

Contents lists available at ScienceDirect

Research in International Business and Finance

journal homepage: www.elsevier.com/locate/ribaf





Making sense and transparency in finance literature: Evidence from trends in readability

Oumaima Lahmar a,b,*, Luca Piras a,b

- ^a University of Cagliari, Italy
- ^b Department of Economics and Business Sciences, Via Sant'Ignazio da Laconi, 17, 09126 Cagliari, Italy

ARTICLE INFO

JEL Classification in this article: G00 G10

Keywords:
Readability
Transparency
Academic writing
Natural language processing
Textual analysis
Finance literature

ABSTRACT

Transparency continues to interest finance scholarship, as regards not just to financial reporting, but to a host of areas. Concomitantly, there is a growing emphasis on the transparency of the finance research process, with journals initiating requirements for uploading data and codes. However, little consideration is given to the transparency of finance scholarly texts, despite new emphasis by academic institutions and accreditation bodies on articles having an impact on practitioners. We use textual analysis to investigate the readability of articles in a selection of finance journals. Results evidence that academic articles are becoming less readable. Whether readability straightforwardly implies transparency is unclear, still, we consider these issues alongside our findings. Our study should be of great interest to those concerned with the state of finance scholarship.

1. Introduction and motivation

Scholars in finance are increasingly employing textual analysis to financial contexts. However, as yet this vein of research has not included examining the texts of finance scholarly articles themselves. For a selection of finance journals, we use textual analysis to investigate trends in articles' readability. Additionally, we use readability indexes for our analysis that produce approximations of the education grade, and the number of years of formal education, required to understand a text at first reading.

Results evidence that finance academic articles in leading journals are trending away from readability. Seeing readability as a component of research transparency, we consider our findings regarding readability as important for understanding trends in the state of finance scholarship.

Textual data analysis has had a longer period of usage in the arts and sciences, including anthropology; linguistics; medicine; and engineering (Shaffer et al., 2018), but until comparatively recent times has been less employed in economics and finance literature. The aim of this paper is to give a closer look at the finance academic narrative by using textual analysis to examine the anatomy of the scientific text in this literature, with reflection on the degree of its understandability by its intended audience. We integrate such reflections with a variety of traditional approaches to regarding textual data.

We note there are three primary approaches to textual data that correspondingly diverge in their assumptions about the nature of texts. First, a *positivist* approach assumes that language conforms to objective reality. The sense and the meaning extracted from the text are assumed to be objective. Words have meaning and potentially enhance actions only to the extent to which they allow some

^{*} Corresponding author at: University of Cagliari, Italy.

E-mail addresses: oumaima.lahmar@unica.it (O. Lahmar), pirasl@unica.it (L. Piras).

satisfactory connection to experience, implying that a meaningful sentence can be empirically verified and, therefore, the semantic content of sentences are defined by logical connections to patterns of experience. The *linguistic* approach posits that reality is embedded in the text, and that it emerges through language Lacity and Marius (1994). In this context it is assumed that language shapes reality, which means that language cannot be a neutral description of reality, but rather there is an interdependency between language and reality. Finally, the *interpretivist* approach assumes that the sense-making of language is subjective, with each side of communications assigning meaning and 'making sense' differently.

Following the seminal work of Bushman et al. (2004) on financial reporting, transparency continues to interest finance scholarship, and, as highlighted by Forssbaeck and Oxelheim (2014), is now valued in a growing array of contexts. Transparency is a concern not just for financial reporting, but for a host of areas within public policy, regulation, and corporate governance. Adding to the latter we discuss the transparency of finance scholarly articles through the lens of readability analysis. The readability analysis, as will be further explained in the following sections, is a quantitative methodology based on textual analysis, that makes it possible to estimate the degree of ease or difficulty an individual might face when reading a text.

We highlight the insightful definition of transparency by Forssbaeck and Oxelheim (2014) that transparency is the absence of asymmetry of information. However, they add to this that asymmetry of information always exists—a notion paralleled by the incomplete contract theory of Hart (1995); Hart and Moore (1999). Consequently, Forssbaeck and Oxelheim (2014) amend their definition of transparency by adding that for transparency to be present there also needs to be vehicles for the alleviation of informational asymmetry. There must be ways for information to be transferred to remove asymmetry. However, particularly in the digital age, transparency cannot be assessed by the volume of information but by its communication efficiency. We need to consider that readers have not just a bounded rationality (Conlisk, 1996) but also bounded time. Finance articles should be read efficiently for the intended audience to further the evolution of finance scholarship, be understandable outside academia, by market participants, and eventually by the general public, having, therefore, an impact on their behavior.

There are two widely spread prejudice about the language of science: the vocabulary used by scientists is complex, hard to spell and pronounce, and it extensively includes difficult words that may, and normally do, intimidate the general public; science has a heavily prosaic language, "objective" sentences that report facts, excluding both sentiments and personal beliefs.

Questioning this prejudice about language in finance academic narrative implies, in fact, the need for a readability analysis to investigate to which extent these claims are true, or false. How difficult to understand finance academic papers are? How has the complexity of scientific language evolved over time? Where does the readability of the scientific narrative fall, compared to other finance-related texts?

Our results suggest that finance academic articles are very difficult to read for most of the population. Not surprisingly given they are the products of advanced academic scholars. However, we also identify that finance articles are trending toward being increasingly difficult to read. Further, specifically finance academic literature is significantly less readable in the examined decades, compared with other finance-related texts. These results are interesting considering recent increasing emphasis by business accreditation agencies such as the AACSB for more inclusion of practitioner scholarship.

The layout of the paper is as follows. In Section 1, we review the literature related to the readability analysis applied to diverse finance-related texts. In Section 2, we describe the applied methodology and extensively portray the sample selection and treatment process. In Section 3, we present our results and discuss the findings. Finally, we outline our conclusions and suggest a path for further development of this stream of research.

2. Literature review

Communication is argued to be about establishing understanding between perspectives. However, any complex idea may be difficult to share and explain. Some communications include a set of scientific prose that may be written in a technical language, showing objectivity, that is indifferent to the audience, posing an additional challenge to them (Green et al., 2018; Padian, 2018; Sharp and Baron, 2011).

Objective information cannot be assumed to be intrinsically appealing during the communication process, as evidence, in fact, does not always speak for itself (Fischhoff and Scheufele, 2014; Green et al., 2018; Schimel, 2012; Sharp and Baron, 2011). That's why, one of the most effective ways to make a piece of information or an idea meaningful, is through the use of narratives and storytelling (Avraamidou and Osborne, 2009; Norris et al., 2005).

Individuals, particularly those unfamiliar with technical terms and mathematical formalizations, find it easier to understand information wrapped into narratives because stories are deeply rooted in our cognition as it has been considered a primary cognitive instrument to make sense (Herman, 2007; Joubert et al., 2019). Correspondingly, Cochrane (2005) advocates for finance articles to be constructed as a story written in a newspaper article fashion, wherein readers can quickly learn the main story with details volitionally. The audience can choose to read details if interested, but the entire article does not need to be fully read to understand the story.

Narratives can substantially encompass new, complex, technical information into a familiar context which enhances the involvement and the attention of the receivers on one side and stimulates their emotions on the other (O'keefe and Daniel, 2003; Zak, 2015; Sanford and Emmott, 2012). In other terms, humans seem to be prone to interpret and process information more efficiently when it is communicated in a narrative context. As a matter of fact, compared to other forms of communication, like descriptive or procedural explanation techniques, stories provoke our narrative appetite, increasing our interest in continuing the story, until the resolution is reached (Downs, 2014). That's why the use of narratives can help surmount the challenges communicators and receivers may face by addressing the barriers both parties may have.

Through finance research narratives, academia tries to build new lines of inquiry by making hypotheses, building explanatory and/

or descriptive models, and testing them (Parwada, 2020). High-quality academic research is conducted to expand knowledge of how financial markets and actors manifest: capital structure decisions, capital market efficiency, investor behaviors, etc. (Broby, 2018), as well as deepen understanding of how such behaviors and 'realities' are underpinned by seminal theories.

We consider that making sense of the financial discourse benefits from review of the nature of this discourse and its qualifying characteristics. It is necessary to understand the anatomy of academic narratives in terms of understandability, readability, and speed of spread among market practitioners. The complexity of the topics debated, and the complicated aspects of some concepts may entail the use of complex writing structures to transmit information accordingly. It may also entail frequent use of specific phrase structures, data analysis, quantitative models, and references.

Additionally, one of the main characteristics of the scientific narrative is the use of what is known as 'jargon' (Sawyer et al., 2008), defined by the Cambridge dictionary as "the set of words or phrases used by a particular group of people, especially in their work [...] and which most of the other people are intimidated with their sense." The quality of a scientific narrative is generally reflected by the correct use of grammatical tools and the appropriate sentence length while avoiding the unnecessary use of technical words, all this in the frame of the technical language shared and accepted within the community. Not surprisingly, finance makes no exception.

At the same time, it can be argued that gathering much of the terminology of finance scholarship into being jargon is overly reductive. Finance scholarship, like all fields, is replete with short-cut intellectual landmarks. Such landmarks are meant to enhance transparency within the academic community, but only among them, as they allow the avoidance of burdensome repetitive explanations of terms to informed readers (e.g., market efficiency, transaction costs, agency concerns, endogeneity, rational expectations, and many other examples).

It is arguable that simply avoiding such terms would act against transparency by reducing the efficiency of information transmission to targeted informed readers. Indeed, regarding the use of jargon, and idiosyncratic terms, literature from the field of sociology highlights that one person's 'cultural hole,' is another person's point of connectivity (Pachucki and Breiger, 2010; Vilhena et al., 2014). Further, as a field such as finance evolves into new topics and concerns, such landmarks must necessarily increase in number. And literature that effectively synthesizes theoretical frameworks across topics would become more populated with such landmarks.

In the context of the linguistic assessment of the scientific narrative in finance, researchers have developed new proxies to evaluate the quality of academic writings such as the readability index, the citation count, and the journal impact factor. Sufficient readability is crucial for textual communication. Readability can describe three different features of a written text. Klare (1974, 2000) suggests that it can describe a written text in terms of its legibility or its desirability and importance for the reader's sake or its understandability. Readability is also defined as the ease readers have with textual cognitive processing and the extent to which textual communication provides a good opportunity for readers to understand and extract meaning from it (Sattari, 2012). In fact, in the sense-making process, individuals need to collect and select readable and clear information to build up their current narrative and act upon it. Thus, readability is particularly crucial both to sense and decision-making. It is intuitively assumed that a text is said to be readable when it is composed of simple familiar words and short sentences (Danielson, 1987).

There are two approaches to measuring the difficulty or the readability of a text. Readability formulas are mainly mathematical equations that attempt to estimate the degree of difficulty of a written text based on its linguistic, semantic, and grammatical characteristics. Using readability formulas provides a predictive value of how challenging or sophisticated the text might be, without involving the reader.

The second approach relies mainly on the readers' judgment. It measures the degree of readability based on testing the readers' comprehension. The major criticism about the reliability and the validity of the second approach emerges when the judgment task becomes more complex with long texts requiring more time and effort (Klare, 1974). Meanwhile, readability formulas address specifically this aspect of text communications by measuring the degree of comprehensibility based on the complexity of words, in terms of the number of characters, syllable composition and number, the frequency of the use of uncommon words (technical, scientific, field-specific), the length and the structure of the sentences.

The evolution of the readability level of scientific papers has been investigated by the means of two types of measurement tools: the readability index which witnessed a decrease over time and the word count which highlighted the increase in the use of scientific and field-specific words at the expense of the commonly used words (Plavén-Sigray et al., 2017). Researchers, studying the readability of scientific journals and articles in different research fields, pointed out that writings are becoming more and more complex, and the readability level is steadily decreasing, which has lessened the accessibility to scientific knowledge (Bauerly et al., 2006).

In economics, studies on readability have been conducted on a wide range of texts and communications, such as academic marketing articles (Bauerly et al., 2006; Sawyer et al., 2008), economic reviews (McCannon, 2019), and tourism articles (Dolnicar and Chapple, 2015). In finance, the study of readability is still a nascent research topic. Studies currently have been testing the readability level in finance textbooks (Plucinski and Seyedian, 2013; Willey et al., 1998), annual reports, and financial disclosures (Aymen et al., 2018; Bonsall and Brian, 2017; Dempsey et al., 2012; Loughran and McDonald, 2014, 2016; Riley and Benjamin, 2015; Tan et al., 2014, 2015).

Readability analysis in finance can be categorized based on the type of text on which the analysis has been performed. From the literature, four types of textual data have been identified. The most extensively analyzed finance-related text is corporate disclosures (financial annual reports). There is quite an extended literature on the association between the readability of annual reports and the firm's performance (Dempsey et al., 2012; Hasan and Habib, 2020; Li, 2008; Loughran and Mcdonald, 2011; Rjiba et al., 2021), potential fraud, investor's feelings, judgment, and trading behavior (Miller, 2010; Tan et al., 2014, 2015), and financial analysts perception (Aymen et al., 2018; Bonsall, Brian, 2017; Lehavy et al., 2011; Rennekamp, 2012).

The second type of text on which readability analysis has been performed is analysts' reports. Surprisingly, the literature is relatively scarce, though in general, literature on financial analysts is rather extended (De Franco et al., 2015; Hsieh et al., 2016).

The third type of finance-related text on which readability analysis has been done is finance textbooks. After careful research of the literature, only two studies have attempted to estimate the readability level of finance and corporate finance textbooks so far (Plucinski and Seyedian, 2013; Willey et al., 1998). The purpose of these works was to identify the more readable textbooks for finance students.

Finally, the fourth type of document on which readability analysis has been conducted is academic papers published in finance. Analysis of the readability of finance scholarly articles is limited. Only two studies, as far as this paper has been written, have considered the issue of the understandability of academic articles (Lee and French, 2011; Berninger et al., 2021). Such studies differ from this one in terms of selected sample, methodology, and purpose.

Table 1 above synthesizes the main findings the previous contributions to literature have provided on readability in different finance-related texts. It displays some relevant examples of studies conducted on such texts, applying the most used readability indexes in the field. The papers have been sorted based on the type of documents studied. The first category is formed by corporations' disclosure documents, such as 10-K forms and earnings press releases. The second grouping contains different types of analysts' reports; the third category is composed of finance textbooks and, finally, the fourth considers academic articles published in finance journals.

Surprisingly, considering the difficulty of academic narratives in finance has so far been thinly addressed. Lee and French (2011) is a small investigation of the readability level of academic articles published in the *Journal of Property, Investment, and Finance*. The authors apply five readability indexes to measure readability, by estimating the educational level required to read and understand the sample articles. Their results suggest that the scientific articles published can be read by college students which ranks them in the range of difficult texts. A second, and more recent, paper investigates the link between citations in top leading finance journals and the readability (Berninger et al., 2021). The authors evidence negative correlation between the readability of the academic articles and the number of citations. Additionally, they find a link between the high complexity of research texts and the high number of citations these papers receive. On the other hand, complex abstracts result in a lower number of citations (Berninger et al., 2021).

Throughout the course of scientific history, researchers intentionally or unintentionally implemented a narrative thought to make sense of reality, and create rational explanations, descriptions, and interpretations of their own experiences with the external world. Narrative thought, in fact, relates human experiences to scientific explanations, leading eventually to better assimilation of both. By shedding the light on the academic discourse, we evaluate and discuss how accessible and understandable are finance academic articles compared to other finance-related narratives such as finance textbooks, corporate annual reports, and financial analysts' reports.

3. Methodology

3.1. Indexes

Among the many indexes developed to estimate readability, three of them appear to be the most widely accepted and used in analyzing social sciences, economics, and business-related corpora: 1) Flesch Reading Ease; 2) Flesch-Kinkaid Reading Grade; 3) Gunning Fog Index.

Readability, as defined previously, is a metric that describes the ease/difficulty to understand a textual communication. Readability

Table 1Results of previous studies on readability.

Author	Sample	Index	Results	
Li (2008)	Corporate disclosure 10-k filling	GFI	19.39	
Miller (2010)	Corporate disclosure 10-K filling	GFI and Plein English	GFI= 19.943; PE= 21.157	
Lehavy et al. (2011)	Corporate disclosure 10-k filling	GFI	19.53	
Lawrence (2013)	Corporate disclosure 10-k filling	GFI	19.02	
Loughran and McDonald, 2014	Corporate disclosure 10-k filling	GFI	1994–2002: 18.44 2003–2011: 18.94	
Bonsall, Brian (2017)	Corporate disclosure 10-K filling	GFI	19.397	
Lo et al. (2017)	Corporate disclosure 10-k filling	GFI	18.020	
Aymen et al. (2018)	Corporate disclosure 10-K filling	GFI - FRE	GFI= 16.82; FRE= 38.2	
Hasan (2018)	Corporate disclosure 10-k filling	BOG index	82.847	
Hasan and Habib (2020)	Corporate disclosure 10-k filling	GFI	19.76	
Dempsey et al. (2012)	Corporate disclosure annual report	FKI	12.89	
Tan et al. (2015)	Corporate disclosure earnings press release	GFI	MD&A: 25.71; Littlefuse Inc: 15.18 / experiment	
De Franco et al. (2015)	Equity Analysts reports	GFI - FKI - FRE	GFI= 18.71; FKI= 51.85 FRE= 8.48	
Hsieh et al. (2016)	Analyst reports	GFI	14.008	
Willey et al. (1998)	Undergraduate Corporate Finance Textbooks	FKI	Between 8.83 and 12.64	
Plucinski and Seyedian (2013)	Introductory finance books	FKI	Block: 12.7; Brignam: 11.0; Gitman:12.8; Keown:12.1; Ross:10.1	
Berninger et al. (2021)	Academic articles from JF, JFE, RFS Academic articles from 11 finance-oriented journals	FKI	Increasing over time	
Lee and French (2011)	Academic papers from the Journal of Investment and Property Management	FRE-FKI	FRE= 30.4 FKI= 15.5	

formulas rely on counts of language variables to provide an estimation of the level of ease/difficulty in making sense of a given text. Readability metrics could be seen as a prediction, in the sense that there is no actual feedback from the readers themselves to judge the extent to which the text is understandable.

Flesch Reading Ease is a score developed by Flesch and Rudolf (1949). In his book "The Art of Readable Writing" he shows the keys and tools to producing accessible and effective textual communication. The Reading Ease score (aka. human interest) is computed based on the following variables: the average number of words per sentence (sentence length); and the average number of syllables per word (word length). The resulting reading ease score is between 0 (unreadable) and 100 (easy for a literate person).

$$RE = 206.835 - 1.015 \left(\frac{\mathbb{N}.ofwords}{\mathbb{N}.ofsentences} \right) - 84.6((\mathbb{N}.ofsyllabes) / (\mathbb{N}.ofwords))$$
 (1)

The table below gives guidance to interpret the index's magnitude, by linking its value to the parameters it uses for the calculation. Higher scores are linked to very easy text, readable by virtually everyone with a very low level of schooling, while a "standard" text is estimated to be readable by individuals who attended seven or eight years of school. That is, also, the level of difficulty a common newspaper article should present. Texts composed of very long and articulated sentences, in which words are, on average, multisyllabic ones (more than 2), tend to require a higher education level, typically a college degree or higher.

Flesch Kincaid Grade is instead a readability score developed to estimate the reading comprehension of the United States Navy personnel and evaluate the understandability of the supplied technical manuals. It estimates the US school grade level required to understand a given text (Kincaid et al., 1975). FKG is one of the most used readability indexes and although it is based on the same variables, it is negatively correlated with the Flesch Reading Ease.

$$FKG = 0.39 \left(\frac{\mathbb{N}.ofwords}{\mathbb{N}.ofsentences} \right) + 11.8 \left(\frac{\mathbb{N}.ofsyllabes}{\mathbb{N}.ofwords} \right) - 15.59$$
 (2)

Gunning (1968) proposed the Gunning-Fog index to compute the readability of a text. Some scholars went as far as to consider the Gunning-Fog index as a measure of the financial statement readability (Biddle et al., 2009). It is a common measure across virtually all fields, and it shows to be the chosen measure in financial research despite the debate around its reliability in indicating the readability of finance-related narratives.

The formula used is straightforward as it is a linear combination of 1) the average number of words per sentence and 2) the proportion of complex words (composed of more than 2 syllables) to the total number of words. The resulting score provides an estimation of the years of education required to comprehend the text at first reading. Texts for the general audience need a Gunning-Fog index below 12 and below 8 for a universal understanding.

$$GFI = 0.4 \left(\frac{\mathbb{N}.ofwords}{\mathbb{N}.ofsentences} + 100 \frac{\mathbb{N}.ofcomplexwords}{\mathbb{N}.ofwords} \right)$$
(3)

All indexes have been identified as in some part flawed. Either they apply better to certain fields than others, or they are affected by intrinsic limits and biases that reduce, sometimes significantly, their efficacy. As much as readability indexes are being used to measure the understandability of business communications and finance texts, they are arguably imperfect. Some claim, for instance, that the Gunning-Fog index does not really reflect the ease or the difficulty of the written text in the business field, given the fact that it is significantly dependent on the number of complex words that are defined as words composed of more than two syllables. This claim is supported by the idea that words used in such documents are generally polysyllabic words (corporation, telecommunication, marketing) that do not usually require readers to have advanced knowledge or consult dictionaries to understand their meaning. However, despite limitations, we consider that these methodologies provide considerable utility for offering insights into how finance scientific literature addresses knowledge among practitioners and enhance financial literacy among investors.

4. Sample selection

Evaluating the understandability of the scientific narrative and eventually, estimating the level of comprehension of the audience requires a textual sample of scientific research in finance. For this purpose, our sample includes academic articles published in a selection of leading journals. These journals are selected based on their citation score and their ranking according to the CABS system. The five selected journals are the ones that satisfy contemporarily two criteria: being in the top 10 most cited and top 10 CABS ranking.

1) Journal of Financial Economics, 2) Review of Financial Studies, 3) The Journal of Finance, 4) Journal of Corporate Finance, 5) Journal of Financial and Quantitative Analysis.

These journals are flagship journals in the field, in the sense that they publish research articles covering relatively the full spectrum of research in finance. Among the articles published in each journal, exclusively, most cited articles were retrieved based on a preset threshold of a minimum of 20 citations per year.

One could claim that citation indicators are not enough to judge whether an article has an impact or not, which is quite common in many papers. However, one fact should be emphasized is that impact in science does not always follow a linear path. In other words, relevant articles, even if not immediately "impactful" when published, may be relevant to developing impactful papers later. In fact, it is argued that citations represent the relative scientific significance and quality of a paper (Crane, 1974). Roles of citations seem to justify this claim as well as the use of citation count as a selection criterion of the sample. As a matter of fact, citing may be used to present background research, use a methodology, acknowledge pioneering insights, identify unique publications and breakthroughs,

criticize, correct, disclaim, and so on (Aksnes et al., 2019).

We started with a data source including 26,332 published articles. The citation-based screening of the sample was conducted considering the citation count available on Scopus based on the predefined threshold of 20 citations/year. Then, we downloaded the portable document format (PDF) of the articles selected and the resulting sample was composed of 1258 articles. We eliminated the articles that are not machine-readable (scanned PDFs) and we ended up with 1055 published articles between 1978 and 2019 (Tables 2–4).

Focusing on the content analysis of the academic narrative, the core instrument to perform semantic-oriented analysis of academic papers is to apply Natural Language Processing (NLP), which is a study that encompasses three major fields: linguistics, data science, and text mining. It is the application of computational and modeling approaches to written texts to analyze various features of the language.

NLP is a discipline that allows computers, through specific algorithms, to perform massive analyses on a virtually infinite number of texts and provide insights about the content and the features of the data. Given that the study is performed on academic articles that were initially available in PDF form, it was necessary to prepare the dataset by extracting the textual content of these PDFs to be able to execute the readability measurement.

Before stepping into the readability measurements, it was also necessary to parse the extracted texts. The challenge at this stage was to scan the files and separate the textual from the non-textual data. Given the size of the selected corpus, such an activity cannot possibly be performed manually. Using Python, the key library by which the data was preprocessed was SciPDF which is dedicated to parsing scientific articles.

The parsing process permits scanning and pulling up the different features of articles: title; year; DOI; abstract; figures; tables; references and dividing the text into sections. The features parsed on which the readability analysis was conducted were the abstract and the textual sections of each article without (tables; figures; references...). To have a complete text we concatenated the abstract to the different parsed sections and then we run the readability measurements.

Textstat library was used to apply computational predefined functions such as the word count, word frequency, readability metrics, etc, to our text sample. As mentioned above, three readability measures have been calculated, resulting in the graphs presented in the following section.

5. Results and discussion

Our evidence suggests that the increase in complexity and length of papers over time demonstrate that scholars make their narrative less intelligible, suggesting they consider a complex narrative as a key, yet difficult to admit, feature to fulfill expected achievements: to gain citations. A complex narrative would impress editors and reviewers more than a simple and easy-to-grasp one, particularly when paired with intricate quantitative methodologies and models.

Articles in the sample have a yearly growing average length. The growing pattern starts to be clear in the first half of the eighties. In 1980 the average length of articles is lower than seven thousand words, ten years later is close to nine thousand, at the beginning of the new century is close to twelve thousand, but in the last decade, it remains clearly above the latter.

Scholars, over time, have written longer papers, engaging the readers in more articulated arguments, using many more words, though maintaining basically the same papers' structure. It takes more words to cover the topics or, perhaps, this is the sign that scientific narrative seeks different goals, other than simply reporting research results Graph 1.

Graph 2 represents the plot of the Gunning-Fog Index (GFI henceforth). Measures appear a bit noisy in the two extremes of the graph, due to the distribution of papers considered in the sample. In fact, the algorithm is run in the initial and final segments of our data series, on a relatively small number of articles, as displayed in Table 5.

Starting from the mid-eighties, with the increase in the number of papers considered in the sample, the trend appears clearer: the index value rises rather regularly, showing a constant tendency of the scientific narrative to evolve into a more complex and difficult

Considering 1985 as the year in which the index starts to be descriptive, the same has a value of 17.67. It thereafter increases continuously, in 1996 it crosses the value of 18.32, arrives to 19.78 in 2006 and it finally reaches the remarkable value of 21.54 (its peak) in the last few years. This implies the need for more than 21 years of formal education to comprehend a finance scientific paper. People with less than a master's degree find it difficult to grasp the meaning and findings of financial studies.

Table 2
Modified from The Art of Readable Writing (Flesch and Rudolf, 1949).

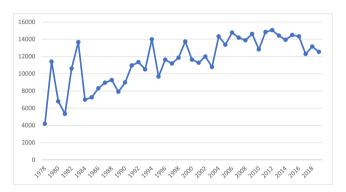
Flesch Reading Ease					
Score	Level	Words/Sentence		Syllables/word	Estimated school grade completed
90–100	Very easy	8 or fewer	1.23 or fewer		4th
80-90	Easy	11	1.31		5th
70-80	Fairly easy	14	1.39		6th
60-70	Standard	17	1.47		7th or 8th
50-60	Fairly hard	21	1.55		Some high school
30-50	Hard	25	1.67		High school or some college
0-30	Very hard	29 or more	1.92 or more		College

Table 3
Modified from The Art of Readable Writing (Flesch and Rudolf, 1949).

Flesch-Kinkaid Reading Grade Level				
Estimated school grade completed	Level	Words/sentence	Syllables/ word	Score
4th	Very easy	8 or fewer	1.23 or fewer	90–100
5th	Easy	11	1.31	80-90
6th	Fairly easy	14	1.39	70-80
7th or 8th	Standard	17	1.47	60-70
Some high school	Fairly hard	21	1.55	50-60
High school or some college	Hard	25	1.67	30-50
College	Very hard	29 or more	1.92 or more	0-30

Table 4
Modified from The Techniques of Clear Writing (Gunning, 1968).

The Gunning-Fog	Index: How do popular consumer publications stack up?	
Fog Index	Reading level by grade	Reading level by publication
20 +	Post-graduate plus	Government information
17-20	Post-graduate	Academic journal papers
16	College senior	Standard medical consent forms are written at the 16th-grade level.
15, 14, 13	College junior, sophomore, freshman	Forbes-University textbooks
Danger line		
12-11	High school senior, junior	The Wall Street Journal
10	High school sophomore	The Sun Magazine
9	High school freshman	National Geographic
8	8th grade	The Bible
7	7th grade	TV Guide
6	6th grade	Parade



Graph 1. Evolution of the average length of papers in n. of words.



Graph 2. Evolution of Gunning Fog Index over time.

Table 5Number of papers in the sample per decade.

years	1978–1985	1986–1995	1996–2005	2006–2015	2016–2019
No. of papers	21	96	324	467	147
	2.00	9.09	30.72	44.26	13.93

At the same time, it is reasonable to imagine that, once finished their education path, market practitioners need to keep updating their capacity to read and understand finance-related texts since the implied narrative evolves rapidly toward a higher level of complexity. Meanwhile, non-alphabetized individuals keep facing mounting struggles in understanding markets' dynamics, news, and modern investment strategies, resulting in misunderstanding and mispricing risks. Hence the call for widespread financial inclusion, through education enhanced by world financial authorities may be in vain if the trend of the understandability of the narrative keeps on such a path.

Estimating the readability level on the same sample using the Flesh Kincaid Index (FKI henceforth), the results plotted in the following graph display an almost equal trend, compared to GFI, with a noisy shape of the curve in the first period of the analysis and a significant increase in the readability index, which reaches its highest value (>20) as it approaches 2019 (Graph 3).

The Flesh Reading Ease index, as illustrated in Graph 4, must instead be read reversed (the higher the index is, the more understandable the text is and vice versa) since it measures how easy it is to read and comprehend a text. Its value reaches the lowest level as it approaches 2019. Such a result is therefore consistent with the two previous indexes, shown before. The graph shows that the index value drops rather sharply from 1985 to 2013 and it stabilizes a little before a further drop in the last few years.

As mentioned above, papers became significantly longer over time, and the number of words largely influence the papers' readability level. The latter are not exclusively meant to report scientific activity and the implied results, but they also seem to tell stories, and as such, they require many more words and rhetorical images.

Lower readability may attract higher citations (Berninger et al., 2021). Or more facile, easier-to-read, text may be simply more challenging to compose. Or, again, trends in readability tangentially reflect evolution in dialogue and writing of subgroups and communities of scholars that gradually establish language connectivity among themselves that furthers intragroup over intergroup communication. Regarding shifting from written explanations to graphical representations, it suffers relevant limitations with such representations intended to support rather than substitute for narratives.

As for the understandability of different types of finance-related texts, based on the results reported in the literature, shown in Table 1, we highlight some further aspects. We consider if and how narratives' readability changes on its path from academic texts to professionals' ones, to investors and the general public. The proposed taxonomy of finance-related texts can be sorted considering the role of texts' protagonists also regarding the willingness to take a given level of reputational risk by the authors and the credibility level granted by the average reader. The description of the texts' protagonists is also important to better depict the sample characteristics. In linguistic studies, scholars always highlight the relationship among communication protagonists. This is considered an essential feature of any text and depinning it is essential for the sense-making process. We, therefore, propose a taxonomy based on three categories of narratives in the financial markets: a) High-Risk, mainly concerning corporations; b) Medium-Risk, mainly regarding financial analysts; c) Low-Risk, mainly considering academia.

Communication protagonists in corporation disclosure documents range from managers, executives, analysts, stockholders, and specialized press to individual investors and so on. The readability level, on average, using GFI is above 18, though not as difficult as the academic literature, we claim that transparency appears to be seldom sacrificed. We consider that authors in this category have a goal to disclose corporations' activities, describe their performances within the industry, announce their future investment projects and provide details on their current situations. In doing so, they rely, extensively, on impression management in their narratives. We believe that an explanation may lay on the fact that corporations are accountable for the disclosures they publish regularly, and the provided information is likely to affect valuations, market expected returns, perceived risk profile, and prices.

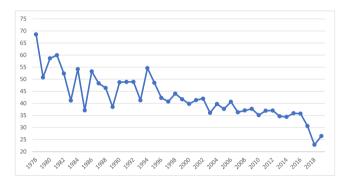
We consider the possibility that authors in this document category might consider both the utility and disutility of transparency. Corporations may seek a balance between full and transparent disclosure on one side and preserving corporation interests on the other (Li, 2008; Rjiba et al., 2021). The more information they are willing to hide, the higher will be the reputational risk they take. For this reason, we consider this document's class a high-risk one (Forssbaeck and Oxelheim, 2014).

The second class of narratives we take into account is produced by financial analysts. There is an extensive literature on financial analysts' behavior, with focus on potential interest conflicts, as markets assign a great deal of attention to analyst reports. Different elements influence market reactions: the number and frequency of reports on a given security and or company may strongly affect the market's interest in that security or company (Bradley et al., 2008). We consider that authors will seek an overall equilibrium between performance and reputation by the means of their narratives. Clearly analysts play a very important role in building 'frames' around information collections and forecasts, by which they tune their personal reputational risk.

Finance-related texts, though very diverse, have few types of protagonists. Academic articles are mainly written by scholars and are meant to be read by scholars, with some exceptions represented by post-grad and Ph.D. students on one side and professional financial analysts on the other. We cannot, of course, exclude that other categories of readers may, from time to time, show some specific interest in scientific papers, but we feel comfortable in considering them exceptions. Narrative, therefore, may well be influenced by the relationship authors are willing to build with their target readers, at the same time it is unavoidable to consider that a clearer exposition of research results and argumentations would better contribute both to the spread of knowledge and faster implementation of new ideas into practice.



Graph 3. Evolution of Flesh Kincaid Index over time.



Graph 4. Values over time of Flesh Reading Ease Index.

6. Conclusions

There is growing emphasis on the transparency of the finance research process, with journals initiating requirements for uploading data and codes. However, little consideration is given to the transparency of finance scholarly texts, despite new emphasis by academic institutions and accreditation bodies on articles having an impact on practitioners. We use textual analysis to investigate the readability of articles in a selection of finance journals. We also examine differences in readability between finance scholarly works and practitioner-authored writings. Results evidence that finance academic articles are becoming less readable. Contemplating readability as a component of transparency, we consider our findings regarding readability as important for understanding trends in the state of finance scholarship.

We investigate scientific literature in the field of finance, along with some other finance-related texts. To do so, we apply a variety of indexes of readability. Our results show, perhaps not surprisingly, that academic articles are very difficult to read, and such difficulty has increased over time. Further, academic articles are significantly less readable in the examined decades compared with other finance-related texts. That is, the full texts may consist of more polysyllabic, longer words, and longer sentences.

Hence, since scientific papers ought to represent the primary vehicle to spread scientific knowledge among the public, even when alphabetized, longer sentences and polysyllabic words are a source of distorted interpretations, biasing understanding, and resulting in an erroneous transposition of the theory into bad practice. Better-written documents produce less opacity. The role of scientific research is to lessen such ambiguity which requires the use of accessible language. A further motivation that we consider, is represented by the fact that the more readable academic articles are, the less susceptible are to distorted narratives Camerer and George Loewenstein (1989). In examining our results, we also reflect on the use of narrative in finance academic writing, as well as the impact of discipline specific catch phrases and jargons. The use of language evolves to specific practices within communities, including the finance scholarly one, but does this gradual shaping of scholarly communication, and establishing of language points of connectivity, also restrict its wider readability? We consider that our results and analysis will be of great interest to academics, practitioners, and business educators.

CRediT authorship contribution statement

Oumaima Lahmar: Conceptualization, Methodology, Data curation, Software, Investigation, Writing — original draft preparation, Revision. **Luca Piras:** Conceptualization, Methodology, Supervision, Validation, Investigation, Writing — review & editing, Revision.

Data availability

Data will be made available on request.

References

Aksnes, Dag W., Liv, Langfeldt, Paul, Wouters, 2019. "Citations, citation indicators, and research quality: an overview of basic concepts and theories.". SAGE Open 9 (1). https://doi.org/10.1177/2158244019829575.

Avraamidou, Lucy, Osborne, Jonathan, 2009. "The role of narrative in communicating science.". Int. J. Sci. Educ. 31 (12), 1683–1707. https://doi.org/10.1080/09500690802380695

Aymen, Ajina, Ben Saad, Sourour, Msolli, Badreddine, 2018. "The effect of annual report resdability on financial analysts' behavior.". J. Econ. 12.

Bauerly, Ronal, J., Don T.Johnson, Mandeep Singh, 2006. Readability and Writing Well. Mark. Manag. J.

Berninger, Marc, Florian, Kiesel, Dirk, Schiereck, Eduard, Gaar, 2021. "Citations and the readers' information-extracting costs of finance articles.". J. Bank. Financ. 131, 106188 https://doi.org/10.1016/j.jbankfin.2021.106188.

Biddle, Gary, C., Gilles Hilary, Rodrigo, S. Verdi., 2009. How Does Financial Reporting Quality Relate to Investment Efficiency? J. Account. Econ 48 (2–3), 112–131. https://doi.org/10.1016/j.jacceco.2009.09.001.

Bonsall, Samuel B., Brian, P.Miller, 2017. "The impact of narrative disclosure readability on bond ratings and the cost of debt.". Rev. Account. Stud. 22 (2), 608–643. https://doi.org/10.1007/s11142-017-9388-0.

Bradley, Daniel J., Bradford, D.Jordan, Jay, R.Ritter, 2008. "Analyst behavior following IPOs: The 'Bubble Period' Evidence.". Rev. Financ. Stud. 21 (1), 101–133. https://doi.org/10.1093/rfs/hhl028.

Broby, Daniel, 2018. "The impact of academic research in finance (impact ratio). In: Kensinger, J.W. (Ed.), Research in Finance, Vol. 34. Emerald Publishing Limited, pp. 217–228.

Bushman, Robert M., Joseph, D.Piotroski, Abbie, J.Smith, 2004. "What determines corporate transparency. In: J. Account. Res, 42, pp. 207–252. https://doi.org/10.1111/j.1475-679X.2004.00136.x.

Camerer, Colin, George Loewenstein, Martin Weber, 1989. "The curse of knowledge in economic settings: an experimental analysis.". J. Political Econ. 97 (5), 1232–1254. https://doi.org/10.1086/261651.

Cochrane, John H., 2005, "Writing Tips for Ph. D. Students." 13.

Conlisk, John, 1996. "Why bounded rationality?". J. Econ. Lit. 34 (2), 669-700.

Crane, Diana, 1974. "Review of social stratification in science". Adm. Sci. Q. 19 (2), 264-266. https://doi.org/10.2307/2393902.

Danielson, Kathy Everts, 1987, "Readability Formulas: A Necessary Evil?" 12.

De Franco, Gus, Ole-Kristian Hope, Dushyantkumar Vyas, Zhou, Yibin, 2015. "Analyst report readability.". Contemp. Account. Res. 32 (1), 76–104. https://doi.org/10.1111/1911-3846.12062.

Dempsey, Stephen J., David M, Harrison, Luchtenberg, Kimberly F., Michael J, Seiler, 2012. "Financial opacity and firm performance: the readability of reit annual reports.". J. Real. Estate Financ. Econ. 45 (2), 450–470. https://doi.org/10.1007/s11146-010-9263-2.

Dolnicar, Sara, Chapple, Alexander, 2015. "The readability of articles in tourism journals.". Ann. Tour. Res. 52, 161–166. https://doi.org/10.1016/j.annals.2015.03.003.

Downs, J.S., 2014. "Prescriptive scientific narratives for communicating usable science.". Proc. Natl. Acad. Sci. 111 (Supplement_4), 13627–13633. https://doi.org/10.1073/pnas.1317502111.

Fischhoff, B., Scheufele, D.A., 2014. "The science of science communication II.". Proc. Natl. Acad. Sci. 111 (Supplement_4), 13583–13584. https://doi.org/10.1073/pnas.1414635111.

Flesch, Blanco, Rudolf, 1949, "Flesch-the-Art-of-Readable-Writing.Pdf." Retrieved July 27, 2022 (https://dc135.files.wordpress.com/2012/11/flesch-the-art-of-readable-writing.pdf).

Forssbaeck, Jens, Oxelheim, Lars, 2014. The Oxford Handbook of Economic and Institutional Transparency. Oxford University Press,

Green, Stephanie J., Kirsten, Grorud-Colvert, Heather, Mannix, 2018. "Uniting science and stories: perspectives on the value of storytelling for communicating science". In: Shanahan, M.-C. (Ed.), FACETS, 3, pp. 164–173. https://doi.org/10.1139/facets-2016-0079.

Gunning, Robert, 1968, The Technique of Clear Writing. McGraw-Hill.

Hart, Oliver, Moore, John, 1999. "Foundations of incomplete contracts.". Rev. Econ. Stud. 66 (1), 115–138.

Hart, Oliver D., 1995, Firms, Contracts, and Financial Structure. Oxford: New York: Clarendon Press; Oxford University Press.

Hasan, Mostafa Monzur, 2018, "Readability of Narrative Disclosures in 10-K Reports: Does Managerial Ability Matter?". Eur. Account. Rev. Forthcom.39.

Hasan, Mostafa Monzur, Habib, Ahsan, 2020. "Readability of narrative disclosures, and corporate liquidity and payout policies.". Int. Rev. Financ. Anal. 68, 101460 https://doi.org/10.1016/j.irfa.2020.101460.

Herman, David, 2007. "Storytelling and the sciences of mind: cognitive narratology, discursive psychology, and narratives in face-to-face interaction. Narrative 15 (3), 306–334.

Hsieh, Chia-Chun, Kai, Wai Hui, Yao, Zhang, 2016. "Analyst report readability and stock returns.". J. Bus. Financ. Account. 43 (1–2), 98–130. https://doi.org/10.1111/jbfa.12166.

Joubert, Marina, Lloyd, Davis, Jennifer, Metcalfe, 2019. "Storytelling: the soul of science communication.". J. Sci. Commun. 18 (05) https://doi.org/10.22323/2.18050501.

Kincaid, J.P., Jr. Fishburne, Rogers Robert P., Chissom Richard L., and Brad S., 1975, Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel: Fort Belvoir, VA: Defense Technical Information Center. doi: 10.21236/ADA006655.

Klare, George R., 1974. "Assessing readability.". Read. Res. Q. 10 (1), 62–102. https://doi.org/10.2307/747086.

Klare, George R., 2000. "The measurement of readability: useful information for communicators.". ACM J. Comput. Doc. 24 (3), 107–121. https://doi.org/10.1145/344599.344630.

Lacity, Mary C., Marius, A.Janson, 1994. "Understanding qualitative data: a framework of text analysis methods.". J. Manag. Inf. Syst. 11 (2), 137–155.

Lawrence, Alastair, 2013. "Individual investors and financial disclosure.". J. Account. Econ. 56 (1), 130–147. https://doi.org/10.1016/j.jacceco.2013.05.001. Lee, Stephen, French, Nick, 2011. "The Readability of Academic Papers in the Journal of Property Investment & Finance" edited by N. Crosby. J. Prop. Invest. Financ. 29

(6), 693–704. https://doi.org/10.1108/14635781111171814.

Lehay, Reuven, Feng Li, Kenneth Merkley, 2011. "The effect of annual report readability on analyst following and the properties of their earnings forecasts.".

Account. Rev. 86 (3), 1087–1115.

Li, Feng, 2008. "Annual report readability, current earnings, and earnings persistence.". J. Account. Econ. 45 (2–3), 221–247. https://doi.org/10.1016/j.

jacceco.2008.02.003.
Lo, Kin, Felipe Ramos, Rafael Rogo, 2017. "Earnings management and annual report readability.". J. Account. Econ. 63 (1), 1–25. https://doi.org/10.1016/j. jacceco.2016.09.002.

Loughran, Tim, Mcdonald, Bill, 2011. "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks.". J. Financ. 66 (1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x.

Loughran, Tim, McDonald, Bill, 2014. "Measuring readability in financial disclosures.". J. Financ. 69 (4), 1643-1671.

Loughran, Tim, Mcdonald, Bill, 2016. "Textual analysis in accounting and finance: a survey." J. Account. Res. 54 (4), 1187–1230. https://doi.org/10.1111/1475-679X.12123.

McCannon, Bryan C., 2019. "Readability and research impact.". Econ. Lett. 180, 76–79. https://doi.org/10.1016/j.econlet.2019.02.017.

Miller, Brian P., 2010. "The effects of reporting complexity on small and large investor trading.". Account. Rev. 85 (6), 2107-2143.

Norris, Stephen, Sandra, P., Guilbert, M., Martha, L., Smith, Shahram Hakimelahi, Linda, M.Phillips, 2005. "A theoretical framework for narrative explanation in science.". Sci. Educ. 89 (4), 535–563. https://doi.org/10.1002/sce.20063.

Padian, Kevin, 2018, "Narrative and 'Anti-Narrative' in Science: How Scientists Tell Stories, and Don't." 11.

Parwada, Jerry T., 2020, "The Scientific Value of Finance after the Crisis: Evidence from Research Usage and Financial Innovation." 57.

O'keefe, Daniel, J., 2003. Message Properties, Mediating States, and Manipulation Checks: Claims, Evidence, and Data Analysis in Experimental Persuasive Message Effects Research. Communication Theory.

Pachucki, M.A., Breiger, R.L., 2010. Cultural Holes: Beyond Relationality in Social Networks and Culture. Annu. Rev. Sociol. 36, 205–224. https://doi.org/10.1146/annurev.soc.012809.102615.

Plavén-Sigray, Pontus, Granville James Matheson, Björn Christian Schiffler, William Hedley Thompson, 2017. "The readability of scientific texts is decreasing over time. ELife 6. https://doi.org/10.7554/eLife.27725.

Plucinski, Kenneth J., Seyedian, Mojtaba, 2013. "Readability of introductory finance textbooks.". J. Financ. Educ. 39 (1/2), 43-52.

Rennekamp, K.R.I.S.T.I.N.A., 2012. "Processing fluency and investors' reactions to disclosure readability.". J. Account. Res. 50 (5), 1319-1354.

Riley, Tracey J., Benjamin, L.Luippold, 2015. Managing investors' perception through strategic word choices in financial narratives. J. Corp. Account. Financ. 26 (5), 57–62. https://doi.org/10.1002/jcaf.22064.

Rjiba, Hatem, Samir Saadi, Sabri Boubaker, Xiaoya (Sara), Ding, 2021. "Annual report readability and the cost of equity capital.". J. Corp. Financ. 67, 101902 https://doi.org/10.1016/j.jcorpfin.2021.101902.

Sanford, Anthony J., and Catherine Emmott, 2012, Mind, Brain and Narrative. Cambridge: Cambridge University Press,

Sattari, Setayesh, 2012. Essays on the Issues of Readability in Business Disciplines. Lulea: Lulea tekniska universitet.

Schimel, Joshua, 2012, Writing Science: How to Write Papers That Get Cited and Proposals That Get Funded. Oxford; New York: Oxford University Press.

Sawyer, A., Juliano Laran, G., 2008. The Readability of Marketing Journals: Are Award-Winning Articles Better Written? Journal of Marketing 72 (1), 108–117. Shaffer, Victoria A., Elizabeth, S.Focella, Hathaway, Andrew, Scherer, Laura D., Brian, J.Zikmund-Fisher, 2018. "On the Usefulness of Narratives: An Interdisciplinary Review and Theoretical Model.". Ann. Behav. Med. 52 (5), 429–442. https://doi.org/10.1093/abm/kax008.

Sharp, J.O.N.A.T.H.A.N.H., Baron, Nancy, 2011. "Review of escape from the ivory tower: a guide to making your science matter, BaronNancy.". Oceanography 24 (2), 216–218

Tan, Hun-Tong, Elaine Ying Wang, Bo Zhou, 2014. "When the use of positive language backfires: the joint effect of tone, readability, and investor sophistication on earnings judgments.". J. Account. Res. 52 (1), 273–302.

Tan, Hun-Tong, Elaine Ying Wang, Bo Zhou, 2015. "How does readability influence investors' judgments? consistency of benchmark performance matters.". Account. Rev. 90 (1), 371–393.

Vilhena, Daril, Jacob Foster, Martin Rosvall, Jevin West, James Evans, Bergstrom, Carl, 2014. "Finding cultural holes: how structure and culture diverge in networks of scholarly communication.". Sociol. Sci. 1, 221–238. https://doi.org/10.15195/v1.a15.

Willey, Thomas, Liane Holliday Willey, Jose Mercado Mendez, 1998. "The readability of undergraduate corporate finance texts.". J. Financ. Educ. 24, 31–34. Zak, Paul J., 2015. "Why inspiring stories make us react: the neuroscience of narrative.". Cerebrum: Dana Forum Brain Sci. 2015, 2.