Original Research

The Correlation Between Annual Reports’ Narratives and Business Performance: A Retrospective Analysis

Yubin Qian and Ya Sun

Abstract
This article investigates the extant literature on the correlation between narratives in corporate annual reports and corporate performance. Prior studies are reviewed for overall characteristics, research topics, theoretical foundations, and methods. Articles published between 2000 and 2018 were analyzed using the content analysis method. The results demonstrated that prior studies generally show an increasing trend with salient interdisciplinarity. Mapping and predictability between annual reports’ narratives and business performance have been the prevailing topics. The impression management and agency theories are the most frequent theoretical references. More importantly, complexity of research methods was found in data, analytical approaches, and variables. The emphasis on narratives in prior research proves the necessity of contextualizing narratives in business communication. Future work would benefit from a “narrative framework” that incorporates linguistic, socio-cultural, and organizational perspectives into the correlation study. The article presents the first study to investigate the correlation studies through content analysis.

Keywords
content analysis, narratives, annual reports, business performance, correlation studies

Introduction
Business performance refers to “the fulfillment of economic goals of a company” (Venkatraman & Ramanujam, 1986, p. 803), which is commonly measured by financial indicators. Nevertheless, the inherited objectivity of financial indicators obscures the evaluation and measurement of management staff, whose sensibility of the business status is often of importance to investment. This issue has prompted increasing research into narratives, that is, a natural language other than financial values in corporate discourse, corporate annual reports in particular, in attempts to mine high-value information that maps onto or predicts business performance. It has been reported that narratives in corporate annual reports are intentionally manipulated by the informant, sending important corporate internal economic signals and conveying high-value business information to the public (Leung et al., 2015; Patelli & Pedrini, 2014; Qian & Sun, 2021). These reports help to uncover current, corporate operating conditions, and reveal future potential from the management point of view (Kloptchenko et al., 2004), thereby providing information relevant to investment decisions. In addition, accounting scholars have engaged in the so-called “narrative turn” as a result of their increasing interest in the rhetorical aspects of voluntary disclosure (Beattie, 2014; Jack et al., 2013). The ways in which business performance is explicitly or implicitly expressed by narratives has given impetus to an increasing number of correlation studies, with particular attention paid to the relation of narratives in corporate annual reports to corporate performance. Narratives under such focus are thus more than an informal mode of communication that writers can either adopt or reject. Instead, it is a technical discourse, constructed in a socially situated context and therefore central to the way individuals give meaning to their lives and link their lives to society (Perkins & Blyler, 1999). An understanding of the discoursal nature of narratives will aid in examining both their complexity and pervasiveness and their connection to the development of selves and communal interests in business communication.

Traditionally, a typical analytical exercise of the correlation between annual reports’ narratives and business performance (hereinafter “correlation”) begins with an extraction

1University of Chinese Academy of Sciences, Beijing, China
2University of International Business and Economics, Beijing, China

Corresponding Author:
Yubin Qian, Department of Foreign Languages, University of Chinese Academy of Sciences, No. 19A, Yuquan Road, Shijingshan District, Beijing 100049, China.
Email: qianyubin@ucas.ac.cn
of language and financial variables from narratives and business indicators, respectively, then works toward the relationship between variables by following manually interpretative or computer-aided analytical methods, to finally arrive at a theoretical or experimental model that presents either the mapping or predictivity relationship between the aforementioned two categories of variables (Figure 1). Approaches from multiple disciplines, linguistics and business in particular, are usually involved in the analytical procedure, establishing an interdisciplinary framework to probe into how managers conceptualize their evaluation through discursive practice to be in concert with the business.

Despite an increasing number of correlation studies in the literature and the importance they have generally attached to the interdisciplinarity in the domain, there is a lack of review articles that investigate the status quo, leading to a constrained knowledge of what characterizes a correlation study, how one is conducted and where future studies may focus. These research gaps may also point to a subconscious devaluing of narratives that prevents us from recognizing the complexity and integrity of narratives in business communication. By taking a content analysis approach, here we aim to partially fill the gap with an attempt to map the structure and unwind the dynamics of extant literature. Three research questions, accordingly, are addressed:

**Research Question 1 (RQ1):** What has been the trend of correlation studies in the past decades?

**Research Question 2 (RQ2):** What are the topics most frequently discussed by correlation studies?

**Research Question (RQ3):** Which theories and methods are most frequently applied in correlation studies?

This article contributes to knowledge in several ways. First, it maps the structure and evaluates the trends of extant literature. We adapted a coding scheme that specifically distinguishes factors characterizing prior studies, thus revealing the main meta-information for presenting the research evolution. Besides, we contribute by uncovering a set of potential research themes based on a systematic interpretation of literature, which paves the way for future correlation studies. More importantly, the issues at the interface between narratives and business that we characterize here may aid professional communicators to “view their field and their discourse in new ways” (Perkins & Blyer, 1999, pp. 1–2). Narratives in professional communication are typically devaluated to informal modes of communication rather than technical discourse (Blyer, 1996), more often than not detached from the context in which the narratives are placed. Correlation study research, however, does not conceive of narratives in this limited way. Instead, correlation studies assert that narratives are a communicative discourse where the discursive practice is given meaning and defines the business community. The way in which discursivity correlates with business examined by this review article may thus improve narratives’ presence in and value to communication, empowering communicators to contextualize their narrative writing and employ the narrative turn in research.

In the remaining sections of this article, “Content Analysis” section reviews prior studies of content analysis in the business domain, and introduces the strategies to conduct the present content analysis. “Results of the Content Analysis” section begins with a description of the overall features of the correlation research and then moves on to the characterization of research topics, theories, and methods. The last section gives a summary of the present research, presents recommendations for future research, and discusses the limitations of the work.

**Content Analysis**

Content analysis has a long history of use in journalism, communication, business, and other social sciences, often defined as “a research technique for making replicable and valid inferences from texts (or other meaning matter) according to their context” (Krippendorff, 2004, p. 18). In other words, content analysis is a method of codifying the text of writing into categories based on selected criteria to derive the pattern of information that is being disclosed in the published report (Guthrie et al., 2004). In the business reporting domain, content analysis has a major application in the examination of organizational practices in managing business reports and in the scientific mapping of literature on business narratives.

With reference to content analysis of business reports, Steenkamp and Northcott (2007) suggest that “its utility for making valid inferences from texts offers considerable
potential for narratives used to communicate the outcomes of accounting activities” (p. 14). There are some useful examples from the general business and accounting literature on the use of content analysis of publicly available reports of companies. For example, there have been many studies of such reports to measure the themes and extent to which companies disclose information. These studies analyze not only mandatory components in corporate annual reports (Beattie et al., 2002) but also voluntary disclosures such as chairman’s statement (Cleary et al., 2019), forward-looking statement (Abed et al., 2016), sustainability reports (Landrum & Ohsowski, 2018), social responsibility reports (Vuontisjärvi, 2006), and reports of management discussion and analysis (Qian, 2020). Studies using content analysis in more recent years have tended to focus on the evaluation of firms’ performance through its linguistic representations in narratives (Gnanaweera & Kunori, 2018; Laskar, 2018; Raucci & Tarquinio, 2020; Solikhah et al., 2020; Uwuigbe et al., 2018; Xie et al., 2019) or on discriminating between financial restatement and accounting fraud (Azis et al., 2020; Blanc et al., 2019; Y.-J. Chen et al., 2017; Zhang et al., 2020).

In review studies of business narratives, content analysis is increasingly used to assess extant knowledge and unwind the structures of research-based in a systematic manner (Gaur & Kumar, 2018). In most cases, its use to review extant literature requires the coding of manifest content (e.g., author names, words relating to particular themes, theories or methods) that focuses on latent-content patterns (Potter & Levine-Donnerstein, 1999). In other words, objective information is first extracted by a coding scheme and then used to infer the underlying dynamics of literature present in the articles being investigated. Elshandidy et al. (2018), for example, reviewed the role of content analysis in risk reporting by coding details of articles such as study name, journal, jurisdiction and theoretical underpinnings, and arrived at an inference that the conceptualization of risk would require further refinement and clarity by researchers in this field. In addition to risk reporting studies, the application of content analysis is mostly demonstrated for reviewing the literature on reporting of corporate social responsibility (Dahlsrud, 2008; Duff, 2016) and sustainability (Asif et al., 2013; Janjua et al., 2021).

While an increasing application of content analysis in the domain of business narratives has captured the disclosure of information not well informed to academics formerly (Castilla-Polo & Ruiz-Rodriguez, 2017), from our search effort, we can only conclude that content analysis is potentially underutilized in the review of the correlation between narratives and business performance. We, therefore, in the present study, focus on the correlation with a highlight of narratives in annual reports and conduct a content analysis of articles published across domains and disciplines. This approach allows a systematic and comprehensive analysis of a broad cross-section of the literature and makes it possible to examine the research dynamics regardless of disciplinary bias. Given the approach, we first collected journal articles by the immediate three criteria: First, we used the core database of the Web of Science as the search engine because the database is well established for its high academic quality, and the emphasis placed on it means that the articles for analysis are represented in the correlation research under focus. Second, we searched for articles that were published between the years 2000 and 2018 as the time duration of years is capable to draw a trajectory characterizing research trends in the past. In addition, we retrieved articles by query sets, that is, TS=(narrati* OR text* OR discourse) AND TS=annual report AND TS=performance AND TS=(compan* OR firm* OR corporat* OR enterprise*) NOT TS=(“narrative review” OR “narrative therapy” OR textile* OR texture* OR textbook*), which hit correlation studies while excluding narrative review, narrative therapy, or any others that do not fall within the scope. According to the aforementioned criteria, a total of 78 academic articles were obtained. Based on a close reading of the title and abstract of each paper, 27 papers were selected as the targets, while the remaining 51, which do not discuss business performance or annual reports, were discarded. The criteria verified that the selected articles were legitimate representatives of the correlation studies, though the number of these articles is relatively small.

We next followed Duriau et al. (2007)’s coding scheme with some variation, according to the objectives of the review and the requirement of exhaustiveness for parent coding categories as needed (Weber, 1990). Whereas Duriau et al. (2007) addressed the methodological issues of the content analysis in the organizational studies domain, we focus on the science mapping of correlation studies by using content analysis as an analytical technique. Thus, we adapted Duriau et al. (2007) coding categories to suit the objective of our study, classifying their categories into four parent categories with each containing one or more subcategories. We also revised some of their codes, for example, content analysis technique, research design, and reliability checks to categories in line with our scope. Table 1 presents our coding scheme and the coding methods.

As shown in Table 1, the coding scheme dissects each article into a collection of 4 parent categories and 15 subcategories. The four parent categories are fundamental to address the research questions aforementioned while their subcategories provide details. Specifically, the parent category Overview consists of four subcategories, including the title of the article, source journal, author, and date of publication, with the aim to elucidate the trend of correlation research over the year. Parent categories Research topics and Research theories, respectively, direct at the prevailing topics and theories in the body of literature. The fourth parent category Research methods contains nine subcategories: Source of language data, source of financial data, number of methods, use of model, language variable, financial variable, result interpretation, performance test, and computer-aided text analysis. By clarifying these subcategories, Research methods elucidate methodological issues highlighted by prior correlation research.
The next step was to code articles to establish the reliability of the coding process (coding samples can be found in the Appendix). We confirmed the reliability of the coding process by repeating the coding 3 weeks later (Milne & Adler, 1999). Matches between the two coding results reached 90.3% agreement. Most of the discrepancies stem from the terminology inconsistency of financial variables across studies. For example, the term profitability in one research work may be paraphrased as return on assets in others, although they are the same metric. With the aid of consulting business professionals, we merged synonyms and eliminated these disagreements. After the code assignment, we characterized and interpreted articles according to their coding, and reported our findings from the perspective of overall characteristics, research topics, theories, and methods highlighted by the extant literature, as presented in the following section.

The next step was to code articles to establish the reliability of the coding process (coding samples can be found in the Appendix). We confirmed the reliability of the coding process by repeating the coding 3 weeks later (Milne & Adler, 1999). Matches between the two coding results reached 90.3% agreement. Most of the discrepancies stem from the terminology inconsistency of financial variables across studies. For example, the term profitability in one research work may be paraphrased as return on assets in others, although they are the same metric. With the aid of consulting business professionals, we merged synonyms and eliminated these disagreements. After the code assignment, we characterized and interpreted articles according to their coding, and reported our findings from the perspective of overall characteristics, research topics, theories, and methods highlighted by the extant literature, as presented in the following section.

### Results of the Content Analysis

#### Overall Characteristics

Figure 2 shows the ratio of the number of correlation research in a given year to the total 27 articles in the years 2000–2018. Looking at the changes in the number of articles during this period, it can be noted that there is a tendency toward an increase which indicates a growing interest in narratives-and-business correlation over the year. In addition, of all the 27 papers, 20, 81.46% of the total, were published between 2010 and 2018, reflecting greater popularity gained in the topic over the last decade.

We also explored the distribution of source journals across disciplines to broadly examine the features of prior research. As shown in Figure 3, the work is primarily published in journals in the fields of business, economics, and computer science. Journals in the fields of business and economics

### Table 1. The Coding Scheme Adapted From Duriau et al. (2007).

| Parent category  | Subcategory     | Coding method                                                                 |
|------------------|-----------------|-------------------------------------------------------------------------------|
| Overview         | Title           | Keep the original title of the article                                        |
|                  | Source journal  | Keep the title of the source journal                                         |
|                  | Author          | Use the surname of the first author                                          |
|                  | Date (year)     | Keep the publication year of the article                                     |
| Research topics  | Research topic  | Sort articles into “mapping” and “predictability” divisions according to the narratives-and-business relationships expressed in the article |
| Research theories| Theories        | Identify the theory the article relies on by recognizing its theoretical framework or assumptions which usually appear in sections of “Literature review” and/or “Methodology”: |
| Research methods | Source of language data | Sort language data into categories according to their source |
|                  | Source of financial data | Sort financial data into categories according to their source |
|                  | Number of methods | Sort articles into “0” and “1” divisions according to the number of methods applied by the article.  
  - 0: only one method;  
  - 1: two or more methods, for example, two quantitative methods, or a combination of quantitative and qualitative methods |
| Model            | Sort articles into “0” and “1” divisions according to the appearance of models in the article.  
  - 0: no models;  
  - 1: one or more models |
| Language variable| Sort language variables into categories according to their measurements |
| Financial variable| Sort financial variables into categories according to their measurements |
| Result interpretation | Sort articles into “0”, “1” and “2” divisions according to how research findings were interpreted by the article.  
  - 0: a qualitative analysis;  
  - 1: a quantitative analysis;  
  - 2: a mixed analysis |
| Performance test | Sort articles into “0” and “1” divisions according to the appearance of performance tests in the article.  
  - 0: no performance tests;  
  - 1: one or more performance tests |
| Computer-aided text analysis | Sort articles into “0” and “1” divisions according to the appearance of computer-aided analytical methods in the article.  
  - 0: no computer-aided analytical methods;  
  - 1: one or more computer-aided analytical methods |
include 21 publications, for example, Accounting and Business Research, Journal of Business Ethics, Sustainability, while those in the field of computer science include 6, for example, Journal of Information Science and Engineering, Neural Computing and Applications, Telematics and Informatics. The involvement of journals from diverse disciplines underscores the interdisciplinarity of correlation research, recognizing the importance of contributions made by professional communicators, such as accountants, management, and engineers from multiple disciplines.

Research Topics

Of the 27 articles, their topics were categorized into a dichotomy of mapping research and predictability research, according to the relationship between narratives and business performance. As presented in Table 2, mapping research includes 20 papers, representing 74% of the total, whereas predictability research includes 7 articles, representing 26% of the total. Thus, mapping research is far more common than predictability research.

The mapping research examines how language representations of narratives, for example, discourse, semantic features, sentiment, and word profile, mirror business performance, for example, performance targets, earning performance, financial health, governance performance, and innovativeness. Among the studies, the mapping between word profile and earning performance is the most prevalent (e.g., Bhana, 2009; Patelli & Pedrini, 2014), accounting for 35%. Bhana (2009), for example, looked into the mapping relation of words in annual reports to corporate profitability. Companies with either improving or declining performance, as he found, would use more positive words, such as “please” and “improved” than words expressing negativity, such as “adversely” and “reversed.”

The predictability research explores the potential of narratives to foresee corporate credit rating, earning performance, and financial health, from the perspectives of semantic features, sentiment, and word profile. Predicting corporate earnings performance on the basis of word profile leads the research (e.g., Balakrishnan et al., 2010; Patelli & Pedrini, 2015), accounting for 44% of the papers. Balakrishnan et al. (2010) applied text classification techniques that represented annual reports with words and term space. Based on the assumption that these representations contain value-relevant information for business earnings, the authors successfully built a statistical model that predicts corporate market performance. They concluded that the word-based classifier shows efficiency in extracting word representatives of the texts, and these key words incorporate features essential to predictive models.

Research Theories

As tabulated in Table 3, we found 10 theories, either systematic frameworks or temporary assumptions, from a wide disciplinary range across management, communication and psychology, explaining why and how narrative discourse conveys value-relevant information.

The volume of research based on impression management is the largest, comprising 5 papers, accounting for 29% of the counts. These studies hold that narratives are an effective tool to manage corporate images in the way that the wording of texts is carefully tuned through strategies, such as making assertions and defenses (Bhana, 2009; Cooper & Slack, 2015), offering concealment (Leung et al.,
Figure 3. Distribution of correlation research across journals.

**Table 2. Research Topics of Correlation Research.**

| Research topic                                      | Count (%) |
|-----------------------------------------------------|-----------|
| Mapping research                                    | 20 (74%)  |
| Discourse and performance targets;                  | 1 (5%)    |
| Semantic features and earning performance;          | 2 (10%)   |
| Sentiment and earning performance;                  | 2 (10%)   |
| Sentiment and environmental performance;             | 1 (5%)    |
| Sentiment and financial health;                     | 2 (10%)   |
| Word profile and corporate governance performance;  | 1 (5%)    |
| Word profile and earning performance;               | 7 (35%)   |
| Word profile and financial health;                  | 1 (5%)    |
| Word profile and Innovativeness                     | 3 (15%)   |
| Predictability research                             | 7 (26%)   |
| Semantic features and credit rating;                | 1 (14%)   |
| Semantic features and earning performance;          | 1 (14%)   |
| Sentiment and earning performance;                  | 1 (14%)   |
| Sentiment and financial health;                     | 1 (14%)   |
| Word profile and earning performance                 | 3 (44%)   |

**Table 3. Theories Applied by Correlation Research.**

| Theory                                      | Count (%) |
|---------------------------------------------|-----------|
| Impression management                       | 5 (29%)   |
| Agency theory                               | 3 (17%)   |
| Communicative action theory                 | 1 (6%)    |
| Dividend theory                             | 1 (6%)    |
| Management entrenchment theory              | 1 (6%)    |
| Signaling theory                            | 1 (6%)    |
| Ambidextrous internationalization           | 1 (6%)    |
| Corporate social responsibility communication| 1 (6%)   |
| Information transparency                    | 1 (6%)    |
| Responsible innovation                      | 1 (6%)    |
| Sensemaking                                 | 1 (6%)    |

2015) or assuming a self-serving manner (Cho et al., 2010; Patelli & Pedrini, 2014). By using this tool, investors’ judgment on the corporate image is likely to be influenced: Their impression of positive business performance will be
enhanced, while their perception of negative features will be obscured. Any alteration of the impression made in this way aligns with the management’s wishes, as they are under obligation to safeguard the economic interest and attract investment. Bhana (2009), in his accounts of the relation between the distribution of positive/negative words in narratives and the corporate performance, pointed out that managers use accounting narratives as a strategy to serve their own interests rather than reporting performance objectively. Overall, managers, in both improving and declining performers, desire credit for positive outcomes but the exemption for negative outcomes.

The volume of research applying the agency theory is the second largest, which includes three papers, accounting for 17% of the total. This theory explains the motivations for management’s manipulation of narratives as being derived from the pressure of shareholder value or personal interests; the management staff wish to shape their image of strong management capabilities by highlighting the positive corporate news on business while downplaying the bad news (Athanasakou & Hussainey, 2014; Michalisin, 2001; Wang & Hussainey, 2013). As reported in Michalisin (2001), annual reports’ assertions about innovativeness were manipulated in opportunistic ways to correspond with corporate innovation, and motivations for the manipulation fitted into the agency theory logic, that is, underscoring positive information helps to maintain managers’ positions, build trust among stakeholders and provide them with the chance to turn-around failing projects.

Despite the diversity of theories, the theories appear excessive and sporadic given their counts benchmarked against the 27 articles we reviewed, suggesting a lack of consensus on a theoretical perspective commonly acknowledged in the literature. Another asymmetry was found in terms of the distribution of theories. While 47%, almost half of the total, were estimated by two theories, that is, impression management and agency theory, the remaining adopted an additional eight. This asymmetry echoes the disagreement of theoretical perspectives mentioned previously. For a close inspection of how theories were applied, we looked into the application of the top two theories.

**Research Methods**

This section informs professional communicators of how to investigate the relationship between narratives and corporate performance from the perspectives of data, data analysis methods, and variables highlighted by prior research.

**Data.** The data used in the prior research include two types, language data and financial data, which, respectively, were drawn from narrative discourse and business databases. Table 4 shows the source of these two types of data. The language data come from the narratives of holistic annual reports or sections extracted from annual reports, such as Management’s Discussion and Analysis, CEO Letters, Chairperson’s Statements, Descriptions of Business, Forward-looking Performance Disclosures, Operational and Financial Reviews and President’s Letters. Given the volume of

| Table 4. Data in Correlation Research. |
|-----------------------------|-----------------------------|
| Source                      | Count (%)                  | Source                      | Count (%)                  |
| Language data               |                            | Financial data              |                            |
| Annual reports              | 18 (63)                    | Compustat                   | 8 (20)                     |
| Management’s Discussion and Analysis | 4 (13)                     | Center for Research of Security Prices | 6 (16)                     |
| Chief Executive Officer (CEO) letters | 1 (4)                     | Datastream                  | 3 (8)                      |
| Chairperson’s Statements    | 1 (4)                      | Value Line database         | 3 (8)                      |
| Descriptions of Business    | 1 (4)                      | Standard & Poor’s database  | 2 (4)                      |
| Forward-looking Performance Disclosures | 1 (4)                  | Institutional Brokers’ Estimate System | 2 (4)                     |
| Operational and Financial Reviews | 1 (4)                   | Annual report               | 2 (4)                      |
| President’s Letters         | 1 (4)                      | Company Regulatory Electronic Filing System CDs | 1 (3)                     |
|                            |                            | Reuters Global Market Data  | 1 (3)                      |
|                            |                            | Fortune 500                 | 1 (3)                      |
|                            |                            | Yahoo! Finance              | 1 (3)                      |
|                            |                            | Service 500                 | 1 (3)                      |
|                            |                            | 1988 Fortune Reputation Survey | 1 (3)                     |
|                            |                            | Thomson Financial           | 1 (3)                      |
|                            |                            | Lexis-Nexis                 | 1 (3)                      |
|                            |                            | Fortune magazine            | 1 (3)                      |
|                            |                            | MarketWatch                 | 1 (3)                      |
|                            |                            | Global Vantage database     | 1 (3)                      |


research on the complete annual report texts, that is, 18 papers, 63% of the total, the integral annual reports are the major source of language data. In terms of the financial data, they are collected from 18 databases, including Compustat and Center for Research of Security Prices (CRSP), among others. Of the 27 papers, 14 of them are based on Compustat (8 articles) and CRSP (6 articles), suggesting a major role of these two databases in providing financial data.

Data analysis methods. Table 5 summarizes the statistical methods employed in the analysis of data from perspectives of the number of methods, model, performance test, interpretation method, and computer-aided text analysis.

As for the number of methods, 24 papers, 89% of the total, used two or more methods. This finding may be attributed to the multiple research objects involved: The textual analysis relies heavily on the bag-of-words method, that is, the use of keywords collected from dictionaries, such as the Diction (Yuthas et al., 2002) or L&M (Hájek & Olej, 2013) dictionaries. The statistical analysis, however, depends on multiple mathematical approaches, tf-idf weights (Hájek et al., 2016), conditional random field models (C.-L. Chen et al., 2013), hypothesis testing (Yuthas et al., 2002), regression models (C.-L. Chen et al., 2013; Michalisin, 2001), support vector machines (Hájek & Olej, 2013), and Bayesian models (Hájek et al., 2016), among others. Michalisin (2001) combined content analysis with statistical modeling to probe into the mapping of narratives onto corporate performance on innovation. The content analysis procedures were used to identify high frequency words about firm innovativeness, whereas multiple regression models were applied to measure the extent to which changes in firm performance explained variation in the number of those key words.

In terms of models, 23 of the 27 papers have applied machine learning models, accounting for 85% of the examined publications (Table 6). The total count of models is 40, with some models combined into the same study. We categorized the models into 10 types: Regression (21 counts), support vector machine (6 counts), naive Bayes (3 counts), neural network (3 counts), decision tree (2 counts), artificial immune network memory classifier (1 count), n-gram classification technique (1 count), random condition field (1 count), k-clustering (1 count), and sequence minimum optimization algorithm (1 count) models. Among these, the regression model is the most widely used, accounting for 53% in the prior research. Michalisin (2001), for example, used ordinary least squares multiple regression to explore the correlation between the narratives in annual reports and corporate innovation. The author found that language emphasizing innovation positively correlated with corporate innovation reputation and trademark reputation. Using the same model, Cho et al. (2010) found that the positive sentiment of annual reports negatively correlated with the corporate environmental performance, an indicator of the corporate business. All 23 papers conducted performance tests by applying the correlation test to verify the accuracy of the results produced by the machine models.

Since most studies used machine learning models, the interpretation of research results is generally quantitative (22 papers, accounting for 81%). There are, however, a small number of studies that have combined qualitative and quantitative perspectives (3 papers, accounting for 11%). Only two papers, from a qualitative perspective, interpret how language strategies (Cooper & Slack, 2015) and narrative topics (Prange & Bruyaka, 2016) relate to the change in business performance. For example, Cooper and Slack (2015), based on their longitudinal data of annual reports, reported a qualitative analysis of 10 water and sewerage companies’ leakage disclosures, mapped against their performance. They found that the level of leakage disclosures markedly changed to reflect their performance against the industry regulator’s target. They established their findings by a close reading of information consistent with the impression management strategies between years of target attainment and failure of business performance.

Computer-aided text analysis is a widely used technique in prior research (25 papers, 92% of the total). Its application ranges from frequency counts (Bhana, 2009), to keyword list
generation (Balakrishnan et al., 2010) and semantic dimension mining (Yuthas et al., 2002). For example, Michalisin (2001) performed a computerized content analysis for annual reports’ emphasis on innovativeness. To this end, a single alphabetical list, along with words’ frequencies, was generated automatically by words compiled from the texts. Compared with a laborious collection of target words, the computer-aided content analysis enabled a deeper focus on interpretation and explanation.

It can be concluded from the aforementioned analysis that the methods of analyzing data are manifold, although machine learning models, performance tests, and computer-aided text analyses are relied upon heavily in the research procedures. In addition, interpreting research results from a quantitative perspective is a common choice in most studies. Since variables are the key to quantitative analysis, we present below the main variables, along with the primary results, in the literature under focus.

Variables. Generally, two types of variables, the language variable and the financial variable, were engaged in the correlation analysis. These variables correspond to the dichotomy of language data and financial data, respectively.

Table 7 presents the language variables in correlation studies. All these six language variables in Table 7 have been further divided into two categories, that is, frequency type and classification type. The frequency type includes four variables with 17 occurrences, while the classification type contains two variables with a total of 13 occurrences. The divergence between the volume and occurrence of variables between the two types indicates that the frequency type is dominant within the categories of language variables.

The frequency type of language variables refers to language units that are quantified by means of frequency counts, such as the frequency-based disclosure index, keyword frequency, semantic dimension index, and text readability. Among all the variables, semantic dimension index, that is, the quantification of semantic meanings, is the most common, with six appearances, accounting for 34% of the total. For example, Kang et al. (2018), in their approach to the measurement of tone in the narrative sections in annual reports, defined tone as a frequency difference between the numbers of words that expressed positive and negative meanings; the determinants of the tone measurement acted as measures of corporate performance.

The classification type of language variables refers to language units that are classified based on features generated from texts, including the semantic dimension classification, which classifies words into categories according to their semantic meanings, and the word classification that presents texts with word vectors. Within these two types, the semantic dimension classification is more common, appearing 8 times and accounting for 62% of the total. For example, Hájek and Olej (2013) divided the lexis of annual reports into six categories, such as negative, positive, and uncertain, based on their semantic meanings. The authors then used a feed-forward neural network and a support vector machine to explore the extent to which semantic categories mapped against investment grade.

Compared with language variables, the financial variable is more complex, with a total of 104 variables occurring 177 times. Due to the length constraints of the present article, Table 8 cannot present an exhaustive list; we have selected 15 variables that meet a minimum occurrence of 3 and classified these variables into two types: The accounting type and the marketing type, each occurring 47 and 22 times, accounting for 68% and 32% of the total, respectively. Thus, the accounting type is the dominant category given its proportion.

The accounting type of financial variables mirrors the corporate profitability through the variables of firm size, return on assets, industry, firm age, return on equity, accrual, liquidity ratio, size-adjusted return, debts, and year. Among these variables, firm size is more frequently used than others, with 12 occurrences, accounting for 26%.

The marketing type of financial variables reflects the corporate sensitivity of market conditions. It includes the variables of price-to-book ratio, dividend yield, stock returns, earnings per share, and the number of analysts following the firm. The price-to-book ratio is more common than the other variables within the same category, with eight occurrences, accounting for 36%.

Despite typological variance, the aforementioned variables have certain relationships, as revealed by statistical models. Some variables are dependent on others, constituting dependent variables, whereas others do not rely on anything else, forming independent variables. There are also control variables that are not of primary research interest and thus constitute extraneous factors whose influence is controlled or eliminated. According to Table 9: (a) The

| Frequency type                        | Counts (%) | Classification type                  | Counts (%) |
|---------------------------------------|------------|--------------------------------------|------------|
| Semantic dimension index              | 6 (34)     | Semantic dimension classification    | 8 (62)     |
| Frequency-based disclosure index      | 5 (29)     | Word classification                   | 5 (38)     |
| Key word frequency                    | 4 (25)     |                                       |            |
| Text readability                      | 2 (12)     |                                       |            |

Table 7. Language Variables in Correlation Research.
dependent variable can either be financial variables, for example, stock return, size-adjusted return, and return on assets, or language variables, for example, frequency-based disclosure index and semantic dimension index. It is thus assumed that the relationship between narratives and business performance is bidirectional; narrative discourse is reflective of business performance or vice versa. (b) The independent variables can also be financial or language variables. All the classification types of language variables, that is, semantic dimension classification and word classification, and one of the frequency type of language variables, that is, text readability, are only used as independent variables. (c) The control variables are all financial variables, among which firm size, firm age, accrual, number of analysts following the firm and year are exclusively used as this type. An exemplary study that shows the relationships between variables can be found in Kang et al. (2018), which described two models to test the relation of narratives’ overtone to business determinants. In their first model that estimated the impact of current earnings upon the proportion of positive words, the authors treated tone as the dependent variable, whereas current earnings were the independent variable, and price-to-book ratio, firm size, and firm age were control variables. In the second model, which measured the degree to which the overtone was reflective of current earnings, the roles of tone and current earnings were inverted to become the independent and dependent variables, respectively. The control variables in the second model were extended to include items such as dividend yield, accrual, and year. To understand the relations between variables is relevant to the interpretation and replication of models applied to prior studies.

**Conclusion and Discussion**

The content analysis of prior studies has successfully characterized the dynamics of correlation research from the perspectives of research trends, prevailing topics, highlighted theories and methods. We found that (a) correlation research in the past 20 years has received increasing attention with salient interdisciplinary features; (b) mapping and predictability between narratives and business

### Table 8. Financial Variables in Correlation Research.

| Accounting type          | Counts (%) | Marketing type                  | Counts (%) |
|--------------------------|------------|---------------------------------|------------|
| Firm size                | 12 (26)    | Price-to-book ratio             | 8 (36)     |
| Return on assets         | 8 (17)     | Dividend yield                  | 4 (18)     |
| Industry                 | 5 (10)     | Stock returns                   | 4 (18)     |
| Firm age                 | 4 (8)      | Earnings per share              | 3 (14)     |
| Return on equity         | 3 (9)      | Number of analysts following the firm | 3 (14) |
| Accrual                  | 3 (6)      |                                 |            |
| Liquidity ratio          | 3 (6)      |                                 |            |
| Size adjusted return     | 3 (6)      |                                 |            |
| Debts                    | 3 (6)      |                                 |            |
| Year                     | 3 (6)      |                                 |            |

### Table 9. Dependent, Independent, and Control Variables in Correlation Research.

| Dependent variable       | Count | Independent variable            | Count | Control variable | Count |
|--------------------------|-------|---------------------------------|-------|------------------|-------|
| Stock return             | 3     | Semantic dimension classification | 6     | Firm size        | 8     |
| Size adjusted return     | 3     | Price-to-book ratio             | 6     | Industry         | 5     |
| Frequency-based disclosure index | 3     | Firm size                      | 4     | Firm age         | 4     |
| Semantic dimension index | 2     | Semantic dimension index        | 3     | Number of analysts following the firm | 3     |
| Return on assets         | 2     | Word classification             | 3     | Return on assets | 3     |
| Earnings per share       | 1     | Key word frequency              | 3     | Dividend yield   | 3     |
| Key word frequency       | 1     | Return on assets                | 3     | Accrual          | 3     |
| Debt                     | 1     | Earnings per share              | 2     | Year             | 3     |
|                          |       | Text readability                | 2     | Debt             | 2     |
|                          |       | Dividend yield                  | 1     | Price-to-book ratio | 2 |
|                          |       | Frequency-based disclosure index | 1     | Liquidity ratio | 2     |
|                          |       | Liquidity ratio                 | 1     | Return on equity | 2     |
|                          |       | Return on equity                | 1     | Stock returns    | 1     |
|                          |       | Semantic dimension index        |       |                  | 1     |
performance are the prevailing topics in the literature, among which mapping research is more prevalent; (c) research theories are drawn from the fields of management, communication, and psychology, with an unbalanced distribution; the impression management and agency theory are more often applied than others; (d) the reported data include two types: Language data, mainly drawn from holistic annual reports, and financial data, primarily collected from Compustat and CRSP databases. Multiple approaches have been applied, largely depending on machine learning models, performance tests, computer-aided text analyses, and quantitative interpretation. As for variables, the frequency type of language variables and the accounting type of financial variables are the most common. The roles of these variables vary across statistical models. Generally, the dependent or independent variables are usually financial or language variables, whereas the control variables consist solely of financial variables.

Taken as a whole, our study has revealed several strategies in which most correlation studies have engaged, thus offering a direction for future research. In addition, the narratives brought forward by the review demonstrate the necessity of contextualizing narratives in business communication. We showed that narratives are neither simply an informal communication mode nor a non-technical discourse but a socially situated construct where the context endows it with meanings, and a construct that constitutes a business community. Prior research theories as mentioned in “Research Theories” section, for example, have justified the dependence of narratives on either individual, organizational, or social environment to generate their economic value, and spelled out the mechanisms of organizational control and community identity behind narratives. The research methods summarized in “Research Methods” Section 3.4, however, have confirmed the existence of discursive strategies reflective of business meaning, under the management of the business community.

Given the findings in research topics, theories, and methods (cf. “Research Topics,” “Research Theories,” and “Research Methods” sections), the emphasis on the mapping and predictivity of narratives to business performance has led to studies focusing on varied themes with a wide range of theoretical and methodological preferences. Although this diversity may be an indication of a dynamic field, motivating the use of many different tools to address the interdisciplinarity in the correlation between narratives and business performance, yet a wide range of theories and methods among a small number of studies may indicate a fragmentation of efforts and a lack of consensus. This may make it difficult for scholars to share an understanding of the overall field. One possible solution to this problem may be setting a study in a narrative framework which incorporates interdisciplinarity into correlation research, that is, by exploring together the linguistic, socio-cultural, and organizational characterizations within and beyond the narrative discourse. A linguistic analysis highlights the appearance of narratives. It focuses on the text’s surface, considering vocabulary usage, generic structure, or discourse strategies (Bhatia, 2017; Brennan & Conroy, 2013; Qian, 2020; Rutherford, 2003). While the linguistic perspective may give a systematic description of discourse, it cannot fully consider the value-related information lying behind the text. To achieve this, a socio-cultural perspective may be useful to find out whether linguistic features in narratives map onto the context where the discourse is situated, and how changes in the social and cultural factors endow narratives with diverse economic values. An organizational perspective may then complement linguistic and socio-cultural findings in that it places narratives in an organizational setting, seeing them as technical discourse and linking the discourse to corporate behaviors—such as legitimation (Rojo & van Dijk, 1997; Vaara & Tienari, 2011), and impression management (Merkl-Davies & Koller, 2012; Ogden & Clarke, 2005). Based on the unique linguistic representations of narratives and their correlation with the business, an interdisciplinary correlation study would conceptualize narrative discourse by referring to how the discourse constitutes the corporate community and whether it is managed at stages of information disclosure and corporate institutionalization (Phillips et al., 2004). The narrative framework, as described above, is an integration of perspectives from multiple disciplines, thus appealing to a joint effort in the field of professional communication. When looking at narratives from linguistic and socio-cultural perspectives, for example, engineers and technical communicators may offer computational methods, such as machine learning and big data technology, for retrieving textual features most prominent in narratives, and quantifying the links between linguistic features and informants’ cognition. When proceeding to the organizational meaning of narratives, viewpoints from the business staff and management may be valued for their claims of a narrative turn in business research.

The present article has several limitations that warrant additional research. First, the size of articles under focus is small. Although the small size guarantees the representativeness of data, the results should be interpreted with caution as they are not inclusive of all studies in the literature. Another limitation comes from the methods of the interpretation of results. The results of the present review are descriptive in nature and thus prone to the limitations of similar review articles. These limitations can be reduced in the future by incorporating data drawn from a broader range of sources with more attention paid to elaborations on the data.
### Appendix. Coding samples.

| Title                                                                 | Source journal        | Author       | Date | Research topic | Source of language data | Source of financial data | Number of methods | Model            | Language variable | Financial variable | Result interpretation | Performance test | Computer-aided text analysis |
|-----------------------------------------------------------------------|-----------------------|--------------|------|----------------|-------------------------|--------------------------|--------------------|------------------|-------------------|--------------------|-----------------------|------------------|--------------------------|
| Communicative action and corporate annual reports                     | Journal of Business Ethics | Yuthas       | 2002 | Financial performance: Semantic features- earning performance | MD&A, President letter | Compustat               | 0                  | Regression model | Key word frequency | Number of trademarks, innovativeness indicators, Industry median trademarks, relative firm performance, firm size | 2                 | 0                        | 1                |
| Validity of annual report assertions about innovativeness: an empirical investigation | Journal of Business Research | Michalisin   | 2001 | Financial performance: Word profile-Innovativeness | Entire annual reports | Lexis-Nexis, 1988 Fortune Reputation Survey, Service 500, Compustat | 1                  | Regression model | Key word frequency | Number of trademarks, innovativeness indicators, Industry median trademarks, relative firm performance, firm size | 1                 | 1                        | 1                |
Acknowledgment
The authors express gratitude to the journal editors and the anonymous reviewers for their valuable feedback on earlier versions of this article.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Yubin Qian https://orcid.org/0000-0001-7523-9647

References
Abed, S., Al-Najjar, B., & Roberts, C. (2016). Measuring annual report narratives disclosure: Empirical evidence from forward-looking information in the UK prior the financial crisis. Managerial Auditing Journal, 31(4–5), 338–361.

Asif, M., Searcy, C., Santos, P. D., & Kensah, D. (2013). A review of Dutch corporate sustainable development reports. Corporate Social Responsibility and Environmental Management, 20(6), 321–339.

Athanasakou, V., & Hussainey, K. (2014). The perceived credibility of forward-looking performance disclosures. Accounting and Business Research, 44(3), 227–259.

Azis, N. K., Marzuki, M. M., & Majid, W. Z. N. A. (2020). Fraud prevention in Malaysia: Maqasid al-Shariah perspective. Global Business and Management Research, 12(2), 104–116.

Balakrishman, R., Qiu, X. Y., & Srinivasan, P. (2010). On the predictive ability of narrative disclosures in annual reports. European Journal of Operational Research, 202(3), 789–801.

Beattie, V. (2014). Accounting narratives and the narrative turn in accounting research: Issues, theory, methodology, methods and a research framework. The British Accounting Review, 46(2), 111–134.

Beattie, V., Melnnes, B., & Fearnley, S. (2002). Through the eyes of management: A study of narrative disclosures. Centre for Business Performance.

Bhana, N. (2009). The chairman’s statements and annual reports: Are they reporting the same company performance to investors? Investment Analysts Journal, 38(70), 32–46.

Bhatia, V. K. (2017). Critical genre analysis: Investigating inter-discursive performance in professional practice. Routledge.

Blanc, R., Cho, C. H., Sopt, J., & Branco, M. C. (2019). Disclosure responses to a corruption scandal: The case of Siemens AG. Journal of Business Ethics, 156(2), 545–561.

Blyler, N. R. (1996). Narrative and research in professional communication. Journal of Business and Technical Communication, 10(3), 330–351.

Breman, N. M., & Conroy, J. P. (2013). Executive hubris: The case of a bank CEO. Accounting, Auditing & Accountability Journal, 26(2), 172–195.

Castilla-Polo, F., & Ruiz-Rodriguez, C. (2017). Content analysis within intangible assets disclosure: A structured literature review. Journal of Intellectual Capital, 18(3), 506–543.

Chen, C.-L., Liu, C.-L., Chang, Y.-C., & Tsai, H.-P. (2013). Opinion mining for relating subjective expressions and annual earnings in US financial statements. Journal of Information Science and Engineering, 29, 743–764.

Chen, Y.-J., Wu, C.-H., Chen, Y.-M., Li, H.-Y., & Chen, H.-K. (2017). Enhancement of fraud detection for narratives in annual reports. International Journal of Accounting Information Systems, 26, 32–45.

Cho, C. H., Roberts, R. W., & Patten, D. M. (2010). The language of US corporate environmental disclosure. Accounting, Organizations and Society, 35(4), 431–443.

Cleary, P., Quinn, M., & Moreno, A. (2019). Socioemotional wealth in family firms: A longitudinal content analysis of corporate disclosures. Journal of Family Business Strategy, 10(2), 119–132.

Cooper, S., & Slack, R. (2015). Reporting practice, impression management and company performance: A longitudinal and comparative analysis of water leakage disclosure. Accounting and Business Research, 45(6–7), 801–840.

Dahlsrud, A. (2008). How corporate social responsibility is defined: An analysis of 37 definitions. Corporate Social Responsibility and Environmental Management, 15(1), 1–13.

Duff, A. (2016). Corporate social responsibility reporting in professional accounting firms. The British Accounting Review, 48(1), 74–86.

Duriau, V. J., Reger, R. K., & Pfarrer, M. D. (2007). A content analysis of the content analysis literature in organization studies: Research themes, data sources, and methodological refinements. Organizational Research Methods, 10(1), 5–34.

Elshandidy, T., Shvires, P. J., Bambar, M., & Abraham, S. (2018). Risk reporting: A review of the literature and implications for future research. Journal of Accounting Literature, 40, 54–82.

Gaur, A., & Kumar, M. (2018). A systematic approach to conducting review studies: An assessment of content analysis in 25 years of IB research. Journal of World Business, 53(2), 280–289.

Gnanaweera, K., & Kunori, N. (2018). Corporate sustainability reporting: Linkage of corporate disclosure information and performance indicators. Cogent Business & Management, 5(1), 1–21.

Guthrie, J., Petty, R., Yongvanich, K., & Ricceri, F. (2004). Using content analysis as a research method to inquire into intellectual capital reporting. Journal of Intellectual Capital, 5(2), 282–293.

Hájek, P., & Olej, V. (2013, September 13–16). Evaluating sentiment in annual reports for financial distress prediction using neural networks and support vector machines [Paper presentation]. Paper presented at the International Conference on Engineering Applications of Neural Networks, Halkidiki, Greece.

Hájek, P., Olej, V., & Prochazka, O. (2016, December 8). Predicting corporate credit ratings using content analysis of annual reports—A naïve Bayesian network approach [Paper presentation]. Paper presented at the International Workshop on Enterprise Applications and Services in the Finance Industry, Frankfurt, Germany.

Jack, L., Davison, J., & Craig, R. (2013). The Routledge companion to accounting communication. Routledge.

Janjua, Z., u, A., Krishnapillai, G., & Rahman, M. (2021). A systematic literature review of rural homestays and sustainability in tourism. SAGE Open, 11(2), 1–17. https://doi.org/10.1177/21582440211007117
Kang, T., Park, D.-H., & Han, I. (2018). Beyond the numbers: The effect of 10-K tone on firms’ performance predictions using text analytics. *Telematics and Informatics, 35*(2), 370–381.

Kloptchenko, A., Eklund, T., Karlsson, J., Back, B., Vanharanta, H., & Visa, A. (2004). Combining data and text mining techniques for analysing financial reports. *Intelligent Systems in Accounting, Finance & Management: International Journal, 12*(1), 29–41.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. SAGE.

Landrum, N. E., & Ohsowski, B. (2018). Identifying worldviews about innovativeness: An empirical investigation. *Journal of Business Research, 53*(3), 151–161.

Michalisin, M. D. (2001). Validity of annual report assertions about innovativeness: An empirical investigation. *Journal of Business Research, 53*(3), 151–161.

Milne, M. J., & Adler, R. W. (1999). Exploring the reliability of social and environmental disclosures content analysis. *Accounting, Auditing & Accountability Journal, 12*(2), 237–256.

Ogden, S., & Clarke, J. (2005). Customer disclosures, impression management and the construction of legitimacy: Corporate reports in the UK privatised water industry. *Accounting, Auditing & Accountability Journal, 18*(3), 313–345.

Patelli, L., & Pedrini, M. (2014). Is the optimism in CEO’s letters to shareholders sincere? Impression management versus communicative action during the economic crisis. *Journal of Business Ethics, 124*(1), 19–34.

Patelli, L., & Pedrini, M. (2015). Is tone at the top associated with financial reporting aggressiveness? *Journal of Business Ethics, 126*(1), 3–19.

Perkins, J. M., & Blyler, N. (1999). *Narrative and professional communication*. Ablex.

Phillips, N., Lawrence, T. B., & Hardy, C. (2004). Discourse and institutions. *Academy of Management Review, 29*(4), 635–652.

Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research, 27*(3), 258–284.

Prange, C., & Bruyaka, O. (2016). Better at home, abroad, or both? How Chinese firms use ambidextrous internationalization strategies to drive innovation. *Cross Cultural & Strategic Management, 23*(2), 306–339.

Qian, Y. (2020). A critical genre analysis of MD&A discourse in corporate annual reports. *Discourse & Communication, 14*(4), 424–437.

Qian, Y., & Sun, Y. (2021). Bibliometric analysis of literature on narrative discourse in corporate annual reports (1990–2019). *Quality & Quantity*. Advance online publication. https://doi.org/10.1007/s11135-021-01135-5

Raucci, D., & Tarquino, L. (2020). Sustainability performance indicators and non-financial information reporting: Evidence from the Italian case. *Administrative Sciences, 10*(1), Article 13. https://doi.org/10.3390/admsci10010013

Rojo, L. M., & van Dijk, T. A. (1997). “There was a problem, and it was solved!”: Legitimating the expulsion of “Illegal” migrants in Spanish parliamentary discourse. *Discourse & Society, 8*(4), 523–566.

Rutherford, B. A. (2003). Obfuscation, textual complexity and the role of regulated narrative accounting disclosure in corporate governance. *Journal of Management and Governance, 7*(2), 187–210.

Solikhan, B., Wahyudin, A., & Rahmayanti, A. A. W. (2020). The extent of intellectual capital disclosure and corporate governance mechanism to increase market value. *The Journal of Asian Finance, Economics, and Business, 7*(10), 119–128.

Steenkamp, N., & Northcott, D. (2007). Content analysis in accounting research: The practical challenges. *Australian Accounting Review, 17*(43), 12–25.

Uwuigbe, U., Teddy, O., Uwuigbe, O. R., Emmanuel, O., Asiriuwa, O., Eyitomi, G. A., & Taiwo, O. S. (2018). Sustainability reporting and firm performance: A bi-directional approach. *Academy of Strategic Management Journal, 17*(3), 1–16.

Vaara, E., & Tienari, J. (2011). On the narrative construction of multinational corporations: An antenarrative analysis of legitimation and resistance in a cross-border merger. *Organization Science, 22*(2), 370–390.

Venkatraman, N., & Ramanujam, V. (1986). Measurement of business performance in strategy research: A comparison of approaches. *Academy of Management Review, 11*(4), 801–814.

Vuontisjärvi, T. (2006). Corporate social reporting in the European context and human resource disclosures: An analysis of Finnish companies. *Journal of Business Ethics, 69*(4), 331–354.

Wang, M., & Hussainey, K. (2013). Voluntary forward-looking statements driven by corporate governance and their value relevance. *Journal of Accounting and Public Policy, 32*(3), 26–49.

Weber, R. P. (1990). *Basic content analysis*. SAGE.

Xie, X., Huo, J., & Zou, H. (2019). Green process innovation, green product innovation, and corporate financial performance: A content analysis method. *Journal of Business Research, 101*, 697–706.

Yuthas, K., Rogers, R., & Dillard, J. F. (2002). Communicative action and corporate annual reports. *Journal of Business Ethics, 41*(1), 141–157.

Zhang, L., Shan, Y. G., & Chang, M. (2020). Can CSR disclosure protect firm reputation during financial restatements? *Journal of Business Ethics*. Advance online publication. https://doi.org/10.1007/s10551-020-04527-z