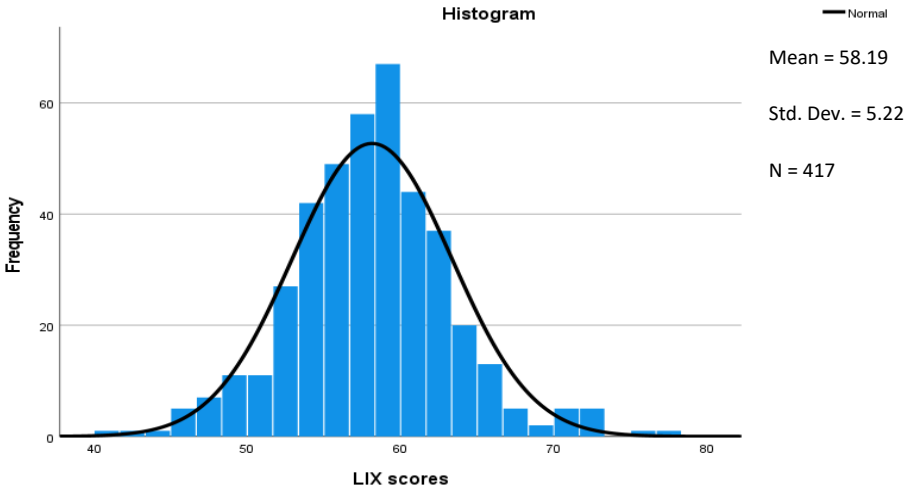
Note: Number of employees is the last value available, Turnover and Profit/Loss in MEUR

Source: own calculation

As displayed in Figure 1 below, the distribution of readability scores across years and companies almost follows a standard distribution (M = 58.19, kurtosis = 1.11, skewness = 0.21). Although some companies provided disclosures with readability levels of around 40, almost indicating an easy-to-read text, extreme outliers had LIX scores of above 75. Texts of this kind are very hard to read, even for professionals. Only 6.2% of the texts examined scored below 50, the threshold where texts are considered to have medium readability. Conversely, 93.8% of the texts were at least difficult to read. Amongst the total, 32.4% of the disclosures had LIX scores above 60, meaning that they are very difficult to read. Even though we identified some positive, readable examples within the sample, most of the disclosures had a high level of difficulty in terms of readability, which hindered a fluent comprehension of their contents. Even though we are aware that the difficulty thresholds are deliberately set and thus, biased, the present results are clear enough to show the levels of difficulty readers are confronted with. Hypothesis 1 is accepted.

Volume **11** | Issue **5** | 2022
https://doi.org/10.18267/j.cebr.307

Figure 1 | Histogram of LIX Scores



Source: own calculation

Table 3 reports the results of the regression model ($F(6, 364) = 1.49$, $p = .18$) where (LIX) readability was our dependent variable. We applied the pairwise deletion approach within the linear regression model of SPSS. The explanatory power of the model ($R^2 = .024$) is at a low level, which is not surprising per se since readability is not expected to largely rely on company performance. Company performance was tested by the annual change in share price, turnover and profit/loss (independent variable). Overall, we found no indication of a significant influence of performance on readability. Only the negative influence of change in profit/loss on LIX (which in turn led to a positive influence on readability) was found to be significant at 8%. Still, we regard this as insufficient to serve as a confirmation of the management obfuscation hypothesis. Similarly, the control variables number of employees, independence from shareholders and time did not show any statistically significant relation with readability. Further, we repeated the procedure and included the absolute, inflation-unadjusted numbers of turnover and profit/loss as performance indicators. We still could not find any statistically significant influence of company performance on readability scores.

Table 3 | Regression Model

Model	*Coefficients*[a]	Unstandardised Coefficients		Standardised Coefficients			
		B	SE	Beta	t	p	R^2
(Constant)		59.35	1.24		47.61	<.001	.024
Change Share Price		-.38	.60	-.03	-.63	.53	
Change Turnover		-.24	.38	-.03	-.64	.52	
Change Profit/Loss		-1.01	.60	-.09	-1.77	.08	
BvD Independence		.22	.21	.06	1.05	.30	
Number Employees		.00	.00	.09	1.59	.11	
Time		-.13	.09	-.08	-1.43	.16	

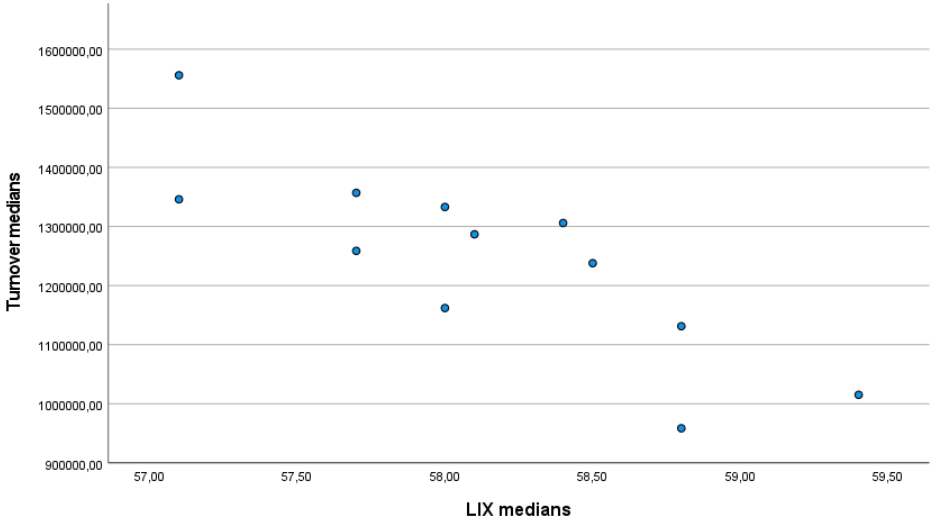
Note: N = 364

a. Dependent Variable: LIX Readability Score

Source: own calculation

To gain a more profound understanding of the relations between readability and performance over time, we calculated Pearson correlations for the median absolute values of turnover and LIX per year to diminish the effects of outliers. Data points were strongly reduced as a result. The relation was highly significant ($r(12) = -.84$, $p < .001$) – see Figure 2. Median profit/loss showed a similar relation with median LIX ($r(12) = -.77$, $p = .004$). When adjusting for inflation at an estimated 2% rate per year, the relations still were significant at the 3% and 0.4% levels, respectively. These analyses would back a positive relationship between performance and readability.

Figure 2 | Scatter Plot Turnover and LIX medians 2009-2020



Source: own calculation

For further analysis, we set up a correlation matrix with annual LIX scores and annual Profit/Loss change. The full results can be found in Table 4. Indeed, some pairs showed significant relations.

https://doi.org/10.18267/j.cebr.307

Nevertheless, if affected years were more than one period apart, we considered the relation a coincidence. Especially the year 2018 showed a significant, negative correlation between LIX and Change profit/loss ($r(37) = - .45$, $p = .006$). This relation indicates that readability increases with profit in this individual year.

Table 4 | LIX and Change Profit/Loss Correlations per Year

	P/L Change 2010	P/L Change 2011	P/L Change 2012	P/L Change 2013	P/L Change 2014	P/L Change 2015	P/L Change 2016	P/L Change 2017	P/L Change 2018	P/L Change 2019	P/L Change 2020
LIX 2009	-,215	-,297	,345	-,219	-,040	,148	-,115	,136	-,202	**-,439**	-,055
	,273	,124	,067	,253	,836	,436	,546	,472	,285	**,015***	,771
	28	28	29	29	29	30	30	30	30	**30**	30
LIX 2010	,117	-,177	,256	-,159	-,061	,192	-,207	,229	**-,384**	-,285	,093
	,524	,333	,150	,377	,737	,277	,240	,193	**,025***	,102	,603
	32	32	33	33	33	34	34	34	**34**	34	34
LIX 2011	,090	,154	-,017	-,150	-,114	-,065	,076	-,039	-,407	-,043	,266
	,622	,400	,924	,406	,528	,713	,670	,827	,017	,810	,128
	32	32	33	33	33	34	34	34	34	34	34
LIX 2012	,026	-,124	-,042	-,197	-,190	,047	-,230	-,157	-,236	,222	**,402**
	,889	,500	,815	,272	,289	,790	,190	,375	,179	,207	**,019***
	32	32	33	33	33	34	34	34	34	34	**34**
LIX 2013	-,041	-,084	,065	-,305	-,215	-,004	-,209	,085	-,300	,054	**,484**
	,819	,644	,713	,079	,221	,983	,228	,628	,080	,759	**,003****
	33	33	34	34	34	35	35	35	35	35	**35**
LIX 2014	,164	,089	,010	-,229	-,275	-,142	-,234	,194	-,190	,167	,025
	,369	,628	,957	,193	,115	,415	,182	,272	,275	,339	,888
	32	32	34	34	34	35	34	34	35	35	35
LIX 2015	,082	,040	,110	-,039	-,138	-,221	-,202	,246	-,007	-,180	,115
	,650	,825	,530	,824	,429	,196	,245	,155	,966	,295	,503
	33	33	35	35	35	36	35	35	36	36	36
LIX 2016	-,158	-,091	-,208	**-,382**	-,208	,160	-,168	-,064	**-,441**	,127	**,437**
	,388	,620	,238	**,026***	,238	,359	,343	,718	**,008****	,466	**,009****
	32	32	34	**34**	34	35	34	34	**35**	35	**35**
LIX 2017	-,069	-,077	,158	**-,378**	-,135	,216	-,249	,148	**-,357**	-,229	**,337**
	,706	,677	,373	**,027***	,445	,212	,156	,395	**,033***	,179	**,044***
	32	32	34	**34**	34	35	34	35	**36**	36	**36**
LIX 2018	,174	-,043	,121	-,286	-,055	,158	-,170	,025	**-,451**	-,260	,305
	,342	,816	,494	,102	,757	,365	,336	,889	**,006****	,126	,070
	32	32	34	34	34	35	34	35	**36**	36	36
LIX 2019	,212	-,024	,233	-,153	-,073	,117	-,212	,062	**-,360**	-,093	,116
	,237	,894	,177	,379	,676	,497	,222	,722	**,029***	,582	,494
	33	33	35	35	35	36	35	36	**37**	37	37
LIX 2020	-,017	-,150	,256	-,100	-,030	,299	-,102	-,151	-,196	-,004	,062
	,926	,419	,150	,581	,868	,086	,572	,395	,258	,981	,724
	31	31	33	33	33	34	33	34	35	35	35

Note: P/L = profit/loss
Source: own calculation

Overall, our analysis has brought mixed results on the relation between performance and readability. Although we did find evidence for obfuscation, we were unable to fully reject the null hypothesis. This contradicts most prior findings. Hypothesis 2 is rejected.

To test Hypothesis 3, we calculated confidence intervals at the 95% level of the LIX readability scores per year with t-tests (see Table 5). Subsequently, we analysed the intersections of the intervals. As all years overlap with each other, we reason that during the observed time span, no significant changes of readability in either direction have occurred. Hypothesis 3 is, thus, accepted.

Table 5 | T-Tests for LIX Readability Scores

			95% Confidence Interval	
Variable	**N**	**Mean**	**Lower**	**Upper**
LIX 2009	30	60.11	58.37	61.85
LIX 2010	34	59.39	58.02	60.77
LIX 2011	34	57.92	56.25	59.59
LIX 2012	34	58.26	56.23	60.28
LIX 2013	35	58.23	56.26	60.19
LIX 2014	35	58.27	56.92	59.62
LIX 2015	36	58.31	56.61	60.01
LIX 2016	35	57.65	55.67	59.63
LIX 2017	36	58.02	56.20	59.83
LIX 2018	36	56.67	54.60	58.73
LIX 2019	37	58.16	56.50	59.83
LIX 2020	35	57.58	55.54	59.63

Note: N total = 417

Source: own calculation

We tested Hypothesis 4 in the same way as Hypothesis 3. Even though sections were rather short in 2018 and longer in 2020, confidence intervals across all years overlap (see Table 6). This indicates that no significant changes in textual length in the period from 2009 to 2020 have occurred. Hypothesis 4 is, thus, accepted.

Volume **11** | Issue **5** | 2022
https://doi.org/10.18267/j.cebr.307

Table 6 | T-Tests for Length

			95% Confidence Interval	
Variable	**N**	**Mean**	**Lower**	**Upper**
Length 2009	30	929.26	760.16	1,098.37
Length 2010	34	919.79	773.36	1,066.22
Length 2011	34	927.29	767.46	1,087.13
Length 2012	34	1,029.91	823.93	1,235.89
Length 2013	35	925.91	738.58	1,113.25
Length 2014	35	852.60	700.29	1,004.91
Length 2015	36	857.11	698.02	1,016.21
Length 2016	35	784.05	628.54	939.57
Length 2017	36	865.22	679.29	1,051,15
Length 2018	36	831.44	657.41	1,005.48
Length 2019	37	952.37	737.94	1,166.82
Length 2020	35	1,041.22	826.32	1,256.14

Note: N total = 417, Length in words

Source: own calculation

4 Discussion, limitations and outlook

4.1 Discussion

Our study does not support the management obfuscation hypothesis tested with performance and readability. This contradicts most of earlier research, particularly considering the variety of different performance indicators that we used in this study. We suppose several possible explanations for that. First, our sample was drawn from German annual reports. This different linguistic setting could make the readability variation less viable (Thoms et al., 2020) and obfuscation attempts less effective. Second, as we found a clear correlation between median values of performance indicators (turnover and profit/loss) and median LIX readability, we assume that outliers in both variables strongly distort potential obfuscation tendencies. Third, we have drawn a sample over a period of twelve years and identified significant obfuscation tendencies in only one period. Past research with limited observation years could have been biased by such single yet coincidental effects. Other scholars could have refrained from publishing the results as they may not have found significant relations. Fourth, companies nowadays have established professional investor relations offices that are responsible for structuring and editing corporate disclosures (Brown et al., 2019). As the management obfuscation hypothesis is not a new concept, these departments could already be avoiding this rather obvious measure of obfuscation. Fifth, we must recognise that stakeholders nowadays obtain information through a variety of text and audio-visual channels and do not solely rely on a single annual report. Company reports in the sample could reflect this in their communication strategies. Our finding that shareholder structure (BvD Independence Indicator) has no effect on readability is a further indication that companies could have emancipated from obfuscating via narrative disclosures.

In line with previous studies, we demonstrate that Austrian stock listed companies also unnecessarily use difficult language in their corporate disclosures. The LIX scores for the present sample (M = 58.19) were slightly higher than those of Thoms et al. (2020) (M = 55.10), who analysed annual reports from Germany published in 2014. Letters to the shareholders and interviews with the board serve as the initial address to the stakeholders, providing them with a short briefing as well as a platform for the managers to present themselves. It is, therefore, surprising that they make so little use of remaining in positive, understandable memory.

Considering the above, we deem it necessary to discuss the necessity of regulations to make disclosures on the Vienna Stock Exchange, and the German-speaking world in general, more reader-friendly. A plain German initiative should enable investors and all other stakeholders of a company to understand disclosures and make informed decisions (Li, 2008).

4.2 Limitations

A general limitation of our study is the use of readability formulas, which provide no guarantee for correctly grasping understandability. They always come with a bias and are only an approximation of how readers perceive a text (Jones & Smith, 2014). Due to the limited number of listed companies in Austria, our sample size is limited in terms of separate entities. In terms of time, we used a decently lengthy period.

Although our research approach follows academic conventions, it is worth questioning whether analysing financial statements with readability formulas is a contemporary means of testing for obfuscation.

In terms of variables, we consider a change in share price the most problematic as it does not encapsulate the full yield of shareholders: dividend pay-outs and sale of subscription rights are not reflected properly. Moreover, measuring company size with the number of its employees is debatable, especially if companies of different industries and business models are compared.

Most of all, our rejection of the obfuscation hypothesis needs further investigation and should not be generalised to other markets.

4.3 Outlook

Analysing the readability of annual reports has been a decade long tradition. Yet, it probably does not reflect the increase in pace and innovations that managers, companies, and investors are confronted with. Almost live communication via social media postings, online Q&A, or video messages is becoming increasingly important – especially during times such as the pandemic. Academics need to screen these communication channels as well and put the management obfuscation hypothesis up for discussion again.

Future research should also pay attention to the major disadvantage of readability formulas – the misuse of punctuation. Writers could deliberately integrate punctuations either visibly (in numbers) or otherwise to outsmart readability analysis tools. In general, we regard this as an issue of English text due to the placement of decimals in numbers, which could appear more frequently if one wants to put good results in the spotlight. A potential improvement of readability formulas could be realised with the help of machine learning techniques (Crossley et al., 2017).

Volume **11** | Issue **5** | 2022
https://doi.org/10.18267/j.cebr.307

Scholars should further investigate the applicability of the management obfuscation hypothesis in German-speaking countries. We encourage other researchers to verify our present work using the same or similar samples. Nevertheless, we also propose to test the obfuscation thesis with different approaches in the future. On the one hand, the management obfuscation hypothesis could be tested with a stronger focus on what is written rather than with which level of difficulty. The Linguistic Inquiry and Word Count (LIWC) is a proper instrument for that and can be used to complement readability formulas (Markowitz & Hancock, 2016; Thoms et al., 2020). On the other hand, with the help of qualitative interviews with (former) board members or investor relations offices, one could get direct access to those who are claimed to have motivation towards obfuscation. Full confidentiality provided, the research could approach the topic from a different direction and evolve from a pure desk analysis of written reports.

From a practical point of view, companies should be aware that investors are frequently inundated with information. Therefore, annual reports should be established and lived as a high-quality, content-rich and easy-to-read source of information.

References

Adelberg, A. H. (1979). Narrative Disclosures Contained in Financial Reports: Means of Communication or Manipulation? A*ccounting and Business Research*, 9(35), 179–190. https://doi.org/10.1080/00014788.1979.9729157.

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. https://doi.org/10.1016/j.ribaf.2016.07.021.

Baker, E. H., & Kare, D. D. (1992). Relationship Between Annual Report Readability and Corporate Financial Performance. M*anagement Research News*, 15(1), 1–4. https://doi.org/10.1108/eb028188.

Barnett, A., & Leoffler, K. (1979). Readability of Accounting and Auditing Messages. *Journal of Business Communication*, 16(3), 49–59. https://doi.org/10.1177/002194367901600305.

Bayerlein, L., & Davidson, P. (2011). The Influence of Connotation on Readability and Obfuscation in Australian Chairman Addresses. *Managerial Auditing Journal,* 27(2), 175–198. https://doi.org/10.1108/02686901211189853.

Bloomfield, R. J. (2002). The "Incomplete Revelation Hypothesis" and Financial Reporting. A*ccounting Horizons*, 16(3), 233–243. https://doi.org/10.2308/acch.2002.16.3.233.

Brown, L. D., Call, A. C., Clement, M. B., & Sharp, N. Y. (2019). Managing the Narrative: Investor Relations Officers and Corporate Disclosure. *Journal of Accounting and Economics*, 67(1), 58–79. https://doi.org/10.1016/j.jacceco.2018.08.014.

Bruce, B., Rubin, A., & Starr, K. (1981). Why Readability Formulas Fail. *IEEE Transactions on Professional Communication*, 24(1), 50–52. https://doi.org/10.1109/tpc.1981.6447826.

Chang, Y.T., & Stone, D. N. (2019). Why Does Decomposed Audit Proposal Readability Differ by Audit Firm Size? A Coh-Metrix approach. *Managerial Auditing Journal*, 34(8), 895–923. https://doi.org/10.1108/MAJ-02-2018-1789.

Courtis, J. K. (1986). An Investigation into Annual Report Readability and Corporate Risk-Return Relationships. *Accounting and Business Research*, 16(64), 285–294. https://doi.org/10.1080/00014788.1986.9729329.

Courtis, J. K. (1995). Readability of Annual Reports: Western Versus Asian Evidence. *Accounting, Auditing & Accountability Journal,* 8(2), 4–17. https://doi.org/10.1108/09513579510086795.

Courtis, J. K. (1998). Annual Report Readability Variability: Tests of the Obfuscation Hypothesis. *Accounting, Auditing & Accountability Journal,* 11(4), 459–472. https://doi.org/10.1108/09513579810231457.

Courtis, J. K. (2004). Corporate Report Obfuscation: Artefact or Phenomenon? *The British Accounting Review,* 36(3), 291–312. https://doi.org/10.1016/j.bar.2004.03.005.

Crossley, S. A., Skalicky, S., & Dascalu, M. (2019). Moving Beyond Classic Readability Formulas: New Methods and New Models. J*ournal of Research in Reading,* 42(3-4), 541–561. https://doi.org/10.1111/1467-9817.12283.

Crossley, S. A., Skalicky, S., Dascalu, M., McNamara, D. S., & Kyle, K. (2017). Predicting Text Comprehension, Processing, and Familiarity in Adult Readers: New Approaches to Readability Formulas. *Discourse Processes*, 54(5-6), 340–359. https://doi.org/10.1080/0163853X.2017.1296264.

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *Academy of Management Review*, 14(1), 57–74. https://doi.org/10.5465/amr.1989.4279003.

Fama, E. F. (1970). Session Topic: Stock Market Price Behavior. *The Journal of Finance*, 25(2), 383–417.

Flesch, R. (1948). A New Readability Yardstick. *The Journal of Applied Psychology*, 32(3), 221–233. https://doi.org/10.1037/h0057532.

Gunning, R. (1969). The Fog Index After Twenty Years. J*ournal of Business Communication*, 6(2), 3–13. https://doi.org/10.1177/002194366900600202.

Habib, A., & Hasan, M. M. (2018). Business Strategies and Annual Report Readability. *Accounting & Finance,* 60(3), 2513–2547. https://doi.org/10.1111/acfi.12380.

Hassan, M. K., Abu Abbas, B., & Garas, S. N. (2019). Readability, Governance and Performance: A Test of the Obfuscation Hypothesis in Qatari Listed Firms. Corporate Governance. *The International Journal of Business in Society*, 19(2), 270–298. https://doi.org/10.1108/CG-05-2018-0182.

Hesarzadeh, R., Bazrafshan, A., & Rajabalizadeh, J. (2020). Financial Reporting Readability: Managerial Choices Versus Firm Fundamentals. *Spanish Journal of Finance and Accounting / Revista Española De Financiación Y Contabilidad*, 49(4), 452–482. https://doi.org/10.1080/02102412.2019.1668219.

Horobet, A., Belascu, L., Curea, Ș., & Pentescu, A. (2019). Ownership Concentration and Performance Recovery Patterns in the European Union. *Sustainability*, 11(4), 953. https://doi.org/10.3390/su11040953.

Hossain, F. F., & Siddiquee, M. (2008). Readability of Management Reviews in the Annual Reports of Listed Companies of Bangladesh. https://doi.org/10.2139/ssrn.1079695.

Hrasky, S., Mason, C. J., & Wills, D. I. (2009). The Textual Complexity of Annual Report Narratives: A Comparison of High- and Low-Performance Companies. *New Zealand Journal of Applied Business Research*, 7(2), 31-45-45.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the Firm: Managerial Behaviour, Agency Costs and Ownership Structure. *Journal of Financial Economics*, 3(4), 305–360. https://doi.org/10.1016/0304-405X(76)90026-X.

Jones, M. J. (1988). A Longitudinal Study of the Readability of the Chairman's Narratives in the Corporate Reports of a UK Company. *Accounting and Business Research*, 18(72), 297–305. https://doi.org/10.1080/00014788.1988.9729377.

Jones, M., & Smith, M. (2014). Traditional and Alternative Methods of Measuring the Understandability of Accounting Narratives. *Accounting, Auditing & Accountability Journal*, 27(1), 183–208. https://doi.org/10.1108/AAAJ-04-2013-1314.

Keller, R. (2006). Der Geschäftsbericht. Wiesbaden: Springer Fachmedien. https://doi.org/10.1007/978-3-8349-9049-5

Lenhard, W., & Lenhard, A. (2011). Berechnung Des Lesbarkeitsindex LIX Nach Björnson. [Calculation of LIX Readability Score According to Björnson]. https://doi.org/10.13140/RG.2.1.1512.3447.

Li, F. (2008). Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics*, 45(2-3), 221–247. https://doi.org/10.1016/j.jacceco.2008.02.003.

Markowitz, D. M., & Hancock, J. T. (2016). Linguistic obfuscation in fraudulent science. *Journal of Language and Social Psychology*, 35(4), 435-445. https://doi.org/10.1177/0261927X15614605.

Oakland, T., & Lane, H. B. (2004). Language, Reading, and Readability Formulas: Implications for Developing and Adapting Tests. *International Journal of Testing*, 4(3), 239–252. https://doi.org/10.1207/s15327574ijt0403_3.

Oliveira, L., Lima Rodrigues, L., & Craig, R. (2006). Firm-Specific Determinants of Intangibles Reporting: Evidence From the Portuguese Stock Market. *Journal of Human Resource Costing & Accounting*, 10(1), 11–33. https://doi.org/10.1108/14013380610672657.

Palan, D. (2011, December 8). Die besten Geschäftsberichte: Mehr Schein als Sein. [The best Business Reports. More Appearance Than Substance]. *Manager Magazin*. https://www.manager-magazin.de/magazin/artikel/a-802303.html.

Pashalian, S., & Crissy, W. J. E. (1950). How Readable Are Corporate Annual Reports? *The Journal of Applied Psychology*, 34(4), 244–248. https://doi.org/10.1037/h0061918.

Redish, J. (2000). Readability Formulas Have Even More Limitations Than Klare Discusses. *ACM Journal of Computer Documentation*, 24(3), 132–137. https://doi.org/10.1145/344599.344637.

Rutherford, B. A. (2003). Obfuscation, Textual Complexity and the Role of Regulated Narrative Accounting Disclosure in Corporate Governance. *Journal of Management and Governanc*e, 7(2), 187–210. https://doi.org/10.1023/A:1023647615279.

Smith, M., Jamil, A., Chik Johari, Y., & Ahmar Ahmad, S. (2006). The Chairman's Statement in Malaysian Companies. *Asian Review of Accounting*, 14(1/2), 49–65. https://doi.org/10.1108/13217340610729464.

Smith, M., & Taffler, R. (1992). Readability and Understandability: Different Measures of the Textual Complexity of Accounting Narrative. *Accounting, Auditing & Accountability Journa*l, 5(4). https://doi.org/10.1108/09513579210019549.

Soper, F. J. & Dolphin, R. (1964). Readability and Corporate Annual Reports. The Accounting Review, 39(2), 358.

de Souza, J. A. S.; Rissatti, J. C., Rover, S., & Borba, J. A. (2019): The linguistic complexities of narrative accounting disclosure on financial statements: An analysis based on readability characteristics. *Research in International Business and Finance,* 48, 59–74. https://doi.org 10.1016/j.ribaf.2018.12.008.

Subramanian, R., Insley, R. G., & Blackwell, R. D. (1993). Performance and Readability: A Comparison of Annual Reports of Profitable and Unprofitable Corporations. J*ournal of Business Communication*, 30(1), 49–61. https://doi.org/10.1177/002194369303000103.

Taylor, W. L. (1953). “Cloze Procedure”: A New Tool for Measuring Readability. *Journalism Quarterly*, 30(4), 415–433. https://doi.org/10.1177/107769905303000401.

Thoms, C., Degenhart, A., & Wohlgemuth, K. (2020). Is Bad News Difficult to Read? A Readability Analysis of Differently Connoted Passages in the Annual Reports of the 30 DAX Companies. *Journal of Business and Technical Communication,* 34(2), 157–187. https://doi.org/10.1177/1050651919892312.

The research article passed the review process. | Received: December 9, 2021**; Revised:** February 17, 2022**; Accepted:** March 15, 2022**; Pre-published online:** May 10, 2022**; Published in the regular issue:** December 2, 2022