Changes in Disclosure Tone after an Incident

Quality and industry spill-over effects

Master Thesis

Authors: Fanny Rydberg and Josefin Dolietis
Supervisor: Emmeli Runesson

ABSTRACT

Background/Purpose: The focus of this report is how the concerned company BP and Oil and Gas (O&G) industry responded to the Deepwater Horizon oil spill in their Sustainability reports. This study examines changes in tone after the incident in order to detect changes in language. The purpose is to examine BP's and the industry's response in order to see if there is a spill-over effect.

Research Design: This is done by analysing the tone in Sustainability reports by conducting a pre-post study design where the computerized text analysis tool DICTION is used to analyse four master variables for tone; activity, optimism, certainty and realism. We use the control industry Food and Beverages (F&B) in order to control that the change is due to the Deepwater Horizon oil spill and nothing else that effects the whole economy. The sample consist of 51 O&G companies and 44 F&B companies. Resulting in 390 Sustainability reports.

Findings: Certainty in Sustainability reports for Oil and Gas companies decreases after the Deepwater Horizon oil spill. Meaning that the industry responds with a less certain language after the incident. Since we expect certainty in the language to decrease, this implies that the sustainability disclosure quality decreases. The other three master variables do not generate statistically significant results. The results are compared to the control industry F&B. We found that industry, distance, age and text length affect the master variables.

Contributions: This study contributes to previous research regarding incidents impact on sustainability disclosure quantity and quality. Moreover, this study contributes by using one of many methods to study disclosure tone. Also, this study contributes by examining if there is a spill-over effect to the industry.

Conclusion/Implications: The tone changed as expected for one of the master variables, certainty. This means that we found small evidence that the industry responds with a less certain language in their Sustainability reports after the Deepwater Horizon oil spill. Indicating that the disclosure quality decreases. Moreover, since the quantity of disclosures increases, there is indication of information overload or corporate greenwashing. The fact that not all master variables changes as expected could be due to impression management.

Keywords: Disclosure tone, disclosure quality, spill-over effect, industry response, text analysis, activity, optimism, certainty, realism, impression management, information overload, greenwashing

TABLE OF CONTENT

1. INTRODUCTION

This section presents the background, problem area with motivation and purpose of this study. Contributions and implications are also be presented. Finally, limitations are discussed.

The aim of this study is to examine if the tone in Sustainability reports is affected by a disaster in the industry where the incident occurred. The chosen incident for this study is the Deepwater Horizon oil spill in the Gulf of Mexico in 2010. We examine if the incident affected the tone on Sustainability reports for companies operating in the Oil and Gas (O&G) industry. More specifically, the aim is to examine if there is a spill-over effect to other companies in the industry or if only the company involved is affected. The change in language is analysed by looking at the tone in Sustainability reports before and after the incident. Examining the tone in the language is one way to evaluate the disclosure quality according to Beattie (2014), Melloni (2015) and Hummel and Schlick (2016). In this study we will look at tone in order to study the industry's response and change in disclosure quality. Hummel and Schlick (2016) suggest that "good sustainability performers" have high-quality sustainability disclosure in order to present their good work to their stakeholders. Moreover, that "bad sustainability performers" have disclosures of low quality since they want to hide their performance and keep their legitimacy intact (Hummel & Schlick, 2016). Also, disclosures should include both good and bad news in order to be classified as high quality (Ammad, Muhammad Bilal & Muhammad, 2018). Building on this idea, if the master variables from DICTION changes as expected based on prior literature, this gives an indication of the industry's response and the disclosure quality. If the tone changes in the expected direction, this implies that disclosure quality increases since the text capture the underlying economics. However, this is only applicable if the change in tone that we expect is positive for the quality. Consequently, if we expect a change in tone that is harmful for the quality, then if the master variable changes in the expected direction, that implies that the disclosure quality decreased. If the tone does not change as expected, one explanation could be impression management.

Expected tone = Actual tone ≈ Disclosure quality

Expected tone that implies higher-quality disclosures = Actual tone ≈ Disclosure quality increased

Expected tone that implies lower-quality disclosures = Actual tone ≈ Disclosure quality decreased

Expected tone ≠ Actual tone ~ Impression management could be one explanation

We also use a control industry not related to the O&G industry, the Food and Beverages (F&B) industry is selected since this industry is less sensitive to economic downturns and because there are available reports in GRI's database. No change for F&B companies is expected because this industry is presumed to be less sensitive to market movements since the industry is well-diversified according to Verbeek (2017). If the changes only occur for the O&G industry, then we can presume that the outcome is due to the Deepwater Horizon oil spill. This was tested using a pre-post study design and a computerized text analysis.

1.1 Background

Sustainability is becoming more and more important and the growing global focus has led to increased focus on companies' sustainability work (Chen, Hung & Wang, 2018). There is also an increasing trend towards requiring companies to disclose their corporate social responsibility (CSR) activities, which are of interest to investors, regulators and stakeholders (Chen et al., 2018). There is a need for voluntary disclosures and sustainability disclosures since there are information asymmetries (Healy & Palepu, 2001; Verrecchia, 2001). One way for companies to communicate their CSR activities is through Sustainability reports, which enables them to present non-financial information about their environmental, social and economic activities and goals (GRI, 2019). Sustainability disclosures is a way for firms to show their awareness surrounding social and environmental issues (Friedman & Miles, 2001; Michelon, 2011), as well as use it to communicate their social engagement to their shareholders and stakeholders (Michelon, 2011). Krasodomska and Cho (2017) argue that the concept of CSR has changed over the years from solely being a part of companies' PR to become an element in companies' long-term strategy since stakeholders becomes more interested in CSR disclosures. However, CSR contributions to sustainability development is not a legal requirement, rather, companies voluntarily present the information (Gamerschlag, Möller & Verbeeten, 2011). Types of non-financial reports are Sustainability reports, Triple bottom line reports, CSR reports and Integrated reports (GRI, 2019). Even if companies are pressured to present CSR disclosures, there are no clear guidelines or regulations for how to do it. The Global Reporting Initiative (GRI) have tried to bridge the gap by issuing the Sustainability Reporting Standards that companies can follow and if they do, the reports can be accessible in GRI's database (GRI, 2019). The GRI framework is used by an increasing number of firms when voluntarily reporting environmental activities and following the framework makes the disclosures more specified according to Raiborn, Butler and Massoud (2011).

A few studies have investigated sustainability disclosures related to major events (Cho, 2009; Vourvachis, Woodward, Woodward & Patten, 2016). Vourvachis et al. (2016) focused on disclosure quantity and examined CSR disclosures in relation to catastrophic accidents in the airline industry, showing that for most companies, the quantity of CSR disclosures increased considerably afterwards with attempts of legitimization. Cho (2009) examined environmental disclosure practice and decisions in one of the largest integrated O&G companies in the world, where the focus was on two oil incidents in the year of 1999 and 2001. They studied legitimation strategies for the companies employed for these incidents (Cho, 2009). We use tone as a proxy for sustainability disclosure quality. Instead of only focusing on the company responsible for the incident, the industry's response will be analysed in order to examine if there was a spill-over effect. Therefore, 390 Sustainability reports from 95 companies in two different industries will be analysed.

1.1.1 Deepwater Horizon Oil Spill

The chosen event for this study is the explosion and sinking of the Deepwater Horizon oil rig, which led to the worst oil spill in the history of the U.S (Smithsonian, 2019). The disaster occurred in the Gulf of Mexico on the 20th of April in 2010, where 11 people were killed and the rig leaked oil into the gulf for 87 days (Smithsonian, 2019). The Deepwater Horizon oil rig was located 50 miles off the coast of Louisiana and leaked roughly 200 million gallons of crude oil into the Gulf of Mexico, which were broadcasted on live camera most of the 3 months when the oil was leaking (Lee & Blanchard, 2012). The oil rig was owned by BP plc., a company involved in everything from finding the oil, producing it, to manufacturing raw materials and fuel (BP, 2019). After the disaster, BP were fined $65 bn for compensating the oil spill (The Guardian, 2018). The seafood industry and the energy industry in Louisiana were affected by this disaster, and due to the tremendous attention to the O&G industry, public interest emerged for the moratorium that were put in place on new drilling (Lee & Blanchard, 2012). The event was chosen for this study since it had major impact on the environment and resulted in that 11 employees lost their lives. The Deepwater Horizon oil spill was an environmental disaster, the worst oil spill in the U.S and it is therefore of interest to study the industry's response to this disaster.

1.2 Problem Area & Purpose

Sustainability reports and sustainability disclosures are unclear since there is a lack of guidelines and regulations (Gamerschlag et al., 2011). This voluntary non-financial information is not as monitored and audited as financial information in Annual reports (GRI, 2019). Thus, GRI have tried to bridge the gap by issuing the Sustainability Reporting Standards which are guidelines that

companies can follow if they want to (GRI, 2019). However, there is no requirement to follow GRI's standards. Also, GRI does not account for the quality of the reports, only that the companies comply with the guidelines (GRI, 2013). Thus, even if companies are following the GRI framework, this is not a guarantee that the disclosures are of high-quality. Since there is a lack of information on how to present the voluntary sustainability disclosures, they might be biased and of low quality. Disclosures are according to Chiu and Wang (2015) normally expressed in general with little concrete information and in vague terms, which leads to poor overall disclosure quality. It is therefore interesting to highlight the qualitative aspects of how firms disclose after a major incident. It is also of interest to see if the quantity increased, since that may give an indication that the quality is worse. This could be explained by information overload or greenwashing. Moreover, if the tone does not change as expected for each master variable, it could be explained by impression management.

Impression management occur when managers use information asymmetries to their advantage by highlighting positive outcomes and avoiding negative outcomes (Merkl-Davies et al., 2011). Managers may act opportunistically and change the tone of the disclosures in order to make them more appealing to their stakeholders. Guillamon-Saorin, Isidro and Marques (2017) argue that tone is an impression management disclosure technique which are used when there is a need to create a positive image of corporate results. Cho, Roberts and Patten (2010) argue that disclosure bias in the language and verbal tone are a consequence of an impression management strategy. Companies might after a disaster only disclose more in order to regain legitimacy. Melloni, Caglio and Perego (2017) argue that an increase in disclosure quantity could be used as smokescreen to hide low firm performance by having low disclosure quality and information overload. Also, greenwashing could explain if the tone does not change as expected. Corporate greenwashing exists when companies present positive environmental information out of context that could be misleading to individuals who lack background information about the company (Lyon & Maxwell, 2011). Or when a company voluntarily discloses environmental information that puts the company in a positive situation and not information that might be negative for the company (Lyon & Maxwell, 2011). According to Leuz and Wysocki (2016) there is a lack of research and evidence about effects of events. They state that disclosure outcomes meaning reporting quality and events need more focus (Leuz & Wysocki, 2016). Consequently, the importance of disclosure quality needs to be highlighted, not only the levels and quantity of social and environmental disclosures. The disclosure quality is important to address in order to avoid information overload and greenwashing. It is also

of importance to study the quality of the disclosures since incidents may drive companies to hide information or create information overload (Lang & Stice-Lawrence, 2015).

Even though it is important to study the quality of the disclosed information, it is according to Helfaya and Whittington (2019) a complex task and there is no consensus about the best way of measuring it. We have decided to examine the industry's response and the disclosure quality by looking at tone which is one way of doing it (Beattie, 2014; Melloni, 2015; Hummel & Schlick, 2016). However, there are many possible ways of doing it and there is no consensus about the best way of doing it. We have made a simplification and assumed that if the master variables change in the expected direction, this indicates that the disclosure quality increases. Our main purpose is to see if the tone and sustainability disclosure quality changes as expected after the Deepwater Horizon oil spill. We examine how the tone for the concerned company BP change and also, if there is a spill-over effect to other companies in the O&G industry. Moreover, we use the control industry F&B. This study sheds light on disclosure tone as well as spill-over effects. Our research question is as follows:

Did tone change in Sustainability reports after the Deepwater Horizon Oil spill for the concerned company and industry?

1.3 Research Design

The industry's response to the incident is investigated by conducting a pre-post study design and a computerized text analysis. A pre-post study design is selected since it enables to investigate the outcome before and after an event (Thiese, 2014). More explicitly, if the event or intervention affects the outcome. Computerized text analysis is also selected since it enables us to analyse the tone of the language in the text which is of interest in this study. Sustainability reports retrieved from GRI's database are used for the two chosen industries. This means that the reports comply with GRI's standards and refer to them. The control industry F&B is chosen since it is assumed to be a stable industry meaning less sensitive to economic downturns (Verbeek, 2017). We included a control industry in order to control that no other event happened at the same time and influenced the outcome. Sustainability reports for five years are extracted, 2008-2012. The sample consists of 214 O&G reports and 176 F&B reports. Parts of the Sustainability reports are analysed using the computerized text analysis tool, DICTION. The text is analysed to see if the tone changes after the event by comparing the outcomes before and after the disaster. This was done by analysing the

DICTION master variables; activity, optimism, certainty and realism. These four master variables are also regressed in order to see if there are any statistically significant results.

1.4 Significant Findings & Conclusion

The results of this study show that one of the four master variables change, certainty decreases after the incident as we expected based on prior literature. This implies that the industry responds with a less certain language after the incident. Since this change in tone occurred in Sustainability reports for the O&G industry, this implies that there is a spill-over effect to the industry. Moreover, that the disclosure quality decreases after the event based on our assumption that if the tone change in the expected direction that gives a signal about the disclosure quality. A less certain language is not good for the quality since companies might be hiding their bad performance. This study provides evidence that O&G companies Sustainability reports becomes less certain in their language after the Deepwater Horizon oil spill. Moreover, one of the master variables does show significant evidence of change. Since the other master variables does not significantly changes, no other conclusion regarding a change in tone can be made. Meaning, since the other variables of tone did not change as expected, we cannot make a conclusion regarding the industry's response and the disclosure quality related to those variables. With the limited empirical evidence, the conclusion of this study is that O&G companies have a less certain tone in their Sustainability reports after the Deepwater Horizon oil spill. This suggests that there is a spill-over effect to the industry. We also found that industry, distance, age and text length affect the master variables.

1.5 Contribution & Implications for Research

This study confirms prior research regarding incidents impact on sustainability disclosures by examining how the language in Sustainability reports change due to the Deepwater Horizon oil spill. Moreover, this study confirms prior studies of the DICTION master variables for tone. One of the master variables changes in the expected direction and this result was statistically significant. Certainty in the Sustainability reports decreases for O&G companies after the Deepwater Horizon oil spill. Also, this study contributes to the research field of disclosure quality by using one of many methods to study disclosure quality. As mentioned, there is no consensus about the best way of examining disclosure tone and quality. In this study we have done it by looking at tone and using the text analysis software DICTION. Moreover, compared to previous research, this study has focused on quality and not only quantity. Disclosure quality may circumvent information overload and corporate greenwashing and it is important to highlight in this study to know if companies are

only complying rather than making a CSR effort. Also, since not all master variables change in the expected direction, this could be due to impression management. This study contributes to the research field by giving explanations to why the tone and quality does not change as expected. Moreover, this study contributes to the research field of spill-over effects. Deegan, Rankin and Voght (2000) argue that incidents which are widely known and covered by media are more likely to threaten the legitimacy of the companies operating in the industry involved. We can find statistically evidence that there was a spill-over effect to the O&G industry for the master variable certainty. Moreover, we can conclude that distance, industry and text length affect the master variables. This also contributes to research of tone and the DICTION master variables.

Organizational stakeholders are concerned of how companies address their environmental impacts and therefore should external environmental reports provide disclosures about future strategies and plans (Raiborn et al., 2011). Consequently, it should be of interest to stakeholder to see how O&G companies handled the Deepwater Horizon incident and what they write about it in their Sustainability reports. Mi, Sheng and Elrod (2016) point out the importance of evaluating the quality of disclosures since there exists many cases where listed companies delay the disclosure or disclose false information which could be of importance to investors. Our findings related to tone and disclosure quality can therefore be interesting to investors. According to Baretta, Demartini and Trucco (2019) there is a growing attention towards content analysis of integrated reporting initiative and how companies report their performance. Baretta et al. (2019) argue that both investors and financial analysts could benefit from research regarding the information in voluntary disclosures, since they then can make less uncertain investment decisions with better forecasts of companies' future performance. Thus, this study may be interesting for investors and financial analysts since this study sheds light on what companies disclose in their Sustainability reports. This study is interesting for scholars and researchers that are in the same research field, as well as investors, shareholders and other stakeholders could find it interesting to see how companies react and handle crises.

1.6 Limitations

The method of this study has some limitations, for example that pre-post studies cannot control for other events that happens at the same time (Thiese, 2014). Meaning, other incidents could have happened at the same time as the Deepwater Horizon oil spill and influenced the outcome. This is the greatest limitation of this study, even if we have tried to find other major events during the

same time frame, it is not possible to fully rule out that other events could have happened at the same time and influenced the outcome. The financial crisis could also have impacted the outcome. We tried to limit this by selecting a stable control industry in the sense that it is less sensitive to market movements and economic downturns (Verbeek, 2017). Moreover, this study cannot answer why sustainability disclosure quality is affected because of its quantitative nature. Furthermore, since there is no specific way to measure industry response and disclosure quality, some measures may not give the same results as other measures. This makes this study difficult to replicate. We assumed that if the tone changes in the expected direction that implies that the sustainability disclosure quality changes. This is a simplification in order to make a connection between changes in tone and changes in sustainability disclosure quality. The selection of paragraphs for the analysis might somewhat be biased with our own interpretation. We have used our own judgment when it comes to selecting parts that describes the firms environmental or social impact related to the keywords. However, we have done the selection systematic and clearly describes the process in order to create replicability. Another limitation with this study is the amount of missing observations from the sample. Some reports are too old and no longer accessible. Another problem is that some companies had reports that covered multiple years, these reports are not included in the sample.

2. LITERATURE REVIEW

This section provides a literature review, theoretical perspective and previous research. Leading to the hypothesis development.

2.1 Incentives to Disclose

This section firstly describes the reason for having sustainability disclosures in order to reduce information asymmetry. Voluntary disclosures arise from information asymmetries and the need to minimize that gap. Thus, information asymmetry is important to discuss since it is one rationale behind having voluntary disclosures. Secondly, theories that describes why companies make sustainability disclosures are presented. Companies need to disclose their sustainability performance in order to gain and maintain legitimacy. Also, if the tone does not change in the expected direction, that could be due to impression management.

2.1.1 Information Asymmetry

Verrecchia (2001) argue that there is a link between information asymmetry and disclosures since the need for more disclosures arises from more information asymmetry. Also, Healy and Palepu (2001) argue that the demand for disclosure arises from information asymmetry. There is a need for voluntary disclosures and more specifically sustainability disclosures since information asymmetries exist. Companies can bridge this information gap by voluntary disclose their environmental and social impact. One explanation for the link between information asymmetry and disclosures is that information asymmetry inhibits investments, which will lead to higher costs of capital for the firm (Verrecchia, 2001). Therefore, by disclosing more information, companies reduce the information asymmetry and by so reduce the cost of capital (Verrecchia, 2001). In this case, the Deepwater Horizon oil spill may have increased the information asymmetry gap between the companies (especially BP) and their investors, leading to a higher demand for sustainability disclosures.

2.1.2 Institutional Theory & Legitimacy Theory

The institutional setting can be helpful to explain the behaviour of an organization, in this case why companies disclose information. Institutional theory explains why organizations are similar to each other and continuously becomes more homogenous through different isomorphic pressures; coercive, mimetic and normative isomorphism (DiMaggio & Powell, 1983). Coercive pressure leads

to organizational change and homogeneity since companies need to gain legitimacy (DiMaggio & Powell, 1983). One of the most dominant theories that tries to explain the reasons behind disclosures is the legitimacy theory (Deegan et al., 2000). A central concept in this theory is the social contract since flaws in the contract may lead to changes in the perception of the organization's legitimacy (Meyer & Rowan, 1977; Deegan et al., 2000). Therefore, Deegan et al. (2000) argue that companies are able to maintain their legitimacy through corporate social disclosures. Chan, Watson and Woodliff (2014) state that the society will react when companies fail to operate in legitimate manner, which will threaten the organizations' contract to continue their business. Therefore, a company's survival depends on society's perception of the social contract, since the society is able to revoke the contract if the company does not operate in a legitimate manner (Deegan, 2002).

Deegan et al. (2000) argue that a major incident like an oil spill, may threaten the company's legitimacy, because then they do not operate within the bounds of their social contract. Therefore, companies must undertake strategies like disclosures in order to change society's view of the company. This is in line with Chan et al. (2014) who state that this theory relies on the assumption of adopting strategies in order to show the society that they will comply with the society's expectations. Companies may damage their reputation if they do not meet their social expectations (Chan et al., 2014). Therefore, voluntary sustainability disclosures may be used by companies to legitimize their business (Chan et al., 2014). In this setting, O&G companies (and especially BP) might have needed to disclose their environmental performance in order to save their reputation and legitimacy. The question is whether they only presented more disclosures or if the quality of the text increased.

Moreover, one competing theory that offer explanation for why companies disclose environmental information are the Voluntary Disclosure Theory (VDT) (Cho, Freedman and Patten, 2012). This theory instead suggests a positive relationship between sustainability disclosure and sustainability performance, while legitimacy theory suggest a negative relationship (Hummel and Schlick, 2016). Hummel and Schlick (2016) results showed that superior sustainability performers prefer sustainability of high quality in order to signal their performance to the market, which they argue are consistent with VDT. Consequently, legitimacy theory suggests that firms with poor environmental performance disclose environmental information in order to regain legitimacy and

bridge the information asymmetry gap. Whereas, VDT imply that firms with good environmental performance disclose environmental information.

2.1.3 Impression Management

In order to maintain organizational legitimacy, Nègre, Verdier, Cho and Patten (2017) argue that companies use impression management strategies. Impression management means that managers opportunistically utilize information asymmetries and emphasize positive outcomes and/or avoid negative outcomes (Merkl-Davies, Brennan & Mcleay, 2011). Leung, Parker and Courtis (2015) argue that minimal narrative disclosure is an impression management strategy used by companies in order to hide information and explanations about poor performance in attempts to distract investors' attention from the company's negative news. According to Leung et al. (2015) impression management could lead to capital misallocation since it has the potential to harm the quality of the reports. Leung et al. (2015) argue that companies may manipulate the presentation and the content of voluntary disclosures by engaging in impression management in order to create a perception that the stakeholders will favour and distract investor attention from negative information. Melloni, Stracchezzini and Lai (2016) argue that companies may adopt these impression management strategies in order to manipulate the tone of disclosures. Where Arena, Bozzolan and Michelon (2015) argue that companies may use a more optimistic tone in order to signal future positive environmental performance. Meaning, the tone in the reports may be manipulated as an impression management strategy in order to maintain their legitimacy.

2.2 Prior Empirical Work

This section describes previous research in the field of sustainability disclosures related to some kind of incident, which will be used as a basis for our research question and hypotheses' development. Clarkson, Li, Richardson and Vasvari (2008) results showed a positive relationship between the level of discretionary disclosures and environmental performance where they examined both environmental and social reports as well as related web disclosures. Clarkson et al. (2008) found that, based on voluntary disclosure theories, better environmental performers where more forthcoming in their discretionary disclosure channels. Moreover, Deegan et al. (2000) results showed that firms that operates in an industry where an incident had occurred, did disclose more social information in their annual reports afterwards. Deegan et al. (2000) claim that this is due to the firm's urge to legitimise themselves by changing society's perception of their business. This is in line with Aureli, Medei, Supino and Travaglini (2017) who analysed sustainability disclosures of

companies facing a legitimacy crisis and how a negative externality which is widely reported in the media affects sustainability communication. Their results show that reporting is a tool of legitimacy, but that companies might decrease the transparency after a disaster in order to regain legitimacy (Aureli et al., 2017). Cho and Patten (2007), and Cho (2009) also found that environmental disclosures are a powerful legitimacy tool. Coetzee and van Staden (2011) examined the mining industry, where their results showed that after an incident, the entire industry had increased disclosure levels, which suggest that companies do respond to increased legitimacy threats and stakeholder pressure. Vourvachis et al. (2016) showed in their study that airline disasters led to increased CSR disclosures for the company involved in terms of pages. However, Vourvachis et al. (2016) study also pointed out that the companies extensively increased their CSR disclosure instead of actually discuss the accidents themselves in their Annual reports.

Summerhays and de Villiers (2012) investigated the disclosure patterns and strategies that major companies in the O&G industry used in response to the Deepwater Horizon crisis to get a better understanding of the disclosure decisions and strategies when a crisis occurs. The authors found that there was an increase of environmental disclosures in six other major oil companies than BP in their Annual reports from 2010 (Summerhays & de Villiers, 2012). However, they also found that the proportions of the disclosures appeared to be similar to the proportion in 2009. Summerhays and de Villiers (2012) conclude that it is likely that companies within an industry where a crisis occurred used the same disclosure strategies (increased the volume of environmental information) in order to regain their legitimacy. However, they investigated the Annual reports for the same year as the crisis, where they argue that the companies do not take responsibility for the crisis (Summerhays & de Villiers, 2012). Even though the volume of environmental disclosures did increase, Summerhays and de Villiers (2012) found that BP also disclosed remedial activities over and over again in the reports.

Previous research indicate that companies are more likely to increase their sustainability disclosures when facing a crisis. However, compared to Summerhays and de Villiers (2012), this study investigates if there is a spill-over effect to other companies in the same industry. This study does not focus on the disclosure strategies, instead it examines the industry's response the years before, but also, the years after the event in order to examine the changes in tone over the years. Compared to Coetzee and van Staden (2011) or Deegan et al., (2000) this study uses a larger sample and a control industry. This study also distinguishes from previous research (Deegan et al., 2000; Coetzee

& van Staden, 2011; Summerhays & de Villiers, 2012; Vourvachis et al., 2016) in the sense that Sustainability reports published in accordance with GRI's guidelines are used, instead of Annual reports. Also, this study does not focus on one area but rather a global sample. Based on prior literature, BP and the O&G industry should react with more CSR disclosures after the Deepwater Horizon oil spill since these findings suggest that the amount of environmental and social disclosures increases after an incident. However, we are also interested if the quality of the disclosures increased or decreased.

2.2.1 Tone in Previous Research

Burks, Cuny, Gerakos and Granja (2018) examined the relationship between changes in competition, changes in tone and level of voluntary disclosures in press releases. They explain that a key element of basic disclosure theory is tone (Burks et al., 2018). They also argue that performance can affect the nature and tone of voluntary disclosures (Burks et al., 2018). Baretta et al. (2019) found evidence that non-financial performance is positively related to optimistic tone. Arena et al. (2015) showed evidence that companies use a more optimistic tone in their 10-K environmental disclosures in order to correctly inform about their future environmental performance. Cho et al. (2010) results show that companies with lower environmental performance use language and tone to bias the message in their environmental disclosures. Moreover, Hummel and Schlick (2016) found that "good sustainability performers" have high-quality sustainability disclosures and "bad sustainability performers" have low-quality sustainability disclosures. Cho et al. (2010) also found that environmental performance and the certainty score of disclosures was negatively related, meaning that "bad performers" attempt to hide this by using convoluted and less certain language.

It is somewhat difficult to measure tone and disclosure quality since there is no clear definition (Leuz & Wysocki, 2016). Ammad et al. (2018) have used the GRI guidelines when assessing disclosure quality of Sustainability reports, stating that the content in the reports needs to be verifiable and covers both good and bad news in order to be classified as high quality. This study focuses on textual attributes in order to measure disclosure quality. One way of measuring disclosure quality is to look at the textual tone (Beattie, 2014; Hummel & Schlick, 2016). According to Melloni (2015) tone is a characteristic used to evaluate disclosure quality and it refers to how information is communicated. This study will focus on tone in order to analyse how the language

changes over time. In order analyse the tone, four master variables from DICTION are used; activity, optimism, certainty and realism.

2.3 Variables and Hypotheses

We used the software DICTION in order to analyse the qualitative aspects of Sustainability reports. The text analysis software calculates the frequency of words related to different categories. The main idea of DICTION is to capture the tone and general understanding of the text. In this study we have focused on four master variables; activity, optimism, certainty and realism. Each of these are calculated using a formula that are built on number of words related to different concepts. DICTION then generates a score for each master variable and adds a constant of 50 to each of the observations as a statistical correction. By looking at the four master variables from DICTION that describes tone we hope to fulfil the aim with this study and see if there is a change in tone after the event.

2.3.1 Activity

The first master variable is activity, DICTION calculates activity by the following formula:

$$Activity: [Aggression + Accomplishment + Communication + Motion] - [Cognitive\ Terms + Passivity + Embellishment]$$

The formula consists of different words that DICTION calculated by using different concepts. The formula for activity builds on words that represent forceful action, energy, goal-directedness and task-completion (DICTION, 2014). The activity score will increase if words related to aggression, accomplishment, communication and motion increases. Consequently, the score will decrease if the amount of words related to cognitive terms, passivity and embellishment increases. In this study, we expect an increase in the activity score after the disaster. Sydserff and Weetman (2002) studied impression management and trust, they concluded that "bad performers" had more words that lead to higher scores for the variable activity. Meaning, that "good performers" had a low activity score. Thorpe, Craig, Hadikin and Batistic (2018) had the same conclusion, high-ranked universities, "good performers" had lower activity scores. This might be because "bad performers" feel the need to present themselves as better performing and therefore write in a more forward-looking and forceful way (Sydserff & Weetman, 2002). This strengthens our idea that companies after a disaster will have a higher activity score. Especially BP since they were the concerned company and therefore a "bad performer". Thorpe et al. (2018) investigated the semantic tone in

environmental submissions for low-ranked and high-ranked universities. They suggest that the activity score might be low since there are well-established and well-settled institutions that are comfortable in their strategy and identity (Thorpe et al., 2018). Thus, if the tone does not increase as expected, it might be because there are well-established companies. However, we expect the activity score to increase for BP after the disaster since they want to present themselves as better performing than they actually were. We also expect a spill-over effect to the O&G industry, meaning that the industry should respond with a more active tone in their Sustainability reports after the incident. Moreover, if the activity score actually increases as we expect based on prior research, then that suggests that the disclosure quality in the Sustainability reports decreased since a higher activity score indicates that they are hiding bad performance. If the activity score increases as we expect that means that the sustainability disclosure quality decreases because of the event[1]. We expect the activity score to increase but that is not good for the disclosure quality. This leads to our hypothesis:

H₁: Activity in the language increases after a disaster for the concerned company and industry

2.3.2 Optimism
The second master variable is optimism, DICTION calculates it by the following formula:

Optimism: [Praise + Satisfaction + Inspiration] − [Blame + Hardship + Denial]

This formula consists of affirmations, positive affective states and abstract virtues (DICTION, 2014). Subtracted from this formula are terms about social inappropriateness and downright evil, which are included in blame, as well as adjectives that describe unfortunate circumstances and unplanned changes (DICTION, 2014). Hardship includes natural disasters such as pollution and human fears like grief and death (DICTION, 2014). Praise, satisfaction and inspiration increase the optimism score while blame, hardship and denial decrease the optimism score. Meaning, it measures the frequency of positive concepts in a text and subtract negative concepts. Melloni et al. (2017) explain that a higher optimism score in DICTION indicate a more optimistic tone in the text. Patelli and Pedrini (2014) concluded that there is a positive association between optimistic tone and firm performance in CEO letters. On the other hand, Cho et al. (2010) found evidence that companies that perform worse use a more optimistic language in their environmental

---

[1] This is built on our assumption that if each master variable changes in the expected direction that implies that the disclosure quality increases. If the master variables change in the expected direction, the Sustainability reports are a good reflection of the underlying economics. However, this is only applicable if what we expect is positive for the quality. Consequently, if we expect a change in tone that is bad for the quality, then, if the master variable changes in the expected direction, that implies that the disclosure quality decreased.

disclosures. Moreover, Melloni et al. (2016) showed that companies with worse performance use a more optimistic and positive tone in order to "balance" the negative perception of the company. We expect a negative relationship and therefore expect the optimism score to decrease after the event. Also, it is understandable to have a less optimistic language after an incident. We also expect a spill-over effect to the O&G industry, meaning that they respond with a less optimistic language after the incident. If the optimism score decreases as we expect that means that the sustainability disclosure quality increases because of the event since the Sustainability reports then would capture the underlying economics[1]. Meaning, if the tone in Sustainability reports are less optimistic after the incident, then the reports would have captured the true underlying economics which means high-quality. Our hypothesis is as follows:

H_2: Optimism in the language decreases after a disaster for the concerned company and industry

2.3.3 Certainty

The third master variable is certainty, DICTION calculates certainty by the following formula:

Certainty: [Tenacity + Levelling + Collectives + Insistence] – [Numerical Terms + Ambivalence + Self-reference + Variety]

The certainty formula consists of tenacity, levelling, collectives, insistence, numerical terms, ambivalence, self-reference and variety (DICTION, 2014). Words that express hesitation or uncertainty are included in ambivalence (DICTION, 2014). If words related to tenacity, levelling, collectives and insistence increases then the certainty formula will increase. Moreover, if words related to numerical terms, ambivalence, self-reference and variety increases then the certainty score will decrease. Cho et al. (2010) results showed that there is a negative relationship between environmental performance and certainty, companies with bad environmental performance used a less certain language in their environmental disclosures. Moreover, a decrease in the certainty score could be an indication of an impression management strategy where companies use a less certain language in their disclosures in order to manage stakeholder impression (Cho et al., 2010). We expect that the certainty score for BP should decrease after the incident since they are "bad performers". Also, we expect the certainty score to decrease for the O&G industry, meaning that there should be a spill-over effect. However, if the certainty score does not decrease as expected it might be because there are large, profitable and old companies in the sample. "Firms that are larger, more profitable, and older tend to use more certainty in the language of their environmental disclosures, whereas companies with higher levels of capital intensity use exhibit lower levels of certainty" (Cho et al., 2010, p. 440). However, we expect the certainty score to decrease for the

O&G industry. If the certainty score for BP and the industry decreased as expected, it implies that the sustainability disclosure quality decreased because of the event since they are hiding their bad performance which is not positive for the disclosure quality[1]. Our hypothesis for certainty is as follows:

H₃: Certainty in the language decreases after a disaster for the concerned company and industry

2.3.4 Realism

The fourth and final master variable that we focus on in this study is realism. It is calculated as follows:

Realism: [Familiarity + Spatial Awareness + Temporal Awareness + Present Concern + Human Interest + Concreteness] − [Past Concern + Complexity]

The formula for realism consists of the most familiar words in the English language, present concern in general, picturable terms and words that occurs frequently in American English (DICTION, 2014). Subtracted in this formula are past concern and complexity (DICTION, 2014). This means that if words related to familiarity, spatial awareness, temporal awareness, present concern, human interest and concreteness increases, then the realism score will increase. Accordingly, if the amount of words related to past concern and complexity increases, the realism score will decrease. Intuitively, realism should decrease in our study since that would indicate that the Sustainability reports becomes more complex and vaguer after the incident. According to Wisniewski and Yekini (2015, p. 288) firms without any manifested successes are more prone to use "vague, abstract and idealistic statements that are not rooted in material reality". Thus, after the disaster it would be likely that the text is more abstract and vaguer since they do not have any manifested success. We expect the realism score to decrease for BP, moreover, we expect a spill-over effect to the O&G industry. Meaning that the entire industry should respond with a less realistic tone. Furthermore, if the realism score decreases as we expect that means that the sustainability disclosure quality decreases because of the event since a lower realism score is not good for the disclosure quality[1]. Leading to our hypothesis:

H₄: Realism in the language decreases after a disaster for the concerned company and industry

3. RESEARCH DESIGN

This study's methodology is presented where pre-post study and computerized text analysis are explained, as well as the sample for the study and data collection.

In order to investigate if the Deepwater Horizon oil spill affected the tone in Sustainability reports for O&G companies, this study examined sustainability disclosures two years before the event and two years after the event, meaning 2008, 2009, 2011 and 2012. Five years were selected in order to analyse the outcome before and after the event. Bryce, Ali and Mather (2015) used five years in their pre-post study. Marra, Mazzola and Prencipe (2011) used two years prior to the event and two years after the event. Baig and Khan (2016) used six years. Thus, we used two years before and two years after the event. We also included the year of the incident in order to see changes over time. The F&B industry were selected as a control industry in order to compare the results, this industry was selected since it was unrelated and less sensitive to economic downturns than O&G (Verbeek, 2017). We wanted to test if the four master variables from DICTION changed as we expected based on prior literature. If the variables for tone changed in the expected direction, that implied that the disclosure quality changed after the incident. This was tested using t-tests and regressions.

3.1 Sample

Using GRI's website, we selected the relevant industries and years. Companies with three, four or five available Sustainability reports in GRI's database for the accurate years were extracted. The O&G industry were included in the energy sector on GRI, therefore, the industry classification for each company in the energy sector was controlled by using S&P Capital IQ in order to select a sample of only O&G companies and extract other energy companies[2]. This resulted in 160 O&G companies and 143 F&B companies. The ones that were not O&G were extracted, 78 companies were operating as energy companies not related to O&G. Also, reports in other languages than English were extracted since DICTION cannot interpret text in other languages than English. For O&G, 14 companies did not present reports in English. The reports could for example be in Chinese, Russian or Norwegian. For F&B, 44 companies did not have English reports. We also screened out Daughter companies and, in those cases, only kept the Mother company. In cases

---

[2] Not the correct classification according to S&P Capital IQ. GRI had broader classifications and we therefore had to check S&P Capital IQ in order to get a sample of only O&G companies.

where there was a Daughter company without a Mother company in the sample, we kept the Daughter company. For O&G companies, we extracted 13 Daughter companies. For F&B, we extracted 33 Daughter companies. Our final sample consisted of 51 O&G companies and 44 F&B companies. As Table 1 shows, our sample resulted in 214 reports for O&G and 176 reports for F&B. Leading to a final sample of 95 companies and 390 reports.

Table 1. Sample O&G companies and F&B companies

	Oil and Gas	Food and Beverages	Total
Number of firms from GRI	160	143	303
Not the correct classification	78	0	78
Not in English	14	44	58
Daughter Company	13	33	46
Not available to download/copy	4	15	19
Reports with multiple years	0	7	7
Total number of firms	**51**	**44**	**95**
Number of reports			
Number of firms with 5 reports	23	14	185
Number of firms with 4 reports	18	16	136
Number of firms with 3 reports	9	14	69
Total number of reports	**214**	**176**	**390**

3.2 Data

GRI have issued the Sustainability Reporting Standards with the aim to increase the quality and harmonize the landscape for sustainability reporting (GRI, 2019). Companies can voluntarily reference to the Sustainability Reporting Standards and all reports that are based on GRI are available in their database (GRI, 2019). This means that the reports comply with GRI's standards and are voluntary. We used the GRI database in order to get Sustainability reports that were more comparable and specified (Raiborn et al., 2011). The Sustainability Reporting Standards are structured as follow, universal standards that includes the foundation, general disclosures and management approach. Then, there are topic specific disclosures which are economic, environmental and social. Companies need to use the universal standards but can choose among the topic specific (GRI, 2019). GRI have an Application Level Check system that makes sure that the reports fulfil the requirements, there are three levels A, B and C (GRI, 2013). The system confirms if the reports fulfil the guidelines for the level it is said to be. If not, GRI urges the company to improve the report (GRI, 2013). However, GRI does not check the quality of the reports, only if they fulfil the guidelines (GRI, 2013). It is the company's own responsibility to make sure that the information is of high quality. All reports were extracted from GRI's website or each company's specific website. However, we had a lot of missing observations in our sample.

Some reports were no longer accessible, some were covering multiple years and others were not possible to copy paste since they were scanned as a picture.

We decided to select parts of the Sustainability reports for our analysis since that is common in text analysis in order to capture relevant sections, so they do not get lost in long reports. We extracted between 150 and 4249 words from each report, the mean was 1615 words[3]. Moreover, since we used Sustainability reports that complied with GRI's guidelines that should make them more comparable. The idea was to extract parts that were comparable and discussing relevant topics. The CEO message was selected in all reports (since it was a common denominator in all reports). Also, parts containing keywords related to environment or safety was selected. Table 2 below shows a list of typical headings containing environmental impact or social impact. We decided to include the heading when we extracted the relevant text. We had 390 Sustainability reports in pdf format and had to manually select paragraphs with headings of interest to this study. DICTION cannot interpret pdf format, so we therefore had to manually insert the text into a Word document. Firstly, the CEO's or chairman's letter/letters to shareholders were selected. Tone in letters to shareholders are more positive than the tone in the rest of the reports according to Hildebrandt and Snyder (1981). However, since we look at changes in tone, we believe that incorporating the letters to shareholders will not affect the outcome of this study. Then, we searched for the keywords and selected paragraphs that contained the word. One requirement was that the paragraph should not only mention the word but actually describe the company's environmental or social impact related to the keyword. For example, a company's environmental impact on water could for example be stated under *Water management* where the company describe their use of water and how they tried to limit it. A company's social impact related to accidents could for example be mentioned under *Occupational health and safety* where they stated how many and what kind of accidents that had happened during the year. The selected parts were then inserted into Word documents in order for DICTION to interpret the text. Table 2 presents the keywords we have used and examples of different headings.

---

[3] Please see Table 4 and Graph 2.

Table 2. Keywords and headings

Topic	Keyword	Headings
Environmental impact	*Pollution*	Enhancing our supply chain Environmental protection Implement clean production Safety production Industrial and environmental risks
	Water	Water Water management Environmental impact Consumption and saving Sustainable use of resources Environmental footprint
	Spill	Prevention of spills Oil spill incident Incident risk Spills Emergency response
Social impact	*Incident*	Pursue zero injuries Safety Serious incident frequency Total recordable incidents Process safety Occupational health and safety
	Accident	Occupational health and safety Zero accidents Health, safety and environment HSE Emergency management Risks related to oil and gas exploration and production
	Injuries	Safety Safe operations Lost-time injury frequency Occupational health and safety

3.3 Pre-Post Study Design

A pre-post study design can be used to study outcomes before and after an event or intervention (Thiese, 2014). The key idea is to investigate if there is a causality between the intervention and the outcome (Harris, McGregor, Perencevich, Furuno, Zhu, Peterson & Finkelstein, 2006). Meaning, if the event or intervention affect the outcome. This design is used to see if the occurrence of outcomes changes because of some event, for example, new regulations (Thiese, 2014). This method suits well with the aim of this study, to see if there is a change due to an event. Moreover, this method enables us to compare the outcome before and after the event in order to see if there is a change in tone. This study design is particularly used within medicine and healthcare to see the benefits of a treatment. Pre-post study design have also been used to study the adaptation of IFRS. One example is Bryce et al. (2015) who studied accounting quality and audit committee effectiveness before and after the IFRS adoption in Australia. Another example is a study that investigated earnings management before and after the IFRS adaption in Pakistan (Baig & Khan,

2016). Other pre-post studies have looked at R&D expenditures, environmental provisions and board monitoring before and after the adaption of IFRS (Marra, et al., 2011; Shah, Liang & Akbar, 2013; Wegener & Labelle, 2017). Pre-post studies can study one-group, a comparison between groups or a control group (Thiese, 2014). The advantage of using a control group is that it strengthens the hypothesis that the outcome is due to the event or intervention (Thiese, 2014). We used the F&B industry as a control group in this industry. One of the drawbacks with pre-post studies is that they cannot control for other events that happens at the same time that might influence the outcome (Thiese, 2014).

3.4 Computerized Text Analysis

Content analysis is a method for analysing documents and text (Bryman & Bell, 2013). For example, Annual reports, newspapers and CEO letters. The characteristics of content analysis is that it is systematic and mostly objective (Bryman & Bell, 2013). One of the key advantages of content analysis is that it enables analysis of organizational values and beliefs (Bryman & Bell, 2013). However, there are some drawbacks as well. For example, the analysis is only as good as the documents (Bryman & Bell, 2013). Also, this method does not really answer why questions. One way of conducting a content analysis is to do a text analysis. A computerized text analysis was used to fulfil the aim of this study. This method was chosen since it is well suited to analyse our research question. We chose to use the software program DICTION that can determine the tone used in different documents. Arena et al. (2015) argue that one of the advantages by using DICTION is that it increases the comparability of disclosures of different lengths. DICTION smooths out the difference in text length by making it to a 500-word norm that is equivalent to the text (Cho et al., 2010). This makes the text comparable. Moreover, Sydserff and Weetman (2002) suggests that DICTION is a useful tool to investigate impression management. This text analysis program searches for text with the following qualities[4]:

Activity: "Language featuring movement, change, the implementation of ideas and the avoidance of inertia"

Optimism: "Language endorsing some person, group, concept or event, or highlighting their positive entailments"

Certainty: "Language indicating resoluteness, inflexibility, and completeness and a tendency to speak ex cathedra"

Realism: "Language describing tangible, immediate, recognizable matters that affect people's everyday lives"

(Citations from DICTION, 2019)

---

[4] For more information about the master variables, please see the formulas in section 2.3 Variables and Hypotheses.

In order to investigate changes in tone we decided to look at textual attributes and conduct a computerised text analysis using the tool DICTION. The aim was to see if the industry responded with a change in language due to the event. We decided to look at tone in order to see how an industry is affected by something that happens one company and by so see if there is a spill-over effect to other companies in the same industry. Moreover, the output from DICTION indicates the tone by looking at the master variables; activity, optimism, certainty and realism. Text analysis and this specific software can therefore give a deeper understanding about the text. We decided to not include commonality in the study since it does not describe the tone of the text[5]. The idea of this study was to compare the output two years before the event with the output two years after the event in order to identify changes in the tone. If the tone changed in the expected direction for the four master variables, that suggest that the disclosure quality increased. The output from the text analysis was analysed using a statistic software (STATA) that enables empirical evidence. We conducted t-tests for each master variable and year. Also, we regressed each master variable. Both industries were tested to examine if there was a change. Table 3 presents the variables we used for our regressions.

---

[5] Commonality is the fifth master variable in DICTION. However, we decided not to include it in our study.

Table 3. Variables

Variable	Type	Description	Source
Activity	Dependent variable	A tone that denotes change and movement, a master variable from the DICTION output	DICTION
Optimism	Dependent variable	If the tone is endorsing and positive, a master variable from DICTION	DICTION
Certainty	Dependent variable	If the tone is resolute and complete, a master variable from DICTION	DICTION
Realism	Dependent variable	A tone that is tangible, relatable and recognizable, a master variable from DICTION	DICTION
POST	Independent dummy variable	Takes the value 0 if it is PRE the disaster and the value 1 if POST the disaster. Meaning, PRE: 2008 and 2009. POST: 2010, 2011 and 2012	Year of the Sustainability report
LONG	Independent dummy variable	The geographical distance between the company's headquarter and the Gulf of Mexico. Takes on the value 0 for short distance and 1 for long distance. Short distance is within 2 000 km and long distance is further than that	S&P Capital IQ
FoodBeverages (Industry)	Independent dummy variable	Takes the value 0 for Oil and Gas. Takes on the value 1 for Food and Beverages	S&P Capital IQ
POSTx LONG	Independent variable	Multiplication of the two variables POST and LONG	
POSTx FoodBeverages	Independent variable	Multiplication of the two variables POST and FoodBeverages	
FoodBeverages xLONG	Independent variable	Multiplication of the two variables LONG and FoodBeverages	
POSTxFoodBeveragesxLONG	Independent variable	Multiplication of POST, FoodBeverages and LONG	
Public (Corporate form)	Control dummy variable	Takes on the value 0 for private and 1 for public	S&P Capital IQ
Employees	Control variable	A proxy for size, approximately how many full-time employees they have	S&P Capital IQ
Age	Control variable	Years between the company was founded and the year of the Sustainability report	S&P Capital IQ
Text length	Control variable	Total words analysed	DICTION

3.5 Restrictions

Text analysis can be done on different kinds of sources, for example, press releases, websites, CEO letters and social media. de Villiers and van Staden (2011) argue that companies use websites to disclose more environmental information after an environmental crisis has occurred. This study has solely looked at Sustainability reports from GRI's database, meaning that companies that do not explicitly refer to GRI was not considered. This is one of the restrictions of this study. Perhaps companies were using other media than Sustainability reports to express their concerns. However, we used the GRI framework since the disclosures are assumed to be more specified (Raiborn et al., 2011) which makes the disclosures more comparable. Another restriction with this study is that we have selected parts of the Sustainability reports even if more text could have been interesting to analyse. We attempted to limit the text to the most relevant sections of each report. Furthermore, one restriction with this study is that we only have one control industry even thought it could have

been interesting to look at one more industry that is more related to O&G. One control industry was enough and gave us a sample that could give significant results. This study is limited to five years, two years before the event, the year of the event and two years after the event. However, the effect of the event could have been more long term. Also, our independent dummy variable LONG could have been a continuous variable rather than a dummy variable. Since we measured the distance from the company's headquarter using their address on S&P Capital IQ to the distance where the incident occurred it made more sense to use a dummy for SHORT and LONG rather than the exact distance. Especially since most of the companies are operating on a global arena with many oil rigs located all around the world. We decided not to include a variable for performance since the sample included both private and public firms, instead did we include the control dummy variable Public. Moreover, it could have been interesting to separate BP and the industry in the t-test to see if the variables changed more for BP. This was not possible to do in STATA and therefore we incorporated graphs in order to separate the results between BP and the industry.

4. EMPIRICS AND ANALYSIS

This part presents the study's empirics and a discussion about the results. The structure is as follows; firstly, descriptive statistics is presented, then tone is presented and discussed. Each master variable is analysed on its own. First, a table that summarizes the scores for BP and the O&G industry average is presented, where the industry average is the average for all companies in the industry for each year. PRE and POST is not the same as year per year. PRE is 2008-2009 and POST is 2010-2013. However, we have decided to include graphs with each year in order to illustrate the change over time. Moreover, the dots indicate the score for each specific year, meaning, the lines do not show movement over time. Normal range low and normal range high are also included which are DICTIONS own values for what can be seen as normal. Secondly, a comparison between the O&G industry and the F&B industry will be presented. Finally, regressions of the different master variables will be presented.

4.1 Descriptive Statistics

The analysis in DICTION resulted in specific scores for each company and year. The results are separated for industries in order to compare the results. Also, the concerned company BP is separated in order to compare with the industry in the graphs. We investigated if the tone changed in Sustainability reports after the Deepwater Horizon oil spill by looking at the DICTION scores; activity, optimism, certainty and realism. Table 4 below presents the descriptive statistics from STATA.

Table 4. Descriptive statistics

Variable	Observations	Mean	Standard deviation	Min	Max
Activity	390	50.425	2.359	34.57	58.20
Optimism	390	51.079	2.684	41.81	67.74
Certainty	390	47.084	3.07	32.77	55.05
Realism	390	44.356	2.68	30.02	53.56
POST	390	.662	.474	0	1
LONG	389	.792	.407	0	1
FoodBeverages	390	.451	.498	0	1
POSTx LONG	389	.527	.499	0	1
POSTxFood Beverages	390	.313	.464	0	1
FoodBeverages xLONG	389	.362	.481	0	1
POSTxFoodBev eragesxLONG	389	.254	.436	0	1
Public	390	.77	.423	0	1
Employees	374	50849.57	84427.69	5.99	472000
Age	390	74.679	79.665	1	646
Text length	390	1615.113	698.140	150	4249

The summary of the variables shows that no extreme values exist, the means are within the normal ranges. Maybe slightly skewed towards the max. We can also note that the dummy variables either has the value 0 or 1. Moreover, the variable Employees misses a few observations and that is due to lack of data from S&P Capital IQ. Also, the number of employees vary broadly, indicating that there are both small and large companies. Moreover, the age of the companies varies as well, from 1 year to 646 years. Also indicating that we have both small and large companies. The correlation between the variables are presented in the table below.

Table 5. Pearson correlation between the variables

	Activity	Optimism	Certainty	Realism	POST	LONG	Food Beverages	Public	Employees	Age	Text length
Activity	1.0000										
Optimism	-0.1332*	1.0000									
	0.0084										
Certainty	0.2520*	0.2056*	1.0000								
	0.0000	0.0000									
Realism	0.1247*	-0.0720	0.0932	1.0000							
	0.0137	0.1557	0.0661								
POST	-0.0452	0.0618	-0.0429	-0.0141	1.0000						
	0.3729	0.2236	0.3979	0.7809							
LONG	0.0408	-0.2720*	0.0512	-0.0198	0.0097	1.0000					
	0.4224	0.0000	0.3141	0.6964	0.8491						
Food Beverages	0.1193*	0.2062*	0.2289*	0.2403*	0.0606	0.0310	1.0000				
	0.0184	0.0000	0.0000	0.0000	0.2321	0.5415					
Public	0.0211	-0.0626	-0.0335	0.0292	-0.0487	-0.1338*	-0.1210*	1.0000			
	0.6776	0.2171	0.5096	0.5653	0.3376	0.0082	0.0168				
Employees	0.1224*	-0.0226	-0.0602	0.0576	-0.0602	0.0434	0.0354	0.2010*	1.0000		
	0.0179	0.6626	0.2458	0.2665	0.2456	0.4033	0.4951	0.0001			
Age	0.0689	0.0060	0.0184	0.1411*	0.0103	-0.0770	0.2848*	0.0852	0.1193*	1.0000	
	0.1745	0.9064	0.7170	0.0052	0.8388	0.1297	0.0000	0.0921	0.0210		
Textlength	-0.0510	-0.1797*	-0.1378*	-0.1092*	0.0646	0.0946	-0.1659*	0.0545	-0.0036	0.0494	1.0000
	0.3152	0.0004	0.0064	0.0311	0.2029	0.0624	0.0010	0.2828	0.9447	0.3309	

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5 presents the Pearson correlation coefficients with significance level; a star indicates that the coefficient is significant at .05. The correlation can vary between –1 and 1, minus for perfect negative linear relationship and 1 for perfect positive linear relationship between the two variables[6]. A value of zero indicates that there is no relationship between the two variables. The correlation coefficient for the relationship between activity and the three other master variables have a small correlation. Activity and optimism have a small negative correlation. The strongest relationship between the master variables is between activity and certainty with a significant r-value at 0.2520,

---

[6] A correlation of -0.5 or 0.5 needs to be further investigated in order to make sure that there is no problem with multicollinearity.

however that is still a small number which means that it is a small correlation. Certainty and optimism also have a small correlation that is significant. Looking at the other variables, LONG and optimism have a significant negative relationship with an r-value of -0.2720 (this is the strongest correlation in the matrix). FoodBeverages have significant correlation with all four master variables, the highest is with realism (0.2403). Moreover, Public have a negative relationship with LONG and FoodBeverages. Employees have a significant relationship with activity and Public. When looking at Age, there is significant correlations with realism, FoodBeverages and Employees. Finally, Textlength have negative significant relationships with optimism, certainty, realism and FoodBeverages. The conclusion from Table 5 is that the correlations are relatively small and there should therefore not be a problem with multicollinearity.

Graph 1 presents the master variables from DICTION in order to compare the scores before and after the event. PRE is the average score before the incident, 2008-2009. POST is the average score after the incident, 2010-2012. The average score for each variable is calculated for the O&G industry[7], BP and the F&B industry. The axis is ranging from 40 to 56 in this graph since DICTION adds a constant of 50 to each observation, it would therefore be hard to see the changes if the axis started at 0.

Graph 1. Comparison of master variables PRE and POST
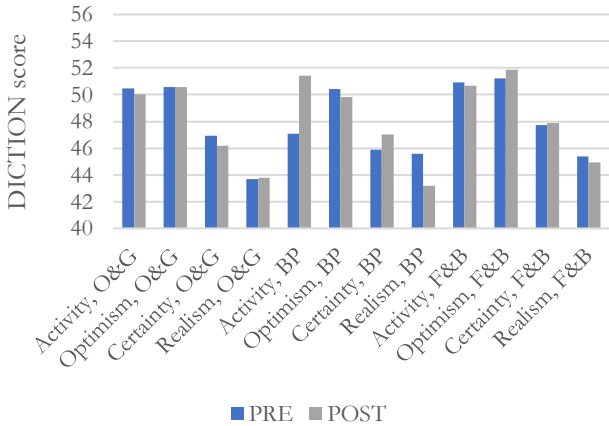
Graph 1 shows that there are some movements in the master variables before and after the Deepwater Horizon oil spill. The change is especially clear for BP. However, there are some changes for both the O&G industry and the F&B industry.

---

[7] BP is not included in the industry average for O&G.

4.1.1 Text Length

It is also of interest to see if the amount of disclosures increased. Graph 2 presents the text length for the parts that we selected and extracted in the Sustainability reports.

Graph 2. Text length between 2008 and 2012 in Sustainability report
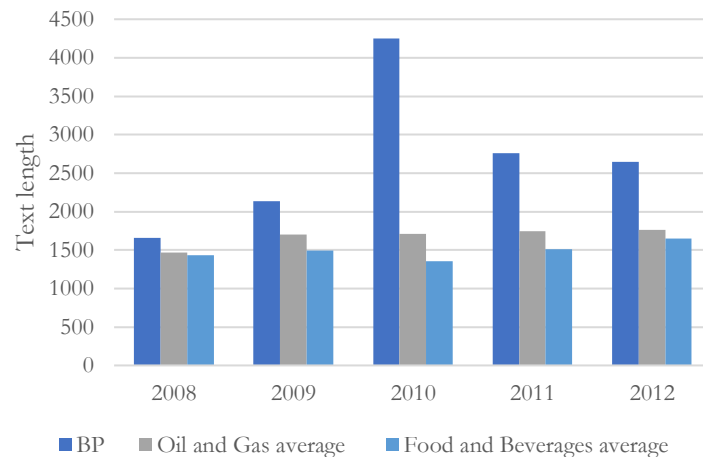
There is a drastic increase in text length for BP in 2010. One explanation for increased disclosure quantity after a disaster is according to Deegan et al. (2000), Chan et al. (2014), and Vourvachis et al. (2016) to regain legitimacy and change the society's view of the company. This might be the case for BP, they might have felt an urge to legitimize themselves after the incident and therefore increased the quantity in their Sustainability reports. Moreover, Vourvachis et al. (2016) also found that the quantity of CSR disclosures increased after accidents, but that they were not actually discussing the accident itself in their Annual reports. This might be the case in our study, the quantity increases but they only mention the death of their 11 employees three times. Thus, they might increase the quantity of their disclosures in order to regain their legitimacy but do not actually discuss the accident. Poor sustainability performers might want to hide their true performance and therefore have low quality disclosures according to Hummel and Schlick (2016). Especially BP might have hidden their low environmental performance behind low quality.

There are different explanations to the fact that the amount of disclosures increased. Greenwashing, information overload and legitimacy are some. Perhaps O&G companies were trying to put themselves in a better position after the incident and therefore not putting more focus than necessary on the incident. Information overload can also be an explanation why the amount

of disclosures increased. Increased amount of disclosures could be a bad signal (Melloni et al., 2017) and it is therefore important to look at the quality of the text. When looking at the development of text length in the Sustainability reports it becomes clear that the text length increased considerably in 2010 for BP (see Graph 2). It is therefore interesting to see if BP increased the quality in their Sustainability reports.
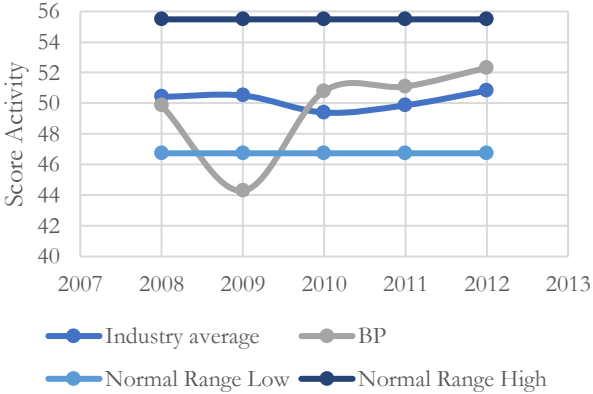
4.2 Activity

The first master variable in DICTION is activity. Our hypothesis is that activity in the text should increase for the concerned company BP and also for other O&G companies after the Deepwater Horizon oil spill. The reason is that it is likely that they want to present themselves as better performers than they actually are (Sydserff & Weetman, 2002; Thorpe et al., 2018). Especially BP since it was their oil rig that suffered the incident. Our hypothesis is as follows:

H₁: Activity in the language increases after a disaster for the concerned company and industry

The Graph 3 below summarises the activity scores for BP and the O&G average. Graph 4 compares the industry averages for O&G to the industry averages for F&B.

Graph 3. Activity score, comparision between average for O&G and BP

Graph 4. Activity score, comparison between industry averages
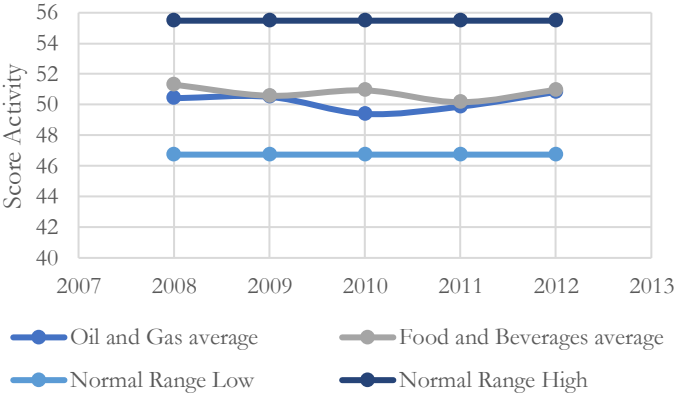
Graph 3 indicates that the activity score increased after 2010 for BP and for the industry. However, the drop for BP in 2009 is difficult to explain, the score is even below what DICTION perceives as normal. Perhaps the straight increase from 2009 to 2010 is because of the oil spill incident that happened in early 2010. The graph indicates that both BP and the O&G industry increased activity in their Sustainability reports after the incident. The rapid increase for BP could potentially be because they were "bad performers" and therefore wanted to present their performance as better

than it actually was. This supports the findings by Sydserff and Weetman (2002) and Thorpe et al. (2018) who finds that "bad performers" have higher activity scores than "good performers". The language and text might be more forceful and forward-looking after a disaster in order to make the reader focus on the future and not the incident. Moreover, what can be observed when breaking down the formula for activity to examine which variables in activity that is changing, is that the variable aggression increased (un-tabulated), which were expected since it includes terms that indicate physical energy with words such as explode or blast, but also words of resistance such as prevent or reduce (DICTION, 2014). However, this increase only occurred in BP, not the O&G industry, which is an indication that BP's reports contained more information about the explosion and ways to prevent this than the industry. Graph 4 shows that the activity score increases for the O&G industry after the incident in 2010. For F&B, there is no clear correlation with the incident in 2010, it first decreases and then increases in 2012.

The master variables where tested using the statistical software STATA in order to examine if there was a significant difference before and after the event. This was done by conducting two t-test, one for O&G and one for F&B. The test separated PRE and POST in order to find if there was a difference in activity before and after the event. Table 6 below presents the output from the t-test for activity and the O&G industry[8] (the output for F&B is un-tabulated).

Table 6. T-test for activity and the O&G industry

Group	Observations	Mean
0	78	50.383
1	136	50.048
Diff		.335
t=0.966		
Ha: diff < 0	Ha: diff !=0	Ha: diff > 0
Pr(T < t) = 0.832	Pr(|T| > |t|) = 0.335	Pr(T > t) = 0.168
Diff = mean (0) – mean (1)	*H0: diff = 0*	

The t-tests did not result in any significant p-values for activity at a .05 significance level. Meaning that the two tests could not find any significant difference in the activity score before and after the event either for O&G or F&B. Consequently, when only analysing the activity score before and after the event, there is no significant difference. Thus, we reject our hypothesis that activity in the

---

[8] The STATA command: ttest Activity if FoodBeverages==0, by(POST), this command was used for each master variable. Also, for the F&B industry: ttest Activity if FoodBeverages==1, by(POST).

language increased after the event and moreover, that this would imply that the quality of the disclosures decreased. Moreover, we cannot say that the industry responded with a more active tone after the Deepwater Horizon oil spill.
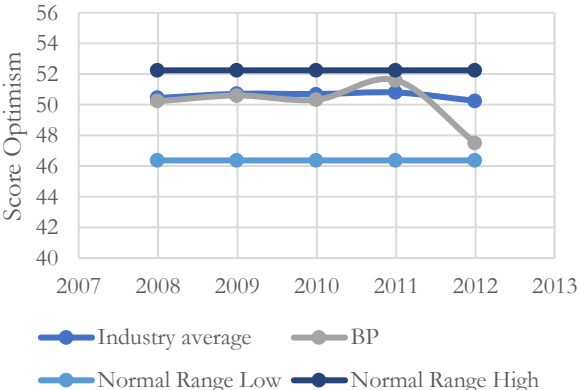
4.3 Optimism

The second DICTION score that indicates the tone of the text is optimism. This variable was expected to decrease after the disaster since the formula extracts blame, hardship and denial. Our hypothesis is as follows:

H₂: Optimism in the language decreases after a disaster for the concerned company and industry

Graph 5 presents the results for optimism scores for BP and the average for O&G between the years 2008 and 2012 in order to identify if there was a change in 2010 when the incident occurred. The optimism score for O&G were also compared with the average for F&B, see Graph 6.

Graph 5. Optimism score, comparison between the industry average for O&G and BP

Graph 6. Optimism score, comparison between industry averages
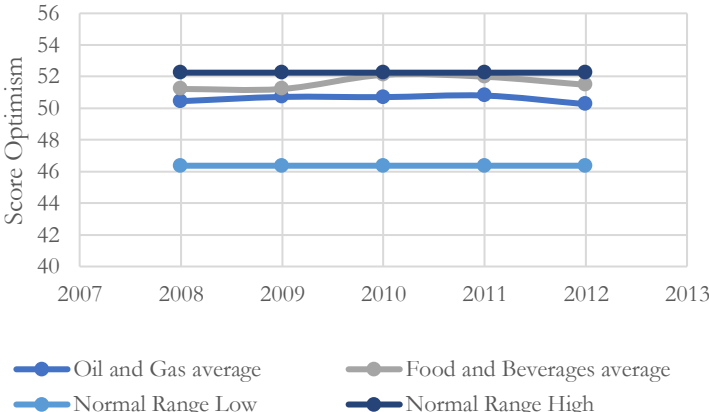
Graph 5 does not show not show any immediate decrease in the optimism score for either BP or the industry. However, in 2012 there is a drastic decrease for BP in the optimism score. The O&G industry average also decreased which could indicate that there was a small spill-over effect. Patelli and Pedrini (2014) found that firms that perform well have a more optimistic tone, while Melloni (2015) found a positive relationship between declining performance and optimistic tone. Due to the spill, BP could be seen as a "bad performer" and it is therefore not surprising that the score decreases in the graph. Since words associated with the terms blame, hardship and denial are subtracted in the formula, intuitively the optimism score should decrease after an oil spill. Hardship might have increased after the event since it includes words like pollution and death. The effect did

not come immediately after the event and this might be because BP wanted to present themselves as better than they were right after the event. Cho et al. (2010) concluded that "bad performers" use a more optimistic tone in their environmental disclosures. Thus, BP might have a more optimistic tone right after the event in order to present themselves as better performers. Then in 2012 the optimistic tone decreased since they actually were "bad performers". When breaking down the master variable into the variables in the formula, we can observe some changes in 2010 for BP (un-tabulated), especially when it comes to inspiration and hardship there are fluctuations around 2010. Hardship includes natural disasters such as pollution and human fears such as grief and dead (DICTION, 2014), therefore, an increase in hardship was expected to be found in BP's Sustainability reports after the incident. However, BP only mention the death of the 11 employees three times and they only mention pollution five times, which could be an explanation for the actual decrease in hardship. This could be an indication that BP tries to bias the message of their environmental performance, which are in line with Cho et al. (2010) who found evidence that bad environmental performers use language and verbal tone in order to bias the message in their environmental disclosures. Looking at Graph 6, the decrease in the optimism score is first in 2011 for F&B companies as for O&G. Moreover, Graph 6 does not indicate that the optimism score for the O&G industry were affected due to the Deepwater Horizon oil spill. Since F&B have the same progress, something else could have happened at the same time and influenced the outcome.

Two t-test were conducted, one for O&G and one for F&B. The test separated PRE and POST in order to find if there was a difference in activity before and after the event. Table 7 below presents the output from the t-test for optimism and the O&G industry (the output for F&B is un-tabulated).

Table 7. T-test for optimism and the O&G industry

Group	Observations	Mean
0	78	50.596
1	136	50.581
Diff		.0149
t=0.040		
Ha: diff < 0	Ha: diff !=0	Ha: diff > 0
Pr(T < t) = 0.516	Pr(|T| > |t|) = 0.968	Pr(T > t) = 0.484
Diff = mean (0) – mean (1)	*H0: diff = 0*	

The t-test for optimism and the O&G industry did not give significant results. When testing for the F&B industry the difference between the group's means were less than zero with an almost significant p-value at .05. The alternative hypothesis that the difference between the groups are less than zero resulted in a p-value of 0.0565. Indicating that for the industry F&B, the optimism score decreases after the event. However, in this study we are interested in the O&G industry. The change in the optimism score for F&B could be because of other events that happened during the same time and affected the F&B. Since there was no change for O&G, we can conclude that the event did not affect optimism in Sustainability reports after the event. Conclusively, we reject our hypothesis that optimism in the language decreased due to the event. This means that since the tone did not change as expected, we could not say that the disclosure quality changed after the event. Consequently, we cannot say that the industry responded to the Deepwater Horizon oil spill with a less optimistic tone.
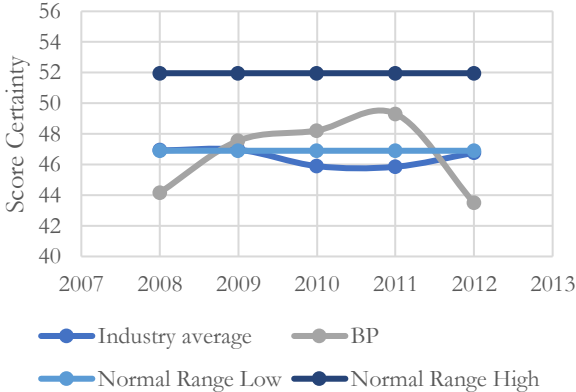
4.4 Certainty

We expected the DICTION score certainty to decrease after the Deepwater Horizon oil spill for the concerned company BP and other O&G companies. Our hypothesis is as follows:

H_3: Certainty in the language decreases after a disaster for the concerned company and industry

The certainty score for BP and the industry average for O&G is presented below in Graph 7. As the other graphs, the certainty score is presented between the years 2008 and 2012 in order to identify the change in 2010 when the accident occurred. Moreover, the averages score for O&G are compared with the average score for F&B in Graph 8.

Graph 7. Certainty score, comparison between the industry average for Oil and Gas and BP

Graph 8. Certainty score, comparison between industry averages
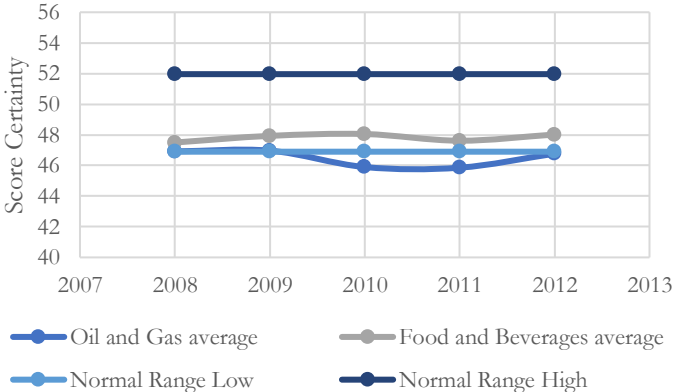
For BP, the certainty score first increases in 2010 and then there is a drastic decrease in 2011. The decrease in 2011 could be explained by Cho et al. (2010) who conclude that "bad performers" have less certain language in their environmental disclosures. Thus, the certainty score may have decreased in 2011 since BP's performance was worse due to the oil spill. When looking at the O&G industry, there is no change between 2010 and 2011. Then in 2011 the score increases slightly. This indicate that there was no clear spill-over effect to other O&G companies. Aureli et al. (2017) found that companies after a disaster might decrease the transparency in order to regain legitimacy. After the Deepwater Horizon oil spill, they might have decreased the transparency in order to legitimise themselves and avoid bad information. This could explain the fact that the certainty score decreased for BP. Also, a less certain tone could be due to an impression management strategy in order to distract stakeholders (Cho et al., 2010). Graph 8 does not give clear evidence that the certainty score for the O&G industry were affected by the oil spill in 2010. When looking at F&B, the progress looks similar to O&G which may indicate that something else happened that influenced the outcome.

Two t-test were conducted, one for O&G and one for F&B. Table 8 presents the output from the t-test for certainty and the O&G industry (the output for F&B is un-tabulated).

Table 8. T-test for certainty and the O&G industry

Group	Observations	Mean
0	78	49.919
1	136	46.448
Diff		.742
t=1.639		
Ha: diff < 0	Ha: diff !=0	Ha: diff > 0
Pr(T < t) = 0.949	Pr(|T| > |t|) = 0.103	Pr(T > t) = 0.051
Diff = mean (0) – mean (1)	*H0: diff = 0*	

The two t-test in STATA generated one significant result. For O&G companies, the certainty score gave a difference between the means that were greater than zero. This means that the certainty score had a significant difference before and after the event for O&G companies. The t-test resulted in a p-value of 0.051 for the alternative hypothesis that the difference between the means are greater than zero[9]. Thus, the O&G industry had a less certain language after the Deepwater

---

[9] We acknowledge that the p-value is 0.051 and not 0.050. However, we believe that this difference is very small and therefore state that the p-value is significant.

Horizon oil spill. This also suggest that there is a spill-over effect to the O&G industry and that they responded with a less certain language. Moreover, we can accept our hypothesis that certainty in the language decreases after the event. These findings are in line with Cho et al. (2010) who found evidence of a negative relationship between environmental performance and certainty, where they argue that companies attempt to hide their bad performance by using less certain language. The less certain language could be explained by the fact that BP and the O&G industry were "bad performers" and therefore had a less certain language in their Sustainability reports (Cho et al., 2010). Moreover, Cho et al. (2010) conclude that "bad performers" want to hide this fact by using a convoluted and less certain language. Also, Roberts and Patten (2010) found that companies with low environmental performance use the tone in environmental disclosures to bias the message and present themselves as better performers. The less certain language after the Deepwater Horizon oil spill could simply be because the O&G industry wanted to present themselves as better performers than they actually were in their Sustainability reports after the incident. The result meant that we could accept our hypothesis and that the tone changed in the expected direction which implies that the sustainability disclosure quality decreased after the event. Moreover, we could see that there was a spill-over effect to the O&G industry and that they responded with a less certain language after the Deepwater Horizon oil spill.

4.5 Realism

The final master variable that we have looked at is realism and we expected the score to decrease after the disaster for BP and other O&G companies. Our hypothesis is as follows:

H₄: Realism in the language decreases after a disaster for the concerned company and industry

The Graph 9 below summarises the realism score for BP and the industry average. Moreover, Graph 10 compares the realism scores between the O&G industry and the F&B industry.

Graph 9. Realism score, comparison between the industry average for O&G and BP
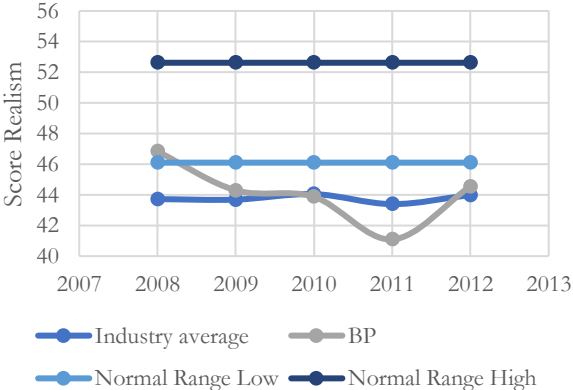
Graph 10. Realism score, comparison between industry averages
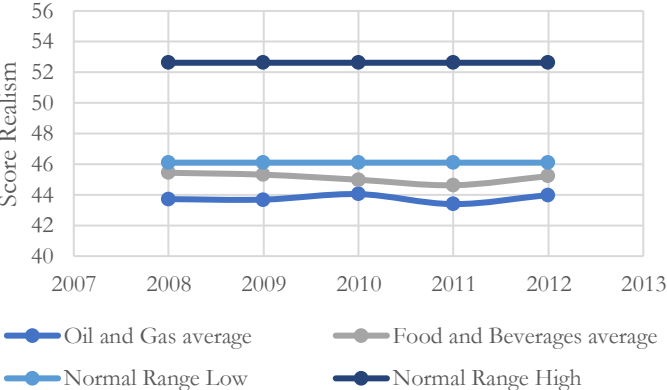
As Graph 9 shows, the scores for both BP and the industry average is below what DICTION perceives as normal values. Moreover, there is a drop in the realism score for BP after the incident. However, the score increases from 2011 to 2012. When looking at the industry average, the curve is fairly straight. Anyway, there is a small decrease in 2010. This suggest that there is a small spill-over effect to other O&G companies. The scores indicate that in the Sustainability reports, words that are related to past concern and complexity has increased after the incident. Especially for BP since there is a rapid decrease from 2010 to 2011. This supports the findings by Wisniewshi and Yekini (2015) who finds that the realism score decreases for firms without any manifested success. The Deepwater Horizon oil spill must have been a catastrophe for the company and might be one of the reasons why the realism score decreased in their Sustainability reports after the disaster. Graph 10 shows two fairly straight lines, however, there is a small drop for both industries in 2010. Both industries are under the normal range low line, this is rather interesting and suggest that both industries have a less realistic language in their Sustainability reports than DICTION suggest that they should have.

Two t-test were conducted, one for O&G and one for F&B. Table 9 below presents the output from the t-test for realism and the O&G industry (the output for F&B is un-tabulated).

Table 9. T-test for realism and the O&G industry

Group	Observations	Mean
0	78	43.749
1	136	43.784
Diff		-.034
t=-0.099		
Ha: diff < 0	Ha: diff !=0	Ha: diff > 0
Pr(T < t) = 0.461	Pr(|T| > |t|) = 0.922	Pr(T > t) = 0.539
Diff = mean (0) – mean (1)	*H0: diff = 0*	

The two t-tests for realism in STATA did not give in any significant p-values. Either for O&G or F&B. Meaning that the t-tests did not identify any difference in the realism score before and after the event. This meant that we reject the hypothesis that realism in the language decreased after the event. Moreover, we could not say that the sustainability disclosure quality changed since the realistic tone did not change as expected. We could not find evidence of a spill-over effect to the O&G industry and that they responded with a less realistic tone.

4.6 Regressions

After looking at the four master variables from DICTION we got one significant result. We could accept our hypotheses that certainty in the language decreased after the incident for O&G companies. However, the next step is regressions and incorporating our independent variables and control variables. The four master variables from DICTION were separately regressed using the variables presented in Table 3. We used three independent dummy variables, POST, LONG and FoodBeverages. The variable POST takes the value 1 if it is after the disaster (2010, 2011 and 2012). The variable LONG takes the value one if the company's headquarters are more than 2 000 km from where the incident occurred, meaning, long distance. FoodBeverages takes on the value 1 for the F&B industry and the value 0 for the O&G industry. We have also used one control dummy variables. The variable Public takes the value 1 if the corporate form is public and the value 0 if the firm is private. We have also used three more control variables, number of full-time employees, the age of the firm and the length of the text DICTION analysed. Each master variable was regressed using the other variables. We have done the following regressions for each master variable:

Regression 1: $Master\ variable = \beta_0 Constant + \beta_1 POST \times LONG + \beta_2 POST + \beta_3 LONG + Public + Employees + Age + Textlength + \varepsilon$

Regression 2: $Master\ variable = \beta_0 Constant + \beta_1 POST \times FoodBeverages + \beta_2 POST + \beta_3 FoodBeverages + Public + Employees + Age + Textlength + \varepsilon$

Regression 3: $Master\ variable = \beta_0 Constant + \beta_1 POST \times FoodBeverages \times LONG + \beta_2 POST \times FoodBeverages + \beta_3 POSTxLONG + \beta_4 FoodBeverages \times LONG + \beta_5 POST + \beta_6 FoodBeverages + \beta_7 LONG + Public + Employees + Age + Textlength + \varepsilon$

The STATA output from each regression is presented below in Table 10. The stars indicate if the p-values are statistically significant. In that case there is some evidence of a linear relationship between the master variable and the other variable controlling for all others. R-square indicates how many percent of the variance of the response variable activity that is explained by our regression. The output can also be interpreted in the following way for each variable, for example, if age increases with one year then activity will increase with 0.00173 (see regression 1). Or if text length increases with 1 then optimism will decrease with 0.000566 (see regression 4) since the coefficient is negative. For dummies they should be interpreted in the following way, on average the value of the variable activity is lower for POST than for PRE by 0.135 (see regression 1) since it is a negative number.

Table 10. Output from STATA, regressions

Variables	(1)Activity	(2)Activity	(3)Activity	(4)Optimism	(5)Optimism	(6)Optimism	(7)Certainty	(8)Certainty	(9)Certainty	(10)Realism	(11)Realism	(12)Realism
POST	-0.135	-0.296	-0.437	0.365	0.0937	0.180	-0.258	-0.578	-0.742	-0.482	0.0152	-0.746
	(0.556)	(0.338)	(0.703)	(0.601)	(0.373)	(0.748)	(0.723)	(0.430)	(0.902)	(0.641)	(0.384)	(0.796)
FoodBeverages		0.352	1.137		0.568	0.261		0.962*	0.611		1.299***	2.385**
		(0.425)	(0.923)		(0.470)	(0.982)		(0.542)	(1.184)		(0.484)	(1.045)
POSTxFoodBeverages		0.160	0.567		0.638	0.362		0.690	1.030		-0.371	0.303
		(0.508)	(1.133)		(0.562)	(1.205)		(0.648)	(1.454)		(0.578)	(1.283)
LONG	0.408		0.738	-1.813***		-2.096***	0.515		0.234	-0.121		0.256
	(0.511)		(0.644)	(0.552)		(0.685)	(0.665)		(0.826)	(0.589)		(0.729)
POSTxLONG	-0.0879		0.176	0.127		-0.138	0.104		0.224	0.505		0.970
	(0.624)		(0.799)	(0.675)		(0.850)	(0.812)		(1.026)	(0.719)		(0.905)
FoodBeveragesxLONG			-1.007			0.492			0.398			-1.361
			(1.039)			(1.105)			(1.333)			(1.176)
POSTxFoodBeveragesxLONG			-0.516			0.445			-0.435			-0.861
			(1.269)			(1.350)			(1.628)			(1.437)
Public	0.0510	0.0725	0.105	-0.517	-0.00296	-0.309	-0.165	0.00952	0.0818	0.182	0.352	0.336
	(0.306)	(0.304)	(0.308)	(0.331)	(0.336)	(0.328)	(0.398)	(0.387)	(0.395)	(0.353)	(0.345)	(0.349)
Employees	2.95e-06**	3.04e-06**	2.58e-06*	3.24e-07	-6.70e-07	3.44e-07	-2.31e-06	-2.34e-06	-2.47e-06	1.00e-06	8.65e-07	3.70e-07
	(1.47e-06)	(1.45e-06)	(1.47e-06)	(1.58e-06)	(1.61e-06)	(1.56e-06)	(1.91e-06)	(1.85e-06)	(1.88e-06)	(1.69e-06)	(1.65e-06)	(1.66e-06)
Age	0.00173	0.000704	0.00137	0.000117	-0.000797	-0.00226	0.00125	-0.00160	-0.00138	0.00494***	0.00288	0.00360**
	(0.00150)	(0.00157)	(0.00159)	(0.00163)	(0.00173)	(0.00169)	(0.00196)	(0.00200)	(0.00204)	(0.00173)	(0.00178)	(0.00180)
Textlength	-0.000179	-0.000102	-6.93e-05	-0.000566***	-0.000522***	-0.000433**	-0.000631***	-0.000405*	-0.000439*	-0.000413**	-0.000290	-0.000208
	(0.000172)	(0.000172)	(0.000176)	(0.000186)	(0.000191)	(0.000187)	(0.000224)	(0.000220)	(0.000226)	(0.000199)	(0.000196)	(0.000200)
Constant	50.24***	50.34***	49.68***	53.46***	51.45***	53.21***	47.98***	47.71***	47.52***	44.62***	43.82***	43.49***
	(0.578)	(0.462)	(0.676)	(0.625)	(0.511)	(0.719)	(0.752)	(0.589)	(0.867)	(0.666)	(0.526)	(0.765)
Observations	373	374	373	373	374	373	373	374	373	373	374	373
R-squared	0.025	0.030	0.046	0.108	0.073	0.155	0.032	0.076	0.079	0.038	0.069	0.091

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

There is no significant difference for PRE and POST. This is rather surprising since we were expecting a change due to the event that took place in 2010. Moreover, the independent dummy variable industry seems to matter since the p-values are significant in the regressions for certainty and realism. Indicating that the values are higher for F&B than for O&G. POSTxFoodBeverages did not generate any significant p-values in any of the regression. It seems like the distance between the companies headquarters to the oil rig is relevant, at least when looking at optimism. The value of optimism is lower for LONG than for SHORT by 1.813 (see regression 4) and 2.096 (see regression 6), these are also significant at .01. This implies that firms with longer distance from the incident had less incentives to regain legitimacy and cover their bad performance with an optimistic tone. What is surprising is that the variable POSTxLONG did not generate any significant results, we were expecting that pre-post and the distance should affect the outcome. Looking at FoodBeveragesxLONG, the regressions did not generate significant p-values. The multiplication of POSTxFoodBeveragesxLONG did not either generate any significant values. Looking at the control variables, we see that Public does not significantly affect the master variables. However, Employees, Age and Textlength did. The variable Employees seems to affect the master variable activity (see regression 1-3). Moreover, Age seems to significantly affect realism (0.00494 and 0.00360, see regression 10 and 12). The control variable text length generated significant p-values for the master variables optimism, certainty and realism, indicating that the amount of words matters. When text length increases with 1 word the value of optimism and certainty decrease with 0.000556 respectively 0.000631 (see regression 4 and 7). This means that when the text increases by one standard deviation (698 words), the optimism and certainty scores increase by 0.39 and 0.44 respectively. All constants are significant at a .01 level, indicating that SHORT and PRE have effect on all master variables. Looking at R-square, the model that explains most is optimism, 15.5% of the variance in optimism is explained by the regression. The other regressions explain between 2-10%. The R-square for activity is fairly low (2.5-4.6%). In these regressions we have tested the effect on the master variables by our independent variables and control variables. This means that we can see what affects the master variables. Conclusively, the four regressions show that industry, distance, age and text length affect the master variables.

5. CONCLUSION

This final section provides a conclusion of this study. Also, contributions, limitations and suggestions for future research are discussed.

The purpose of this study was to examine if tone in Sustainability reports changed after a major incident for the concerned company, and if there was a spill-over effect to the industry. This was tested by analysing the text in Sustainability reports and using a pre-post study design. We chose the Deepwater Horizon oil spill for this study and two industries, O&G and F&B. Selected parts from the Sustainability reports were analysed using the text analysis software DICTION, where focus was on tone, which is a characteristic of disclosure quality (Beattie, 2014; Melloni, 2015; Hummel & Schlick, 2016). Tone was analysed through four master variables from DICTION; activity, optimism, certainty and realism. If the tone changed in the expected direction, that implied that the sustainability disclosure quality increased after the Deepwater Horizon oil spill. The main focus was on how BP and the industry responded to the incident in their Sustainability reports.

The graphs (Graph 1, Graph 3, Graph 5, Graph 7 and Graph 9) indicate that there is a change in activity, optimism, certainty and realism over time. Especially for BP and that there is a small spill-over to the O&G industry. The graphs indicate that there is a change in tone over time. For BP it is easier to identify the change since the increases and decreases are sharper than for the O&G industry. However, a small spill-over effect to the industry could be observed. Our t-tests in STATA did result in one significant conclusion. Certainty in the language decreased after the event for O&G companies. This meant that we could accept our hypothesis *H_3: Certainty in the language decreases after a disaster for the concerned company and industry.* Certainty in the text decreased as we expected based on prior literature and this meant that we could say that the sustainability disclosure quality decreased after the Deepwater Horizon oil spill. Moreover, this suggest that there was a spill-over effect to the O&G industry since they responded with a less certain language in their Sustainability reports after the Deepwater Horizon oil spill. Since the other t-test did not generate significant results, we rejected the other hypotheses related to the DICTION master variables. Furthermore, we could therefore not say that the industry responded or that the sustainability disclosure quality changed for those variables. Our empirics generated one conclusion, certainty in the language decreased for the O&G industry after the Deepwater Horizon oil spill. Meaning that words related to tenacity, levelling, collectives and insistence decreased and/or that words related

to numerical terms, ambivalence, self-reference and variety increased. The reason that not all master variables changed as expected could be because of impression management. Or by the fact that the amount of text increased which could be explained by information overload or greenwashing. One explanation is impression management, the companies might have used the tone to give the reader a positive image and avoid the accident itself. Tone could be used as an impression management disclosure technique to manipulate the disclosures and create a positive image of the company (Merkl-Davies et al., 2011; Leung et al., 2015; Melloni et al., 2016; Guillamon-Saorin et al., 2017). BP and the O&G industry might have used the tone to present a positive image about themselves and shift focus from the incident. The reasons behind that not all master variables changed was not tested in this study. However, we present some ideas for why the tone changed as expected and why it did not.

We expected that the Deepwater Horizon incident would increase the information asymmetry between the companies and their investors since the demand for disclosure arises from information asymmetry and that information asymmetry creates the need for more disclosures (Healy & Palepu, 2001; Verrecchia, 2001). Mi et al. (2015) state that disclosures are key to improve transparency and bridge the information asymmetry gap. The incident would therefore have increased the information gap between BP and investors or other stakeholders and created the need for more disclosures. We were expecting to find that the tone changed because of the event since the O&G industry and especially BP needed to minimize the information gap after the accident. We could conclude that certainty in the language decreased after the event. However, we could also see that the amount of disclosures increased in the Sustainability reports (see Graph 2). This could be due to information overload or greenwashing. The fact that the amount of disclosures increased could also be a signal that the quality decreased (Melloni et al., 2017).

This study found some evidence that the O&G industry responded with a less certain language after the Deepwater Horizon oil spill. This indicates that there was a spill-over effect to the industry and that not only BP changed the tone in their Sustainability reports after the event. We could also find empirical evidence that industry, distance, age and text length affect the master variables. To conclude, certainty in Sustainability reports changed as expected in Sustainability reports for O&G companies.

5.1 Contributions

The results from this study have some practical implications. First, this study concluded that certainty in the language decreased after the Deepwater Horizon oil spill for the O&G industry. These results contribute to research related to tone and spill-over effects. Moreover, this study shows that distance, industry, age and text length affect the master variables. Also, this study confirms prior research regarding companies increasing the quantity of environmental disclosures after an incident. Summerhays and de Villiers (2012) found that O&G companies increased the amount of disclosures after an incident, but they were only repeating the same information over and over again. Our findings are in line with Summerhays and de Villiers (2012) findings, the quantity of disclosures increased, and we found some evidence that the quality did not. Especially BP increased the amount of words after the Deepwater Horizon oil spill. This is interesting findings since it might suggest that the increasing quantity may be because of information overload or greenwashing. Moreover, three master variables did not change as expected and that could be explained by impression management, disclosures could be used as an impression management strategy in order to hide information regarding poor performance as an attempt to distract investor's attention from their negative news. This study is therefore useful for researchers in the field of impression management strategies and disclosures. Finally, this study could be interesting to stakeholders such as investors in order for them to understand how companies discuss incidents in their Sustainability reports.

5.2 Limitations & Suggestions for Future Research

The main limitation with this study is that pre-post studies cannot control for other events that happened at the same time and influenced the outcome. We could not fully control that no other events happened at the same time and affected or results. For example, we almost got a significant result that the optimism decreased after the event for the F&B industry. This result could be due to another event that happened at the same time and affected the optimistic tone for F&B. Moreover, since there is no consensus about the best practice to study disclosure tone and quality, we have only done it in one of many possible ways. Thus, this study might be difficult to replicate if not using the same measures as we did. We have therefore explained the process in detail in order to make this study replicable. Also, we decided to study sustainability disclosure quality by looking at tone, this is only one possible way of examining the industry's response in their Sustainability reports. We assumed that if the tone changed in the expected direction that implied that the sustainability disclosure quality increased, this is a simplification.

We have touched upon some explanations for why not all master variables changed as expected. This could be interesting to develop in another study. For example, investigating if there are boilerplate statements in the reports, meaning if they use a standardized text every year. This would explain why there is no expected change in the tone. This could be tested using a text analysis tool and a benchmark text. Moreover, we present impression management as a reason why not all master variables changed as expected. This could be interesting to examine in a future study. For example, the connection between low-quality sustainability disclosures and the existence of impression management. Another suggestion in this area could be to investigate how companies disclose information about an incident in other media than Sustainability reports afterwards. This study only focused on Sustainability reports, but the tone in disclosures could have changed in other medias as well, for example other reports or their own website.

6. REFERENCE LIST

Ammad, A., Muhammad Bilal, F. and Muhammad, N. (2018). 'Sustainability reporter classification matrix: explaining variations in disclosure quality, *Meditari Accountancy Research*, 26(2), pp.332-352.

Arena, C., Bozzolan, S., and Michelon, G. (2015). 'Environmental reporting: Transparency to stakeholders or stakeholder manipulation? An analysis of disclosure tone and the role of the board of directors. *Corporate social responsibility and environmental management,* 22(6), pp. 346-361

Aureli, S., Medei, R., Supino, E. and Travaglini, C. (2017). 'Sustainability disclosure and a legitimacy crisis: Insights from two major cruise companies', *European Journal of Tourism Research*, 17, pp.149-163.

Baig, & Khan. (2016). Impact of IFRS on Earnings Management: Comparison of Pre-Post IFRS Era in Pakistan. *Procedia - Social and Behavioral Sciences, 230*, 343-350.

Baretta, V., Demartini, C. and Trucco, S. (2019). 'Does environmental, social and governance performance influence intellectual capital disclosure tone in integrated reporting?'. *Journal of intellectual capital,* 20(1), pp.100-124.

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.

BP. (2019). *BP at a glance. BP.* Retrieved from, https://www.bp.com/en/global/corporate/what-we-do/bp-at-a-glance.html, 2019-01-10.

Bryce, Ali, & Mather. (2015). Accounting quality in the pre-/post-IFRS adoption periods and the impact on audit committee effectiveness — Evidence from Australia. *Pacific-Basin Finance Journal, 35*, 163-181.

Bryman, A., & Bell, E. (2013). *Företagsekonomiska forskningsmetoder* (2., [rev.] uppl. ed.). Stockholm: Liber.

Burks, J., Cuny, C., Gerakos, J., and Granja, J. (2018). 'Competition and voluntary disclosure: evidence from deregulation in the banking industry. *Review of accounting studies*, 23(4), pp.1471-1511.

Chan, M., Watson, J. and Woodliff, D. (2014). 'Corporate governance quality and CSR disclosures', *Journal of Business Ethics*, 125(1), pp.59-73.

Chen, Y-C., Hung, M. and Wang, Y. (2018). 'The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China', *Journal of Accounting and Economics,* 65(1), pp. 169-190.

Chiu, T-K. And Wang, Y-H. (2015). 'Determinants of social disclosure quality in Taiwan: an application of stakeholder theory', *Journal of business ethics*, 129(2), pp.379-398.

Cho, C.H. (2009). 'Legitimation Strategies Used in Response to Environmental Disaster: A French Case Study of Total SA's Erika and AZF Incidents', *European Accounting Review,* 18(1), p.33-62.

Cho, C.H., Freedman, M. and Patten, D.M. (2012) 'Corporate disclosure of environmental capital expenditures', *Accounting, Auditing & Accountability Journal*, 25(3), pp.486-507.

Cho, G.H. and Patten, D.M. (2007). 'The role of environmental disclosures as tools of legitimacy: A research note', *Accounting, Organizations and Society*, 32(7), pp.639-647.

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

Clarkson, P.M., Li, Y., Richardson, G.D. and Vasvari, F.P. (2008). 'Revisiting the relation between environmental performance and environmental disclosure: an empirical analysis. Accounting, organizations and society, 33(4), pp.303-327.

Coetzee, C.M. and van Staden, C.J. (2011). 'Disclosure responses to mining accidents: South African evidence', *Accounting Forum*, 35(4), pp.232-246.

de Villiers, C. and van Staden, C.J. (2011). 'Where firms choose to disclose voluntary environmental information', *Journal of Accounting and Public Policy*, 30(6), pp. 504-525.

Deegan, C. (2002). 'Introduction: the legitimising effect of social and environmental disclosures – a theoretical foundation', *Accounting, auditing and accountability journal*, 15(3), pp282-311.

Deegan, C., Rankin, M. and Voght, P. (2000). 'Firms' Disclosure Reactions to Major Social Incidents: Australian Evidence', *Accounting Forum*, 24(1), pp.101-130.

DICTION. (2014). *Help manual.* Retrieved from, http://www.dictionsoftware.com/download.php?file=wp-content/uploads/2014/02/DICTION-7-Manual-2-26-14.pdf, 2019-04-01.

DICTION. (2019). *Diction overview*. Retrieved from, https://www.dictionsoftware.com/diction-overview/, 2019-01-09.

DiMaggio, P. J. and Powell, W. (1983). 'The iron cage revisited: institutional isomorphism and collective rationality in organizational fields*', *American Sociological Review*, 48(1), 147-60.

Eccles, R., Krzus, M., Rogers, J. and Serafeim, G. (2012). 'The need for sector-specific materiality and sustainability reporting standards', *Journal of Applied Corporate Finance*, 24(2), pp.65-71.

Gamerschlag, R., Möller, K. and Verbeeten, F. (2011). 'Determinants of voluntary CSR disclosure: empirical evidence from Germany', *Review of Managerial Science*, 5(2), pp.233-262.

GRI (2013). *Taking responsibility: A discussion about report quality*. Retrieved from, https://www.globalreporting.org/information/news-and-press-center/Pages/Taking-responsibility-a-discussion-about-report-quality.aspx, 2019-03-28.

GRI (2019). *About Sustainability Reporting*. Retrieved from, https://www.globalreporting.org/information/sustainability-reporting/Pages/default.aspx, 2019-01-11.

Guillamon-Saorin, E., Isidro, H. and Marques, A. (2017). 'Impression management and non-gaap disclosure in earnings announcements'. *Journal of business finance & accounting*, 44(3-4), pp.448-479.

Harris, A., McGregor, J., Perencevich, E., Furuno, J., Zhu, J., Peterson, D., & Finkelstein, J. (2006). The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics. *Journal of the American Medical Informatics Association, 13*(1), 16-23.

Helfaya, A. and Whittington, M. (2019) 'Does designing environmental sustainability disclosure quality measures make a difference?', *Business strategy and the environment.*

Hildebrandts, H.W. and Snyder, R.D. (1981). 'The pollyanna hypothesis in business writing: Initial results, suggestions for research'. *Journal of business communication*, 18(1), pp.5-15.

Hummel, K. and Schlick, C. (2016). 'The relationship between sustainability performance and sustainability disclosure – Reconciling voluntary disclosure theory and legitimacy theory', *Journal of Accounting and Public Policy*, 35(5), pp.455-476.

Krasodomska, J. and Cho, C.H. (2017). 'Corporate social responsibility disclosure', *Sustainability Accounting, Management and Policy Journal*, 8(1), pp.2-19.

Lang, M. and Stice-Lawrence, L. (2015). 'Textual analysis and international financial reporting: Large sample evidence', *Journal of accounting and economics,* 60(2-3), pp.110-135.

Lee, M.R. and Blanchard, T.C. (2012). 'Community attachment and negative affective states in the context of the BP deepwater horizon disaster.(Initial Perspective on the 2010 BP Deepwater Horizon Disaster).(Author abstract)', *American Behavioral Scientist*, 56(1), p.24-47.

Leung, S., Parker, L. and Courtis, J. (2015). 'Impression management through minimal narrative disclosure in annual reports'. *The British accounting review*, 47(3), pp.275-289.

Leuz, C., & Wysocki, P. (2016). 'The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research', *Journal of Accounting Research, 54*(2), 525-622.

Lyon, T.P. and Maxwell, J.W. (2011) 'Greenwash: Corporate environmental disclosure under threat of audit', *Journal of Economics and Management Strategy*, 20(1), p.3(39)

Marra, A., Mazzola, P., & Prencipe, A. (2011). Board Monitoring and Earnings Management Pre- and Post-IFRS. *The International Journal of Accounting, 46*(2), 205-230.

Melloni, G. (2015). 'Intellectual capital disclosure in integrated reporting: an impression management analysis', *Journal of intellectual capital*, 16(3), pp.661-680.

Melloni, G., Stracchezzini, R. and Lai, A. (2016). 'The tone of business model disclosure: an impression management analysis of the integrated reports'. *Journal of management & governance*, 20(2), pp.295-320.

Melloni, G., Caglio, A., and Perego, P. (2017), 'Saying more with less? Disclosure conciseness, completeness and balance in integrated reports', *Journal of accounting and public policy,* 36(3), pp.220-238.

Merkl-Davies, D.M., Brennan, N.M. and Mcleay, S.J. (2011). 'Impression management and retrospective sense-making in corporate narratives: a social psychology perspective'. *Accounting, auditing and accountability*, 24(3), p.315-344.

Meyer, J.W. and Rowan, B. (1977). 'Institutionalized Organizations: Formal Structure as Myth and Ceremony', *American Journal of Sociology, 83*(2), pp.340-363.

Mi, J., Sheng, H. and Elrod, C. (2015). 'Financing, reputation and information disclosure quality in the Chinese market', *Applied Economics*, p.1-10.

Michelon, G. (2011). 'Sustainability Disclosure and Reputation: A Comparative Study', *Corporate Reputation Review,* 14(2), p. 79.

Nègre, E., Verdier, M-A., Cho, C.H. and Patten, D.M. (2017). 'Disclosure strategies and investor reactions to downsizing announcements: A legitimacy perspective'. *Journal of Accounting and Public Policy*. 36(3), pp.239-257.

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.

Raiborn, C.A., Butler, J.B. and Massoud, M.F. (2011) 'Environmental reporting: Toward enhanced information quality'. *Business Horizons*, 54(5), pp.425-433.

Shah, Syed Zulfiqar Ali, Liang, Shuang, & Akbar, Saeed. (2013). International Financial Reporting Standards and the value relevance of R&D expenditures: Pre and post IFRS analysis. *International Review of Financial Analysis, 30*, 158.

Smithsonian. (2019). *Gulf oil spill*. Retrieved from, https://ocean.si.edu/conservation/pollution/gulf-oil-spill, 2019-01-09.

Summerhays, K. and de Villiers, C. (2012). 'Oil company annual report disclosure responses to the 2010 Gulf of Mexico oil spill', *Journal of the Asia-Pasific Centre for Environmental Accountability*, 18(2).

Sydserff, R., & Weetman, P. (2002). Developments in content analysis: A transitivity index and DICTION scores. *Accounting, Auditing & Accountability Journal, 15*(4), 523-545.

The Guardian. (2018). *BP's Deepwater Horizon bill tops $65bn*. Retrieved from, https://www.theguardian.com/business/2018/jan/16/bps-deepwater-horizon-bill-tops-65bn, 2019-01-10.

Thiese, M. (2014). Observational and interventional study design types; an overview. *Biochemia Medica, 24*(2), 199-210.

Thorpe, A., Craig, R., Hadikin, G., & Batistic, S. (2018). Semantic tone of research 'environment' submissions in the UK's Research Evaluation Framework 2014. *Research Evaluation, 27*(2), 53-62.

Verbeek, M. (2017). *A guide to modern econometrics* (5.th ed.). Hoboken, NJ: Wiley Custom.

Verrecchia, R.E. (2001) 'Essays on disclosure', *Journal of accounting and economics*, 32(1), pp.97-180.

Vourvachis, P., Woodward, T., Woodward, D.G. and Patten, D.M. (2016). 'CSR disclosure in response to major airline accidents: A legitimacy-based exploration', *Sustainability Accounting, Management and Policy Journal,* 7(1), pp. 26-43.

Wegener, M., & Labelle, R. (2017). Value Relevance of Environmental Provisions Pre- and Post-IFRS. *Accounting Perspectives, 16*(3), 139-168.

Wisniewski, & Yekini. (2015). Stock market returns and the content of annual report narratives. *Accounting Forum, 39*(4), 281-294.