

Harmony Between The Information Contained In The Text And Figures Of Brazilian Companies' Annual Reports

Abstract

The purpose of this study is to analyze the harmony between information provided in the narrative sections of annual reports and the corporate financial performance. The sample consisted of 120 companies listed on BM&FBovespa (Brazilian Stock Exchange) in 2009 -- 60 with greater positive variation and 60 with greater negative variation in net accounting profit. Keywords relating to three central topics were selected: profitability, growth and management. Through content analysis, the meaning of these keywords in the reports was evaluated (quantitative positive, qualitative positive, quantitative negative and qualitative negative). Once the frequencies were obtained, two logistic regressions were performed to compare the text to the numbers, one on the raw frequency of words and another by a weighting of terms. The results indicate that, in the information linked to the topic of profitability, the text is harmonious with the numbers. In the information relating to growth, harmony is partial. In addition, when it comes to information linked to management, there is conflict between the narrative sections and corporate performance. Finally, it was found that, the more subjective the information, the greater the conflict.

Keywords: Annual Reports; Cognitive Dissonance; Content Analysis. "Acknowledgements to FAPES – Fundação de Apoio à Ciência e Tecnologia do Espírito Santo for funding this research."

Acknowledgements to FAPES – Fundação de Apoio à Ciência e Tecnologia do Espírito Santo for funding this research.

Marcelo Sanches Pagliarussi

PhD. in Materials Science and Engineering from Universidade Federal de São Carlos and Professor at the University of São Paulo - Ribeirão Preto. **Contact:** Av. Bandeirantes 3900, Monte Alegre, Ribeirão Preto, SP, CEP: 14040-905.
E-mail: marcelosp@usp.br

Thiago Neiva Guimarães

Master's Degree in Accounting at Fucape Business School and Assistant Professor at Faculdades Unidas do Norte de Minas. **Contact:** Rua Lírio Brant, 787, Melo, Montes Claros, MG, CEP: 39401-063.
E-mail: thiaguimaraes@gmail.com

Tiago Alves Ferreira

Master's student in Accounting at Fucape Business School. **Contact:** Av. Fernando Ferrari, 1358, Boa Vista, Vitória, ES, CEP: 29075-505.
E-mail: tiagorev@gmail.com

1. Introduction

Traditionally, the performance assessment of an organization is based on an analysis of numerical information, such as accounting figures. However, textual information may also be useful to assess the potential future performance of an organization (Clatworthy & Jones, 2003). Such information contextualizes the accounting figures through a narrative that describes the activities and actions companies develop in their target markets, future plans, projects, product lines, cost reduction efforts, sale of assets, etc.

Recently, a substantial body of research was developed with a focus on the analysis of textual information disclosed by the organizations and its usefulness in decision making in a capital market context. The studies involve various phenomena, such as the readability of corporate annual reports according to performance (Li, 2008); the use of corporate reputation as management tactics in annual reports (Vargas, Almeida & Maria Junior, 2014; Ribeiro, Pagliarussi, Silva & Silva Junior, 2011; Tessarolo, Pagliarussi & da Luz, 2010; da Luz, Pagliarussi, Teixeira & Baptista, 2009); the tone of financial texts (Loughran & McDonald, 2011); and the disclosure of strategic background information in annual reports (Pagliarussi & Liberato, 2011).

The focus of this study is the harmony between the textual and numerical information presented in the reports of companies listed in Brazil. The relevance of questioning the harmony between the two sets of information is derived from the greater discretion organizational managers are allowed in producing the narrative (Balata & Breton, 2005). When developing a similar study for use with Canadian companies, the authors observed a moderate level of divergence between the text and the numbers, but only analyzed the letters of the presidents in 30 companies.

In another study on the subject, Smith (1998) showed by means of an experiment that the existence of conflicting messages influenced the decisions the research subjects made. Beynon, Clatworthy and Jones (2004), sought to verify the correspondence of the text with the figures, developed a data-mining method to try to classify, by means of textual analysis of the letters of company presidents, profitable and unprofitable companies in the United Kingdom, and succeeded in applying the method with high predictive accuracy and a high level of correct classification.

In order to contribute to the understanding of the relationship between the text and the numerical information presented in the reports of the listed companies, the present study intends to analyze whether or not there is harmony in the messages conveyed in the narrative sections of the annual reports, in relation to accounting figures concerning the company's performance. The theoretical framework of this study is based on the propositions of Milgrom (1981), concerning the "persuasion game" in which a party interested in providing information attempts to influence the decisions of others; and on Grossman and Hart (1980), who claim that the party who has relevant information has nothing to gain by hiding or lying about the information, since it is possible for the receiving party to verify whether or not the information is true and accurate.

This study contributes to the literature on the consistency between textual information and numerical information through the analysis of different thematic groupings, based on the previous studies by Vargas, Almeida and Maria Junior (2014); Ribeiro, Pagliarussi, Silva and Silva Junior (2011), Tessarolo, Pagliarussi and Da Luz (2010), and Da Luz, Pagliarussi, Teixeira and Baptista (2009). In addition to an analysis of the topics of Profitability, Growth and Management, the entire narrative section contained in the annual reports was reviewed, and not just the presidents' letters, as in the studies by Balata and Breton (2005).

In order to achieve the proposed objectives, the annual reports of all companies listed in 2009 were collected from the BM&FBovespa website and 120 were selected with greater variation in net accounting profit, comprised of 60 with greater positive variation and 60 with greater negative variation. The content analysis technique was applied to the 120 reports, by analyzing the frequency of keywords. These keywords were preselected in correspondence with the topics of Profitability, Growth and Management; and their frequencies were categorized, taking into consideration their nature as either quantitative or qualitative,

and their polarity, either positive or negative. From the frequency distribution, the variables for application of the logistic regression technique were obtained.

Two logistic regressions were performed, one on the raw frequency of the occurrence of each keyword in the texts, and the other by weighting terms developed by Manning and Schütze (2009). The results showed that, for the data referring to profitability, the texts are consistent with the numbers. As for the data referring to Growth and Management, the two sections demonstrate conflict.

2. Theoretical Framework And Development Of Hypotheses

2.1. Brazilian legal provisions about annual management reports

In Brazil, according to Art.133 of Law No. 6.404 (1976), all enterprises classified as corporations shall publish annually the management report on corporate business and the main administrative factors of the previous year, along with the financial statements, among other reports. In the narrative sections of the reports, companies can explain, in great detail, the events that occurred during the financial year, as well as future projects.

Recommendation No. 15 of the Brazilian Securities and Exchange Commission (Comissão de Valores Mobiliários - CVM) of December 28, 1987 (CVM, 1987), also defines the topics and minimal content of the narrative sections, such as the description of the businesses, products and services; comments on the general economic situation; human resources, number of employees, investments in training, security funds and social plans; investments made; research and development; new products and services; environmental protection; administrative reforms; investments in subsidiaries and affiliates; shareholders' rights and market data; prospects and plans for the current year and future years.

Yet, according to this Recommendation, p.2: "the disclosure of useful, reliable and detailed information that enables the knowledge of the company and of its objectives and policies is an essential right of the shareholder. Another excerpt says that the text of the reports shall:

[...] be written with a simplicity of language to be accessible to the greatest possible readership, and adjectives and phrases such as, "excellent result", "great performance", "low debt", "excellent prospects", unless supported by comparative data or facts, shall be avoided (CVM, 1987, P.2).

In addition, Art. 25 of the CVM Instruction No. 308 (CVM, 1999) says that the independent auditor shall verify, "...if the information and accounting and financial analysis submitted in the entity's management report are consistent with the financial statements audited."

The importance of the reports as a source of information for all stakeholders of the company is stressed. For instance, most of the information mentioned by the analysts is taken from the narrative sections (Rogers and Grant, 1997).

2.2. Previous Studies

In the case of the narrative sections, studies in this area can be categorized into two types: readability and content analysis (Beynon, Clatworthy & Jones, 2004). Studies in readability seek to assess the difficulty to understand the texts while, in content analysis, there is an array of methods, including the interpretation of the content of the text by the producer of the work, with word counts or phrases of pre-determined topics with the aid of software (Beynon, Clatworthy & Jones, 2004).

Among the studies involving content analysis, which is the focus of this article, Balata and Breton (2005) analyzed letters of the presidents of 30 Canadian companies between 1993 and 1998. They ques-

tioned the harmony between the information contained in the presidents' letters and the accounting figures, checking for conflicting information that tended to produce a divergent effect. The authors created an index of the level of optimism contained in the letters and financial statements, analyzing three central topics: Profitability, Growth and Management. Keywords were determined for each topic in order to evaluate the optimism in the textual part and the financial indicators associated with topics to assess the optimism contained in the financial statements. As a result, they found a moderate level of inconsistency between narrative sections and accounting figures.

Another study that found conflict between the information contained in the narrative sections and in the financial statements was developed by Smith (1998). The author submitted 60 students with degrees in accounting, in pursuit of MBAs, to an experiment. In this trial, students had to make decisions based on financial information and, subsequently, they also had to make decisions based on the narratives of failed and successful businesses. As a result, the author concluded there was evidence of significant conflicts between the messages conveyed by accounting information and by the narratives. For Smith, the narrative can be misleading in decision making.

Clatworthy and Jones (2003) analyzed 100 letters of presidents from firms listed in the United Kingdom, 50 with positive performance variation and 50 with negative variation, in order to know whether or not firms with increasing or declining performance report good and bad news in different ways. Overall, the results suggested that firms prefer to emphasize the positive aspects of their performance, accepting the credit for good news while blaming the external environment for the bad news.

Tessarolo, Pagliarussi and Da Luz (2010) used content analysis to investigate whether or not there was a pattern of causal attribution in the narrative sections of the annual reports and how the economic environment and the company's performance would influence such a pattern. The results suggest that, in Brazil, annual reports are marked by significant levels of assignments on their own account, that is, they try to create a positive corporate image, even when receiving a negative performance in a favorable external environment. The authors also state that it became clear that the managers intended to disclose the narrative sections of the annual report in an opportunistic way. In a study on a similar topic, Da Luz et al. (2009) found that the companies listed on BM&FBovespa make use of impression management tactics in their annual reports, regardless of their adherence to different levels of corporate governance. As Ribeiro et al. (2011) observed, in a case study of a pulp and paper company, the textual information presented sought to prioritize and emphasize the institutional identity of the company; to intensify, give details and strengthen their attitudes, rather than inform in a neutral way their achievements or plans.

Finally, Vargas, Almeida and Maria Júnior (2014) observed that the existence of print management in the reports of the companies listed on BM&FBovespa. The authors found that there is difference in the number of lines and words present in the sections discussing profit and loss. The same study noted that, when dealing with negative results (losses), the frequency of terms that express this result are reduced, and there were increases in an emphasis of terms such as the EBITDA and the gross margin. Evidence of attempts to divert the user's attention from the results to other indicators was also observed.

2.3. Development of Hypotheses

The results of studies using narrative sections are still contradictory; there is no consensus regarding the consistency of the texts with performance or, if there is, it is a skewed presentation of information aimed at managing the corporate reputation.

In Smith's view (1998), reputation management prevails in the documents companies disclosed. Based on the agency and signaling theories, the authors state that companies seek to build a narrative in order to obscure the perception of the audience when the financial performance is poor. Abrahamson and Amir (1996), in turn, found evidence of consistency between the two sets of information, both in the numbers and in the narrative.

As Spence theorizes (1973), in a market environment, there are characteristics that are permanent and others that are changeable. The author argues that, in a context of asymmetric information, these changeable characteristics are transmitted by means of signs, capable of changing beliefs and the impressions of individuals who receive them. That is, if there is information asymmetry, individuals and/or companies can provide stakeholders with the information necessary in order to eliminate and/or minimize the asymmetry. However, the information must be reliable, where the interests of those who provide it are a potential source of bias.

In what Milgrom (1981) calls “the game of persuasion”, those who provide information to an individual in an attempt to influence his decisions should reveal all of the information they have about the “product,” since it is assumed that the information can be checked and the “buyer” can detect whether or not there is information that has been hidden. However, if there is information that cannot be checked, the “seller” can pass it on as works best for him. This idea is confirmed by Grossman and Hart (1980) when they say that it is of no advantage for the “seller” to withhold information in an attempt to deceive. The authors suggest that, assuming that defrauding or lying is illegal, and the cost of the transmission of information is irrelevant, then the “seller” will voluntarily disclose all of the information he/she has about the “product.”

In the context of the present study, it is assumed that if information is verifiable, it will be transmitted in a transparent way. Otherwise, there may be reputation management attempts through the selection of the sets of information to be transmitted about the company’s situation.

Therefore, in order to meet the objective of the research, which is to determine whether or not there is harmony in the messages transmitted through the text, represented by the narrative sections of the annual reports, as compared to the company’s performance, represented by the variation in its net profit; we assume the following hypotheses:

H1a: *The better the company’s performance, the greater the frequency of quantitative positive information relating to the topics of Profitability, Growth and Management;*

H1b: *The worse the company’s performance, the greater the frequency of quantitative negative information on the topics Profitability, Growth and Management;*

It is expected that the company’s performance has a directly proportional relationship with the frequency of quantitative information in its report. Therefore, it is understood that, as for the quantitative information, the text of the reports will be transparent and consistent with the figures.

H2a: *Regardless of the company’s performance, it is expected that there will be no difference in the frequency of qualitative information, either positive or negative, for the topics of Profitability, Growth and Management.*

Because qualitative information is more subjective and not directly associated with figures, it is expected that there will be a lack of harmony between the qualitative and quantitative information in the case of companies with a negative growth in profit, because it is expected that, even if they are not performing well, they will submit positive and negative qualitative information with a frequency similar to companies with a positive growth in profits.

3. Method

3.1. Data and Sample

The selection in the initial sample used all the companies listed on BM&FBovespa whose 2009 annual reports were published. The year 2009 was selected for two reasons: first, the companies' performance was not yet seriously affected by the global crisis that began late 2008, which would probably affect the narrative constructed to justify the performance of organizations. Second, from 2010 on, due to the adoption of international accounting standards in Brazil, companies listed on BM&FBovespa began to disclose their narrative information using the Reference Form, with a very different structure from the annual information required until then. To ensure comparability with previous studies, the year 2009 proved to be the most appropriate for conducting the study.

The collection of reports was performed directly on the BM&FBovespa website and included an initial sample of 389 reports. According to Clatworthy and Jones (2001), annual reports with a total content of less than 300 words are not necessarily representative, and will therefore be excluded from consideration in the sample. Based on this, 25 reports in this condition were excluded, leaving a total of 364 reports.

In the next stage, data referring to the companies' variation in net accounting results (profit/loss) were extracted from the Economática database, in order to determine those with greater positive and negative variations. Reports without the necessary information to calculate the variation of the net accounting result (net income for the years of 2008 and 2009), a total of five, were excluded, ultimately leaving 359 reports for analysis.

Then, 120 reports were selected, 60 with greater positive variation of the net accounting result and 60 with greater negative variation of the net accounting result. These two extreme groups were chosen to capture the companies that were, at that moment, going through more radical changes in financial performance (Clatworthy & Jones, 2003, 2006).

3.2. Content Analysis

The content analysis technique was utilized to extract information from the narrative section of annual reports. Content analysis attempts to utilize the content of messages to highlight indicators that imply a different reality other than the message alone (Bardin, 2004).

To evaluate the types of messages that were transmitted in the narrative sections of these reports, groups of keywords that refer to the topics of profitability, growth and management of companies were determined, following a study by Balata and Breton (2005). The keywords used for profitability were cost, expense, billing, gain, profit, margin, loss, price, prejudice, revenue, profitability, result, return and value. For growth, we used addition, alienation, acquisition, growth, demand, development, consolidation, incorporation, investment, market, offer, sales and volume. For management, the keywords used were competitiveness, control, differentiation, strategy, focus, management, initiative, innovation, leadership, productivity, restructuring and reorganization.

To locate the keywords in the reports, the software ATLAS.ti was used for content analysis, which, in addition to locating excerpts in the texts where the keywords appear, generates a report with a word count, both with a total of the keywords in each report and a total of each of the keywords in all the reports. In this process, only the roots of words were entered into the software, so that derivative words related to keywords such as, "rent," which can be related to the terms "profitability," "cost effectiveness" or "profitable," could be captured.

Next, a report was generated with the frequency of keywords, totaling 17,348 appearances of some of the words mentioned above. Another report was generated from the software with excerpts from annual reports in which each of the keywords would appear, resulting in a text editor file of 1,405 pages.

The authors manually assessed this file to determine what context each of the words would appear in throughout the text (quantitative positive, qualitative positive, negative quantitative or qualitative negative) for each of the core subjects (profitability, growth and management). Words that were not inserted into any of the contexts mentioned, that is, those that appeared in headings and subheadings, those that had a neutral meaning in the text or those with similar root but that would not correspond to the keyword (for instance, “rent”: current) were discarded, leaving a total of 7,173 valid words.

To illustrate how this evaluation occurred, the following presents some excerpts where there appear keywords related to the proposed topics in each of the meanings evaluated. Regarding the topic of Profitability, there is an example of quantitative positive Profitability: “In 2009, the company’s revenue was R\$ 299,195, 131% higher than that reported in 2008”. In this excerpt, the keyword “billing” is related to the highlighted words “131% more,” that is, a positive notation in a quantitative form. An example of a qualitative positive Profitability is “the net financial expenditures obtained a reduction due to the improvement of the working capital...” In this excerpt, the word expenditure is associated with “reduction”; therefore, it appears in a positive context and in a qualitative form.

To illustrate the classification quantitative negative Profitability, there is “The gross profit of R\$85.1 million in 2008 dropped 9.5% for R\$77.1 million in 2009.” In this excerpt, the keyword “profit” appears as associated with the word “dropped”, that is, in a negative context and followed by numerical values, which gives a quantitative form to the excerpt. The classification of qualitative negative profitability is apparent in: “...and the high costs of raw material, such as, sulfur, ammonia and oil products led to a considerable reduction in the company’s profitability margins in 2009.” In this excerpt, the term Profitability appears associated with the word “reduction,” that is, a negative context for the company, in a qualitative form.

Related to the topic Growth, here is an example with a quantitative positive context: “... and sales that totaled 2.9 million tons, a record volume, which represents an increase of 16.7%, in comparison to the previous year.” In this excerpt, the keywords “sales” and “volume” appear to be associated with the word “increase,” that is, a positive meaning and in a quantitative form. For qualitative positive Growth, an example would be, “...our presence in all regions of the country and exposure to different income segments allows us to continue growing in a solid and meaningful way in the coming years.” In this citation, the term “growing” appears in a qualitative positive form.

In the context of negative quantitative Growth, there is: “In 2009, domestic sales totaled 3.2 tons, representing a 22% retraction compared to the same period in 2008, a result of the slowdown in the registered demand, mainly, in the first half of the year.” In this excerpt the term “sales” is associated to the word “retraction,” which accompanies numerical values, that is, in the quantitative negative form. The topic Growth, with a qualitative negative form: “The investments in 2009 were substantially reduced...” In this excerpt, the keyword “investments” is associated to the word “reduced”, which shows a negative note on the part of the company and in a qualitative form.

For the topic Management, an example in quantitative positive context is: “We kept Kaizen projects in various production sectors, by making appropriate processes and reducing time operations, with productivity improvement in results between 50% and 200%.” In this passage, the term “productivity” appears associated with the word “improvement” and with the percentages “50% and 200%,” which provide a positive and a quantitative form to the phrase. The topic of Management, with a qualitative positive form appears in the following excerpt: “...creating value for all audiences with which it relates, in addition to consolidating its leadership in the markets into which it operates.”

In this example, the word “leadership” appears to be associated with “consolidating,” that is, a positive context for the company and in a qualitative form.

This is an example of the topic of Management in a negative quantitative context: "... fall in productivity, especially in soybeans, from 10.5%, for crop-year 2008/09 compared to the previous crop-year." In this sentence, the term productivity appears to be associated with the word "fall" and to values of "10.5%." Therefore, it is in a quantitative negative context. As for the topic Management, in a qualitative negative context, there is "...the consequent fall of the productive activity forced Usiminas to make adjustments to its workforce." In this case, the keyword "productive" is associated to the word "fall," giving the excerpt a negative meaning and qualitative form.

3.3 Hypothesis Testing

After tabulating the frequencies and the meanings in which all keywords appeared, both by topic and by report, logistic regressions were performed, which is an appropriate technique when the dependent variable is binary (nonmetric) and the independent variables are metric (Hair, Babin, Anderson & Tatham, 2007). Thus, the two interest groups are coded as binary variables with values 0 and 1. In this paper, we designated companies with negative variation in profits as 0 and companies with positive variation in profits as 1.

To perform the regressions, we have, on the one hand, the binary variable referring to the variation of profit (0 for firms with negative variation in profits and 1 for companies with positive variation in profits) and, on the other hand, the coefficients relating to the characteristics of keywords (quantitative and qualitative positive, quantitative or qualitative negative) are extracted from the narrative sections. The model representing regression is as follows:

$$y = \beta_0 + \beta_1 \text{Rentpq} + \beta_2 \text{Rentpql} + \beta_3 \text{Rentnegqt} + \beta_4 \text{Rentnegql} + \beta_5 \text{Crescpqt} + \beta_6 \text{Crescpql} + \beta_7 \text{Crescnegqt} + \beta_8 \text{Crescnegql} + \beta_9 \text{Gestpqt} + \beta_{10} \text{Gestpql} + \beta_{11} \text{Gestnegqt} + \beta_{12} \text{Gestnegql} + \varepsilon$$

Where:

- y:** 0 = companies with negative variation in profits; 1 = companies with positive variation in profits;
- Rentpq:** positive quantitative Profitability = sum of the frequencies of all keywords associated with the topic Profitability which appear in the report text with positive and quantitative meaning;
- Rentpnq:** positive qualitative Profitability = sum of the frequencies of all keywords associated with the topic Profitability which appear in the report text with positive and not quantitative meaning;
- Rentnegq:** negative quantitative Profitability = sum of the frequencies of all keywords associated with the topic Profitability which appear in the report text with negative and quantitative meaning;
- Rentnegnq:** negative qualitative Profitability = sum of the frequencies of all keywords associated with the topic Profitability which appear in the report text with negative and not quantitative meaning;
- Crescpq:** positive quantitative Growth = sum of the frequencies of all keywords associated with the topic Growth which appear in the report text with positive and quantitative meaning;
- Crescpnq:** positive qualitative Growth = sum of the frequencies of all keywords associated with the topic Growth which appear in the report text with positive and not quantitative meaning;

- Crescnegql** negative quantitative Growth = sum of the frequencies of all keywords associated with the topic Growth which appear in the report text with negative and quantitative meaning;
- Crescnegq:** negative qualitative Growth = sum of the frequencies of all keywords associated with the topic Growth which appear in the report text with negative and not quantitative meaning;
- Gestpq:** positive quantitative Management = sum of the frequencies of all keywords associated with the topic Management which appear in the report text with positive and quantitative meaning;
- Gestpnq:** positive qualitative Management = sum of the frequencies of all keywords associated with the topic Management which appear in the report text with positive and quantitative meaning;
- Gestnegq:** negative quantitative Management = sum of the frequencies of all keywords associated with the topic Management which appear in the report text with negative and quantitative meaning;
- Gestnegq:** negative qualitative Management = sum of the frequencies of all keywords associated with the topic Management which appear in the report text with negative and not quantitative meaning;

Two logistic regressions were performed. The first was by use of the raw frequency of the keywords, with the coefficients obtained by the sum of the absolute number of occurrences of each keyword in the reports.

The second regression calculated its coefficients by the weighting of the terms. The weighting reflects the importance of a term in a document, based on the number of occurrences of this term (Manning, Raghavan & Schütze, 2008; Loughran & McDonald, 2011). For this, we assign to each term in a document (report) a weight according to its number of occurrences in the text. The simplest approach would be to assign the weight to be equal to the number of occurrences of term t in document d , since a greater frequency suggests a greater importance of the word in the text (Manning, Raghavan & Schütze, 2008).

However, Manning and Schütze (1999) state that this importance needs to be mitigated. An example cited by the authors is where a term that appears three times does not mean that it is three times more important than another that appears only once. Another important detail that is commonly used to define the importance of a term is to know how many reports it appears in, in relation to the total number of reports. Defined by these criteria, we have the following formula for weighting terms developed by Manning and Schütze (1999).

$$W(i, j) = \begin{cases} (1 + \ln(\text{tf}_{i,j})) \ln \frac{N}{\text{df}_i}, & \text{se } \text{tf}_{i,j} \geq 1 \\ 0 & , \text{ se } \text{tf}_{i,j} = 0 \end{cases}$$

Where:

$W_{(i,j)}$: term weighting i , in the report j ;

$\text{tf}_{i,j}$: keyword frequency i , in the report j ;

N : total number of reports;

df_i : number of reports containing the keywords (i).

4. Data Analysis

The following is the research data analysis, descriptive statistics and logistic regressions with their respective statistical classifications.

Table 1 shows descriptive statistics for the sum of the raw frequency of keywords.

Table 1

Descriptive Statistics by Raw Frequency

Variables	Total			Negative variation in profit			Positive variation in profit		
	Obs.	Avg.	SD.	Obs.	Avg.	SD.	Obs.	Avg.	SD.
Rentpqt	120	5.167	5.402	60	3.917	5.099	60	6.417	5.447
Rentpql	120	9.292	9.467	60	7.717	8.567	60	10.867	10.115
Rentnegqt	120	3.392	4.235	60	4.367	4.723	60	2.417	3.456
Rentnegql	120	2.742	3.781	60	3.517	4.048	60	1.967	3.349
Crescpqt	120	6.675	8.385	60	5.733	8.273	60	7.617	8.459
Crescpql	120	18.092	19.527	60	15.567	19.727	60	20.617	19.155
Crescnegqt	120	1.375	2.521	60	1.483	2.734	60	1.267	2.306
Crescnegql	120	2.083	3.034	60	2.45	2.861	60	1.717	3.179
Gestpqt	120	0.025	0.157	60	0.033	0.181	60	0.167	0.129
Gestpql	120	10.775	15.749	60	9.350	15.491	60	12.200	16.005
Gestnegqt	120	0.075	0.735	60	0.150	1.039	60	0	0
Gestnegql	120	0.833	0.306	60	0.117	0.372	60	0.050	0.220

Legend: Rentpqt (positive quantitative Profitability); Rentpql (positive qualitative Profitability); Rentnegqt (negative quantitative Profitability); Rentnegql (negative qualitative Profitability); Crescpqt (positive quantitative Growth); Crescpql (positive qualitative Growth); Crescnegqt (negative quantitative Growth); Crescnegql (negative qualitative Growth); Gestpqt (positive quantitative Management); Gestpql (positive qualitative Management); Gestnegql (negative qualitative Management).

Source: Prepared by the authors

According to Table 1, it was observed that the frequency of non-quantitative information appears to be higher than the quantitative information in the reports. For example, on the topic of growth, we have a mean of 18.10 for the non-quantitative data and a mean of 6.68 for the quantitative data. Regarding the positive or negative information, the positive data appear with a higher frequency. For example, on the topic of profitability, the positive information has a mean of 9.29 and the negative information for the same topic has a mean of 2.74.

It was apparent that the reports generally provided more positive information about the companies. In the present study, of all the keywords mentioned, 83.7 % were in a positive direction. If only the companies with negative variations in profit are considered, then the percentage of keywords with a positive direction remained at 77.8%.

After calculating descriptive statistics, Table 2 shows the results of the logical regression using the raw frequencies of the keywords in the reports.

Table 2

Logistic Regression by Raw Frequency

Variable	Coef.	Standard Error	Z	P>z
Rentpqt	0.1872	0.0777	2.41	0.016**
Rentpql	0.1642	0.0695	2.36	0.018**
Rentnegqt	-0.2697	0.1102	-2.45	0.014**
Rentnegql	-0.2721	0.1216	-2.24	0.025**
Crescpqt	-0.0279	0.0507	-0.55	0.581
Crescpql	0.0315	0.0341	0.93	0.355
Crescnegqt	0.1414	0.1547	0.91	0.361
Crescnegql	0.0309	0.1439	0.21	0.830
Gestpqt	-1.9688	1.7561	-1.12	0.262
Gestpql	-0.0824	0.0404	-2.04	0.041**
Gestnegql	-1.9356	1.1630	-1.66	0.096*
N° Obs = 120	LR chi2(11) = 42.74	P>chi2 = 0.0000	Pseudo R2 = 0.2569	

Legend: Rentpqt (positive quantitative Profitability); Rentpql (positive qualitative Profitability); Rentnegqt (negative quantitative Profitability); Rentnegql (negative qualitative Profitability); Crescpqt (positive quantitative Growth); Crescpql (positive qualitative Growth); Crescnegqt (negative quantitative Growth); Crescnegql (negative qualitative Growth); Gestpqt (positive quantitative Management); Gestpql (positive qualitative Management); Gestnegql (negative qualitative Management).

Significance levels: *** = 1%; ** = 5%; * = 10%

Source: Prepared by the authors

In the regression analysis, the software discarded the variable Gestnegqt (Management negative quantitative) due to a lack of variability in the data and, for this reason, the regression was run again without this variable. The results are shown in Table 2, although without the discarded variable, showing that the model was statistically significant, because $P > \chi^2 = 0.0000$. In addition, the variables Rentpqt, Rentpql, Rentnegqt, Rentnegql and Gestpql were significant at 5%. The variable Gestnegql was also significant, although at 10%.

In analyzing the results, hypothesis 1a predicts that higher frequencies of positive quantitative information about the issues of Profitability, Growth and Management increased the probability of a company belonging to the group of companies with positive variation in profits. Only the variable Rentpqt had a statistically significant coefficient, with the sign as expected, indicating the confirmation of this hypothesis. Yet, the variables Crescpqt and Gestpqt did not have significantly significant coefficients, rejecting the hypothesis.

Hypothesis 1b -- the higher the frequency of negative quantitative keywords about the issues of Profitability, Growth and Management, the greater the likelihood of the observations belonging to the group 0 (companies with negative change in profit) -- showed evidence that confirmed the hypothesis for the variable Rentnegqt, which had a statistically significant coefficient and expected sign. The other variables (Crescnegqt and Gestnegqt) lacked statistically significant coefficients and, therefore, provided evidence for the rejection of the hypotheses.

In hypothesis 2a, it was expected that the frequencies would be similar. Some variables had coefficients suggesting evidence confirming the hypothesis, while others indicated that the hypothesis was rejected. With this hypothesis, the analogy is reversed, that is, the variables that had significant coefficients (Rentnegql, Gestnegql, Rentnegql and Gestnegql) generate evidence that the hypothesis was rejected because they indicated that the frequencies are not similar, as suggested by the hypothesis.

For the variable Rentpql, the higher frequencies of keywords imply a greater probability of an observation belonging to the group of companies with positive growth in profits, that is, companies with good performance presented more information in their texts that was qualitative positive on the subject of Profitability, demonstrating consistency between text and numbers.

For the other variables (Gestnegql, Rentnegql and Gestnegql), higher frequencies increased the chances that a specific observation would be part of companies without a positive variation in profit. With the variables Rentnegql and Gestnegql, there is no conflict of information because companies in worse financial situations tend to have more negative information. With the variable Gestpql, there is conflict in the messages transmitted, and higher frequencies of Management information with a positive direction and qualitative form, increase the chances of happening in companies with negative growth in profit.

For the variables Crescpql and Crescnegql, the coefficients confirmed the hypothesis; because they were not statistically significant and suggested similar frequencies for companies with either positive or negative performances. That is, it is not possible with these results to identify the frequency of keywords that belong to a group of companies in a particular observation. In this case, there is conflict; it was expected that companies with a positive variation in profit would evidence information that is more positive and companies with poor performance would evidence information that is more negative, due to their financial situation.

In addition to logistic regression, in which the variables were constructed through the raw frequency of keywords, another regression was conducted, with a weighting of terms, assigning different weights according to the number of occurrences of the term, as shown in Table 3.

Table 3
Logistic Regression by Weighting of Terms

Variable	No. Obs = 120	LR chi2(11) = 48.18	P>chi2 = 0.0000	Pseudo R2 = 0.2896
	Coef.	Standard Error	z	P>z
Rentpqt	0.0926	0.0788	1.18	0.240
Rentpql	0.2768	0.0974	2.84	0.004***
Rentnegqt	-0.3654	0.1127	-3.24	0.001***
Rentnegql	-0.2316	0.0880	-2.63	0.008***
Crescpqt	-0.0036	0.0759	-0.05	0.962
Crescpql	0.2045	0.0904	2.26	0.024**
Crescnegqt	0.2928	0.1369	2.14	0.032**
Crescnegql	-0.0663	0.1038	-0.64	0.523
Gestpqt	0.2306	0.4287	0.54	0.591
Gestpql	-0.1557	0.0707	-2.20	0.028**
Gestnegql	-0.3655	0.3061	-1.19	0.232

Legend: Rentpqt (positive quantitative Profitability); Rentpql (positive qualitative Profitability); Rentnegqt (negative quantitative Profitability); Rentnegql (negative qualitative Profitability); Crescpqt (positive quantitative Growth); Crescpql (positive qualitative Growth); Crescnegqt (negative quantitative Growth); Crescnegql (negative qualitative Growth); Gestpqt (positive quantitative Management); Gestpql (positive qualitative Management); Gestnegql (negative qualitative Management).

Significance levels: *** = 1%; ** = 5%; * = 10%

Source: Prepared by the authors

During regression analysis by raw frequencies, the variable Gestnegqt (Management negative quantitative) was not considered by the software in the regression using the weighting of terms due to a lack of variability in the data, and the regression was run again without this variable. The results obtained show that the model was statistically significant, although $P>chi2 = 0.0000$. In addition, the variables Rentpqt, Rentpql, Rentnegqt were significant at 1%. The variables Crescpql, Crescnegqt and Gestpql were also significant, although at 5%.

Analyzing the results, hypothesis 1a expects that the higher the frequency of positive quantitative keywords related to Profitability, Growth and Management increases the chances of an observation belonging to the group of companies with a positive variation in profits. None of the variables had statistically significant coefficients, indicating that the hypothesis was rejected.

Hypothesis 1b -- the higher the frequency of keywords relating to topics Profitability, Growth and Management, in a quantitative negative way, the greater the likelihood of the observation belonging to the Group of companies with a negative variation in profit. Statistically significant coefficients were found for the variables *Rentnegqt* and *Crescnegqt*. For the variable *Rentnegqt*, the sign of the coefficient was as expected, confirming the hypothesis. As for the variable *Crescnegqt*, the sign of the coefficient was contrary to expectations, indicating that the hypothesis was rejected. This result suggests that the higher the frequency of keywords relative to the themes of Growth, in a negative quantitative form, the greater the probability that an observation belongs to the group of companies with poor performance.

The variable *Gestnegqt* did not have a statistically significant coefficient and this indicates that the hypothesis for this variable was rejected. Thus, we cannot say that the higher frequencies of keywords in this direction will increase the chances of the observations being part of the group of companies with negative variation in profit.

Hypothesis 2a, which suggested that average frequencies are similar to the qualitative information of the coefficients of the variables *Rentpql*, *Crescpql*, *Gestpql* and *Rentnegql* that were statistically significant, indicated that the hypothesis was rejected. For the variables *Rentpql* and *Crescpql*, the greater the frequency of keywords, the greater the probability of an observation belonging to the group of companies with a positive variation in profits, that is, companies with good performance supplying more textual information in a positive qualitative form about the themes of Profitability and Growth, demonstrating consistency between narrative sections and financial performance.

With the variable *Gestpql*, as in the previous regression, there is conflict in the messages transmitted, because the higher the frequency of information on Management, with a positive direction and a non-quantitative form, the greater the likelihood of the observation belonging to companies with a negative variation in profit. In the analysis of the variable *Rentnegnq*, higher frequencies indicate a greater probability that a specific observation belongs to companies with negative variations in profits, that is, companies with poor performance supplying more information on Profitability, with a negative direction, and a qualitative form. This demonstrates consistency between textual and financial information.

For the variables *Crescnegnq* and *Gestnegnq*, the coefficients indicate the confirmation of the hypothesis, and indicate similar averages between companies with good performance and those with bad performance. Thus, it is not possible, based on these results, to identify by the frequency of keywords which group of companies a particular observation belongs to. In this case, there is conflict between the two types of information, because it was expected that companies with a negative variation in profit, due to their lower performance, would provide more information with a negative direction on the topics of Growth and Management.

Comparing the results of the two regressions performed, in the case of statistical classification, the two models are very similar, showing no major differences. The model for raw frequencies had 79.17% of correct classifications, while the model using weighted terms had 80% of correct classifications.

The Pseudo R2 also increased with the weighting of terms from 0.2569 in the regression rate to 0.2896, showing a greater explanatory power for this model.

As for the statistically significant variables, the raw frequency model had five variables significant at 5% and a significant variable at 10%, for a total of six variables. Yet the model for weighted terms had six significant variables, three at 1% level and three at 5%.

However, if only the variables are observed that were expected to be significant according to the hypothesis, the model for raw frequencies had two variables at 5%, and the model for weighted terms had one variable that was significant at 1% and another at 5%, in which the sign of the coefficient went against expectations.

Thus, in general, we can say that, although the model using weighted terms obtained better results in statistical classification and in the pseudo R², with respect to the statistically significant variables expected, the two models were similar. This shows that the weighting of the terms suggested by Manning and Schütze (1999), which mitigates the importance of a specified term that appears frequently in the report, does not show any significant improvements compared to the simple absolute count of keywords.

Table 4 provides a summary of the hypotheses, showing which variables were statistically significant and which confirmed or rejected the hypotheses in each of the regressions.

Table 4
Summary of Results

Hypotheses	Variables	Freq. Raw	Summary of Results	Freq. Weighted	Summary of Results
1a	Rentpqt	**	confirmed	-	rejected
	Crescpqt	-	rejected	-	rejected
	Gestpqt	-	rejected	-	rejected
1b	Rentnegqt	**	confirmed	***	confirmed
	Cresnegqt	-	rejected	(**)	rejected
	Gestnegqt	-	rejected	-	rejected
2a	Rentpql	**	rejected	***	rejected
	Crescpql	-	confirmed	**	rejected
	Gestpql	**	rejected	**	rejected
	Rentnegql	**	rejected	***	rejected
	Crescnegql	-	confirmed	-	confirmed
	Gestnegql	*	rejected	-	confirmed

Legend: Rentpqt (positive quantitative Profitability); Rentpql (positive qualitative Profitability); Rentnegqt (negative quantitative Profitability); Rentnegql (negative qualitative Profitability); Crescpqt (positive quantitative Growth); Crescpql (positive qualitative Growth); Crescnegqt (negative quantitative Growth); Crescnegql (negative qualitative Growth); Gestpqt (positive quantitative Management); Gestpql (positive qualitative Management); Gestnegql (negative qualitative Management).

Significance levels: *** = 1%; ** = 5%; * = 10%

Note: (**) variable with a sign of the coefficient contrary to expectations.

Source: Prepared by the authors

Analyzing the results for the proposed hypotheses, we have: Hypothesis 1a, where the better the performance of the company, the greater the frequency of positive quantitative information on the topics of Profitability, Growth and Management, presented confirmation of the hypothesis only for the variable Rentpqt in the raw frequency regression. With the other variables and the regression using weighted terms, the variables did not have statistically significant coefficients; hence, the hypothesis was rejected.

These results suggest that the message transmitted was consistent only with respect to the topic of Profitability and, for data referring to this topic, there was no attempt to employ impression management by the companies. This result can be explained by the fact that, when it comes to information that leads to profitability, the keywords are, in most cases, directly associated with the numbers, making it difficult to pass a false picture of the company's situation, as highlighted in the following excerpt: "In 2009, there was a 78% increase in gross profits compared to 2008, reaching R\$ 70,388.00 in 2009, compared to R\$ 39,592.00 in 2008" (Excerpt taken from the report of a company with positive growth in profit).

For the topics of Growth and Management, the results suggest similar frequencies of information for companies, independent of performance, indicating that the companies with negative variation in profits transmit positive information relating to these topics as much as those with positive variation in profits.

As for Hypothesis 1b, which envisaged that the worse the performance of the largest company in the frequency of negative quantitative information concerning the topics of Profitability, Growth and Management, only the variable *Rentnegqt* had a statistically significant coefficient in the two regressions, which implies that there were indications that the hypothesis was confirmed, suggesting consistency between the text and performance when it comes to this variable.

The variable *Crescnegqt* had a statistically significant coefficient in the weighted regression, but with a sign opposite to expectations, that is, companies with a greater frequency of this type of information are more likely to belong to the group of companies with positive variation in profits. Companies with a greater frequency of this type of information are more likely to belong to the group of companies with positive variation in profit. This suggests that the companies with a negative variation in profits avoid mentioning this topic to avoid emphasizing their financial situation.

Nevertheless, the variable *Gestnegqt* provided reasons for rejecting the hypothesis, since this variable was discarded in the regressions because of a lack of variability in the data. This result indicates conflict in transmitted messages, for it suggests evidence that the frequencies of this type of information is similar in companies, independent of performance, and different than expected in the proposed hypothesis.

Based on these results, it can be seen that, as in Hypothesis 1a, only the information about the topic Profitability showed consistency between narrative sections and numbers.

Hypothesis 2a, which predicted that, independently of a company's performance, there will be no difference in the frequency of positive and negative qualitative information on the topics of Profitability, Growth and Management, presented evidence to reject the hypothesis for some variables and to confirm it for others. The analogy in this case is reversed, since the hypothesis is rejected when the coefficients of the variables are statistically significant, indicating that the frequencies are not similar.

Regarding the variable *Rentpql*, the coefficient was statistically significant in both regressions and always with a positive sign, indicating a higher frequency for companies with positive variation in profit, which indicates consistency between the text and the figures and evidences that the hypothesis was rejected. The variable *Crescpql* only had a statistically significant coefficient in the regression using weighting, with a positive sign, just like the previous variable.

The variable *Crescpql* had a statistically significant coefficient in both regressions. It is noteworthy, however, that the variable coefficient is negative in both regressions. This indicates that the average frequency of keywords on the topic of Management in a qualitative positive way was higher in companies with negative variation in profit, indicating inconsistency between narrative and financial sections, and indicating that the hypothesis was rejected. The *Rentnegql* variable also had a statistically significant coefficient in both regressions with a negative sign, indicating a greater frequency for companies with negative growth in profits, signifying that the 2a hypothesis was rejected. With this result, the text is consistent with the company's financial performance.

The variable *Crescnegql* did not have a statistically significant coefficient in any regression, confirming the hypothesis of the similarity between the narratives of companies, regardless of performance, and thus demonstrating a conflict between the information provided, because it was expected that companies with negative variation in profits would have a greater frequency of information with a negative meaning. The *Rentnegql* variable had a statistically significant coefficient in the regression by raw frequency, with the negative sign indicating a greater frequency of information for companies with negative variation in profits, revealing that the hypothesis was rejected, and indicating consistency between statements.

Analyzing this hypothesis, we found that, even when it comes to qualitative information, the information related mainly to profitability and growth was consistent with company performance, while for the information related to management, we consider that, the more information was provided, mainly in the positive direction, the greater the chances of being a company with a negative change in profit.

This result can be explained by the fact that, unlike the keywords relating to profitability and a few words related to growth, the words that refer to management are not directly linked to the "numbers", making it possible to transmit a message about the company's situation that is not really the truth, as high-

lighted in the following passages: “the nationwide service offering gives the company the necessary scale to compete on equal terms with the major international competitors present in Brazil” (excerpt taken from the report of a company with positive growth in profit).

In an excerpt from another company, we note: “The adjustments made and all actions taken in 2009 served to make the most productive and mainly more competitive, which qualifies us to seek new customers, markets with higher added value and product development of high complexity” (excerpt taken from a company report with negative variation in profit).

In these passages, it is noted that, regardless of the real situation of the company, when it comes to management, the text can also be equally positive and optimistic for companies, no matter what their performance is, as stated Hildebrandt and Snyder (1981) suggesting that regardless of the company's financial performance, the language in the annual report will be predominantly positive.

It was also found that some companies ignore the recommendation made by CVM to avoid adjectives and phrases such as “excellent result”, “great performance”, “excellent prospects”, when those are not supported by facts or comparative data. Some examples of the violation of this recommendation can be seen in the excerpts: “[...] that defines the energy service conditions in isolated systems and brings excellent prospects for restoring the financial health of companies” (excerpt from the report of a company with negative change in profit).

5. Conclusion

This study aimed to analyze the harmony in the messages disclosed in the narrative sections of the annual reports of companies listed on BM&FBovespa, compared to the financial performance presented by these companies. To achieve the goal, we selected 120 companies listed on BM&FBovespa in 2009: 60 with the highest positive variation and 60 more with a negative variation in net accounting profit. Content analysis was applied to the reports of these companies, analyzing the frequency, nature and polarity of a set of predetermined keywords related to the themes of Profitability, Growth and Management.

The logistic regression analysis was applied to assess the association between variables obtained from the content analysis with the performance measure selected. Overall, the results indicate that, when dealing with information regarding Profitability, there is harmony between text and numbers. As for the information linked to growth, harmony between text and numbers decreased. And in the case of information on Management, disharmony is more evident. It is concluded, therefore, that there is not total consistency between the text and the numbers, that is, there is a conflict of information, but partially.

The keywords related to management, in most cases, besides being more subjective, give an idea of future perspective, that is, even if the performance of a particular company has been bad, its report tries to bring into focus a discussion of future prospects. It was noticed that the more subjective the nature of the information, the greater the disharmony with the company's figures. In addition, the reports generally transmit predominantly positive information about the companies, minimizing any negative situation with mitigating arguments.

The results allow us to conclude that there is a moderate degree of disharmony between the textual and numerical information disclosed by the companies in the sample, and that the disharmony observed in some of them is associated with the attempt to manage the corporate reputation with regard to performance and, especially, regarding the management of the company, with attempts to mitigate the negative focus on previous performance.

This study contributes to the understanding of the aspects that make up the quality of financial and non-financial information presented in the narrative sections of the annual reports of the companies listed. The results support the conclusion recognized by previous studies, that discretion in the presentation of textual information in the narrative sections may be excessive. That is, there is a high potential for bias in the presentation of qualitative information on the company's performance and management. This pre-

sentation can also be done opportunistically. Consequently, the definition of formal mechanisms for the disclosure and auditing of textual information in the company's annual reports should be considered as costs and benefits.

In the context of a growing importance of the Brazilian capital market, it is expected that this paper will contribute to support the activities of regulators and market analysts. The Brazilian capital market, characterized by weak protection of the property rights of minority shareholders, a high concentration of ownership and inefficient market from an informational standpoint, stimulates the needs of understanding the role of textual information presented by the companies listed in this context.

Among the limitations present in the article, the year 2009 was the year following the crisis, and the results of some companies may still be impacted by this situation, as well as the fact that the companies have not been separated by sectors. Future studies can analyze longer periods of time, with a larger number of companies, or use other measures to assess the performance of companies, such as ROE (Return on Equity) or ROA (Return on assets), for example, to verify the results of this study. There are also opportunities for future research on the same topic, using published reports after the full adoption of IFRS (International Financial Reporting Standards).

6. References

- Abrahamson, E., & Amir, E. (1996). The information content of the president's letter to shareholders. *Journal of Business Finance e Accounting*, 23(8), pp.1157-1182, doi: <http://dx.doi.org/10.1111/j.1468-5957.1996.tb01163.x>
- Balata, P., & Breton, G. (2005). Narratives vs numbers in the annual report: are they giving the same message to the investors?. *Review of Accounting and Finance*, 4(2), pp.5-14, doi: <http://dx.doi.org/10.1108/eb043421>
- Bardin, L. (2004). *Análise de conteúdo*. Lisboa: Edições 70.
- Beynon, M. J., Clatworthy, M. A., & Jones, M. J. (2004). The prediction of profitability using accounting narratives: a variable-precision rough set approach. *Intelligent Systems in Accounting, Finance and Management*, 12(4), pp. 227-242, doi: <http://dx.doi.org/10.1002/isaf.256>
- Breton, G., & Taffler, R. J. (2001). Accounting information and analyst stock recommendation decisions: a content analysis approach. *Accounting and business research*, 31(2), pp. 91-101, doi: <http://dx.doi.org/10.1080/00014788.2001.9729604>
- Clatworthy, M., & Jones, M. J. (2001). The effect of thematic structure on the variability of annual report readability. *Accounting, Auditing e Accountability Journal*, 14(3), pp. 311-326, doi: <http://dx.doi.org/10.1108/09513570110399890>
- _____. (2003). Financial reporting of good news and bad news: evidence from accounting narratives. *Accounting and business research*, 33(3), pp. 171-185, doi: <http://dx.doi.org/10.1080/00014788.2003.9729645>
- _____. (2006). Differential patterns of textual characteristics and company performance in the chairman's statement. *Accounting, Auditing e Accountability Journal*, 19(4), pp. 493-511, doi: <http://dx.doi.org/10.1108/09513570610679100>
- Cunha, R. K. C. D. (2009). Silva, César Augusto Tibúrcio. Análise da Facilidade de Leitura das Demonstrações Contábeis das Empresas Brasileiras: Uma Investigação do Gerenciamento de Impressões Nas Narrativas Contábeis. *Anais eletrônicos Congresso USP Controladoria e Contabilidade*, São Paulo, SP, Brasil, 9.
- CVM Comissão de Valores Mobiliários. (1987). *Parecer de Orientação CVM nº 15*, de 28/12/1987. Retrieved on Abril 25, 2013 from <http://www.cvm.gov.br/>
- CVM Comissão de Valores Mobiliários. (1999). *Nota explicativa à Instrução CVM nº. 308, de 14 de maio de 1999*. Retrieved on July 8, 2013 from <http://www.cvm.gov.br/>.

- Da Luz, A. T. M., Pagliarussi, M. S., Teixeira, A. M. C., & Baptista, É. C. (2009). Uma Investigação sobre a Relação entre Atribuições em Causa Própria e Governança Corporativa. *BBR-Brazilian Business Review*, 6(2), pp. 191-208, doi: <http://dx.doi.org/10.15728/bbr.2009.6.2.5>
- Grossman, S. J., & Hart, O. D. (1980). Disclosure laws and takeover bids. *The Journal of Finance*, 35(2), pp. 323-334, doi: <http://dx.doi.org/10.2307/2327390>
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., e Tatham, R. L. (2007). *Análise multivariada de dados*. Bookman.
- Hildebrandt, H. W., & Snyder, R. D. (1981). The Pollyanna hypothesis in business writing: initial results, suggestions for research. *Journal of Business Communication*, 18(1), pp. 5-15, doi: <http://dx.doi.org/10.1177/002194368101800102>
- Hyland, K. (1998). Exploring corporate rhetoric: metadiscourse in the CEO's letter. *Journal of Business Communication*, 35(2), pp. 224-244, doi: <http://dx.doi.org/10.1177/002194369803500203>
- Lei n.º 6.404, de 15 de dezembro de 1976. *Lei das Sociedades por Ações*. Lei 6.404, de 15/12/1976. Recuperado em 23 agosto, 2010 de http://www.planalto.gov.br/ccivil_03/Leis/L6404consol.htm.
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and economics*, 45(2), pp. 221-247, doi: <http://dx.doi.org/10.1016/j.jacceco.2008.02.003>
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), pp. 35-65, doi: <http://dx.doi.org/10.1111/j.1540-6261.2010.01625.x>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval* (Vol. 1, p. 6). Cambridge: Cambridge university press.
- Manning, C. D., & Schütze H. (1999). *Foundations of statistical natural language processing* (Ed.). MIT press.
- Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12(2), pp. 380-391 doi: <http://dx.doi.org/10.2307/3003562>
- Pagliarussi, M.S, & Liberato, G. B. (2011) Disclosure de estratégia em relatórios anuais: uma análise de dimensões culturais, de Sistema legal e de governança corporativa em empresas de quatro países. *RAM. Revista de Administração Mackenzie (Online)*, 12(4), pp. 155-181.
- Ribeiro, J. F., Pagliarussi, M. S., Silva, A. R. L., & Silva Junior, A. (2011). Escolhas linguísticas, custos políticos e gerenciamento da imagem corporativa: o caso da Aracruz Celulose S. A. *Cadernos EBAPE. BR*, 9(4), pp. 1136-1157.
- Rogers, R. K., & Grant, J. (1997). Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis*, 3, pp. 17-31, doi: <http://dx.doi.org/10.2469/dig.v28.n2.254>
- Smith, M. (1998). Conflicting messages in annual reports, *Accountability and Performance*, 4(2), pp. 43-60,
- Smith, M., & Taffler, R. J. (2000). The chairman's statement-a content analysis of discretionary narrative disclosures. *Accounting, Auditing e Accountability Journal*, 13(5), pp. 624-647, <http://dx.doi.org/10.1108/09513570010353738>
- Spence, M. (1973). Job market signaling. *The quarterly journal of Economics*, 87(3), pp. 355-374, doi: <http://dx.doi.org/10.2307/1882010>
- Tessarolo, I. F., Pagliarussi, M. S. & da Luz, A. T. M. (2010). The justification of organizational performance in annual report narratives. *BAR-Brazilian Administration Review*, 7(2), pp. 198-212, doi:<http://dx.doi.org/10.1590/S1807-76922010000200006>
- Vargas, L. H. F., Almeida, J. E. F., & Júnior, E. M. (2014). Lucro e prejuízo sob a perspectiva da administração: como os resultados são apresentados nos relatórios de administração? *Revista de Educação e Pesquisa em Contabilidade (REPeC)*, 8(4), pp. 351-368.