

[This is the final version of the article that is included in my PhD thesis. NB! In the PhD thesis table 5 and 6 were missing, and they are included here.]

To cite: Fallan, E. (2013). *The representativeness of the annual report as data source in CSR reporting research*. In E. Fallan: Issues on supply and demand for environmental accounting information. LIMAC PhD School, Department of Accounting and Auditing, Copenhagen Business School, PhD Series 41-2013, pp. 21-60.

THE REPRESENTATIVENESS OF THE ANNUAL REPORT AS DATA SOURCE IN CSR REPORTING RESEARCH

Even Fallan¹,

(e-mail: ibn@turmusic.no)

ABSTRACT

Purpose: Selection of data source is an important methodological issue related to validity. What data sources are used in CSR reporting research, and how is the choice explained? The diffusion of internet and separate CSR reports has made this issue even more relevant. What does this mean to the representativeness of annual reports relative to total disclosure using all media types? Is it valid to use the annual report as the only data source in CSR reporting research?

Design/methodology/approach: A comprehensive literature study is conducted to examine previous research practices. Content analysis – 13 information content categories, each for two datasets – is employed to collect data on environmental (ENV) and working environment/human resources (WEHR) disclosure, respectively. Both datasets consists of companies listed on Oslo Stock Exchange (OSE). Annual reports, separate reports, and websites are used as proxies for total disclosure. Annual report disclosure and total disclosure is compared for each content category for each company, and the differences are aggregated to reveal the proportion of total disclosure covered by annual reports.

Findings: The annual report has been and still is the data source of choice in research. , but the use of several sources increases. Empirical studies of the representativeness of data sources are rarely used to guide the choice of source. Annual reports include approximately all information content of total disclosure provided in all disclosure media, irrespective of CSR theme and industry, and whether the information is mandatory or

¹ The author gratefully acknowledges the constructive comments from Jesper Møller Banghøj and Thomas Riise Johansen, Copenhagen Business School; Philip Shrivess, Newcastle Business School; Katerina Hellström, Stockholm School of Economics; Lars Fallan, Trondheim Business School; and participants at conferences and staff seminars at Copenhagen Business School; Newcastle University Business School, and Trondheim Business School.

voluntary. Information content is disconnected from volume of disclosure, indicating that either of these two measures cannot be used as a proxy for the other.

Research limitations/implications: The annual report can be used as a proxy for total disclosure of information content, and as the only data source in ENV and WEHR disclosure research aiming to address frequently asked research questions.

Originality/value: This comprehensive study provides timely guidance on data source selection, based on the premise that information content (what companies are saying) is more important than volume of disclosure (how many pages they use to say it). The results and implications are clear, and mostly in contrast to previous studies, which are found to be of limited current relevance to this research question due to issues of timeliness and/or material weaknesses in the research design.

INTRODUCTION

“In an era when companies produce stand-alone reports reflecting aspects of their environmental performance and/or social impact, future studies focusing exclusively on annual reports might not produce particularly relevant results” (Unerman, 2000:674).

Companies use various communication media to disclose corporate social responsibility (CSR) information: annual reports, websites, separate CSR reports, press releases, advertisements, brochures etc. Which media do users have to choose to get a representative picture of the information content of total disclosure? This knowledge is important for users in order to get the information easily and in a cost effective way. The importance is illustrated by an extreme case: data source selection in CSR reporting research. Researchers might face a cost-benefit trade-off between the need for large amounts of valid data and feasible data collection. The representativeness of collected data, compared with total disclosure in all media, increases with each additional data source containing unique information. However, content analysis – the most common data collection method for such disclosure (Milne and Adler, 1999) – is extremely resource-demanding. Hand collected data generally causes samples to be quite small (or data collection to be expensive). Hence, it would be advantageous to include as few data sources as possible while keeping up representativeness. In addition to the representativeness, a cost-benefit perspective of researchers’ data source selection will also consider other characteristics: it is easier and less time consuming to collect longitudinal and historic data from annual reports than from websites; the annual report is a clearly defined document, whereas the website content is more difficult to define; while annual reports are issued regularly, the content of websites is often not dated; etc. From most criteria, the annual report stands out as the obvious candidate, if only one medium is to be chosen. This paper focuses on the methodological issue of data source selection in research, aiming to answer the

question: Is the annual report representative of the total disclosed CSR information content in all relevant media?

The paper is motivated by a comprehensive literature study of CSR disclosure research. The process gave a clear indication that total disclosure is used in or is relevant for most published studies of CSR reporting, either directly – as a measure of disclosure practice – or as a proxy for, e.g., CSR performance. Annual reports have been, and still are, the most used data source in CSR reporting research. Until the 1990s, it was often the only source as well, but is changing. However, research papers seldom refer to empirical data of representativeness to substantiate the choice of data sources. This might partly be due to the fact that research containing findings has obvious methodological weaknesses limiting their relevance. This is illustrated by Tilt's (2008) call for website studies. These methodological challenges are reviewed to improve the current research design.

In this paper, annual reports, separate CSR reports² and websites are used as a proxy for total disclosure, which is compared to annual report disclosure. A common weakness in previous research is that volume is often used as a proxy for information content in CSR reporting. If volume and content of disclosure are disconnected, the conclusions from these studies may be questionable.

This study contributes to previous research by showing the lack of consideration of empirical evidence in data source selection, and by providing both a more timely and methodologically improved advice on annual reports' representativeness of total disclosure than previous research. While there has been considerable attention towards reliability issues concerning content analysis (Milne and Adler, 1999), this paper contributes by addressing validity.

MEDIA USED FOR CSR DISCLOSURE

The annual report is an important medium for CSR disclosure. Over the recent decades, it has become more comprehensive, in volume, information content etc. (Lessem, 1979, Tinker and Neimark, 1987, Tinker et al., 1991, Adams and Harte, 1998, Ljungdahl, 1999, Unerman, 2000, Deegan et al., 2002, Beattie et al., 2008, and Fallan and Fallan, 2009). However, companies are using several media for CSR disclosure (Zéghal and Ahmed, 1990, Unerman, 2000). Lessem (1979) found that American companies started issuing separate CSR reports in the 1970s. Worldwide, an increasing frequency in issuance of such reports has been apparent from the early 1990s until at least 2008

² The terminology 'separate CSR reports' or 'separate reports' is used here for reports specifically devoted to one or several issues concerning CSR, issued under various names such as community-, environmental-, sustainability-, HSE reports, etc.

(Ljungdahl, 1999, ACCA, 2004, KPMG, 2002, Frost et al., 2005, KPMG, 2005, KPMG, 2008). Though, there are indications of reduced use in some countries at a relatively high reporting level – e.g., USA Central- and Northern Europe (KPMG, 2002, KPMG, 2005, Fallan and Fallan, 2009). Separate reports are quite common among the largest companies, but not generally. They are not even always issued on an annual basis. Therefore, the most important reporting innovation, media-wise, since the annual report, is probably the internet. Following the launch in 1991, the World Wide Web was estimated to have about 30 million users in 1995, 361 million in the year 2000, and 2.4 billion by year end 2010³. The number of websites grew from 130 in early 1993 to 320 million in 1998 (Patten and Crampton, 2003). Estimates suggest that it is currently a to-digit billion number⁴. Studies show that a large and growing proportion of large and listed (Western) companies are disclosing CSR information on the websites (Del Bosco, 2004, Jose and Lee, 2007, Del Bosco, 2011, Moroney et al., 2011).

To sum up, companies use several media for CSR disclosure. According to KPMG (2011) only 10% of the 250 largest corporations in the world rely solely on the annual report. All the available communication channels, and especially the diffusion of the internet and separate reports, questions the validity of using the annual report alone as a proxy for total disclosure in CSR reporting research. Claims of insufficiency can be read in explanations for the choice of data sources in academic papers and heard in conference presentations. However, that does not in itself prove anything about the annual reports' representativeness of total disclosure from all media. The next section addresses the data source selection in CSR disclosure research, and empirical research on representativeness of data sources, as a basis for hypotheses development.

LITERATURE REVIEW

A literature study is conducted to explore the selection of data sources in content analysis studies of CSR disclosure over time. Research databases and reference lists are used to identify articles. Some early period works were hard to obtain, resulting in some missing data. Articles addressing the current research question were deliberately sought, which might give a small bias in the results presented in Tables 1-3. Influential papers from journals are also included. All the papers are part of the CSR disclosure research population. The final sample consists of 116 papers, excluding missing data. 107 articles are collected from 38 journals. The rest are book chapters etc. Accounting Auditing and Accountability Journal (AAAJ) and Accounting, Organizations and Society (AOS) are represented with 20 and 13 articles, respectively.

³ www.internetworldstats.com/stats.htm (date: 13-8-2013)

⁴ <http://www.worldwidewebsite.com> (date: 12-4-2012)

Which data sources a researcher should select depends on the research question, the analysis and the desired degree of certainty with which conclusions are reported in each individual study. In studies looking at e.g. mandatory CSR reporting, the selection is confined to a few sources identified in, e.g. the Accounting Act, etc. However, the literature study indicates that a large proportion of studies use, or should ideally have used, a measure of total disclosure as variable. This is the perspective of this paper. This requires knowledge of how to capture total disclosure.

Selected data sources in CSR reporting research

Table 1 indicates that the annual report has been and still is the data source of choice in CSR disclosure research. The first (or the only) number in each cell is the proportion of papers published in a given time period that uses each medium as data source. This result is supported by observations in many studies, though without backing of empirical findings (Wiseman, 1982). The main change in this period is that, while the annual report was frequently the single source until the 1990s, the use of websites and separate CSR reports, and especially the use of more than one data source, has increased since 2000. The proportion of the times each medium was used in a time period where it was the only data source, is shown in brackets. Developments illustrated in Table 1 are supported by other data (Fifka, 2012, Fifka, 2013). The observed change in the use of data sources in research might be a response to the increased use of websites and separate reports as media in corporate reporting. This assumption is supported by claims in several papers (Aerts et al., 2008, Clarkson et al., 2011). It supports the view that focus on the methodological challenge of valid data source selection is highly relevant.

Table 1: The selection of data sources in CSR reporting research

Year	1975-1979	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2009	2010-2013	Whole period
No. of papers reviewed	12	9	8	14	18	19	25	11	116
Annual report ⁵	100% (83%)	100% (78%)	100% (100%)	93% (69%)	100% (83%)	89% (65%)	88% (36%)	82% (67%)	93% (69%)
Separate report	0%	22% (0%)	0%	14% (0%)	17% (0%)	32% (0%)	52% (0%)	45% (0%)	27% (0%)
Web	0%	0%	0%	0%	6% (0%)	21% (50%)	36% (22%)	36% (0%)	16% (22%)
Advertisements	0%	0%	0%	7% (0%)	0%	0%	4%	0%	2% (0%)
Brochures	8% (0%)	0%	0%	14% (0%)	0%	0%	4%	0%	3% (0%)
Other	8%* (0%**)	11% (0%)	0%	21% (33%)	0%	16% (0%)	12%	9% (0%)	10% (8%)

* 8% of the 12 reviewed papers published in this time period use "Other" as data source.

** 0% of the 12 reviewed papers use "Other" as the only data source.

⁵ "10-K reports", filed with the Security and Exchange Commission in the US, and equivalents in other countries, are included in the annual report numbers in this table.

Discussion or justification of data source selection

The literature study reveals whether or not papers explain why the chosen data sources are selected. There are many potential reasons. Kuasirikun and Sherer (2004:635) chose the annual report due to its “credibility; usefulness to various stakeholders; regularity; accessibility and completeness in terms of the company’s communication on social issues”. Quite a few papers refer to the choice made by (the majority of) other researchers as an argument (Gray et al., 1995, Kuasirikun and Sherer, 2004). Because the majority of studies were assumed to have used annual reports, the possibility of comparison with other studies turns up as a reason in itself (Deegan et al., 2002). Practical, pragmatic, or economically feasible considerations are also emphasized: e.g. that annual reports are available in English whereas separate reports where not consistently so (Beck et al., 2010), or that it is impossible to identify all media used for disclosure by each individual company in a sample, so it is easier to use just one data source (Gray et al., 1995). Another topic addressed in some papers is the audience of CSR disclosure in different media (Neu et al., 1998, De Villiers and Van Staden, 2011). The upper result-row in Table 2 shows that it has become common to explain why data the source(s) is chosen. A quarter of the papers did so in the 1970s, while about 80% did after the millennial-change.

Table 2: *The proportion of papers explaining their choice of data source(s)*

Time period	1975-1979	1980-1984	1985-1989	1990-1994	1995-1999	2000-2004	2005-2009	2010-2013	Whole period
No. of papers reviewed	12	9	8	14	18	19	25	11	116
It is explained why the data source is selected	25%*	78%	38%	50%	67%	84%	80%	82%	66%
Several media are discussed when data source selection is explained	17%**	33%	13%	43%	50%	79%	56%	73%	50%

* 25% of the 12 reviewed papers published in this time period explain why the chosen data source(s) is (are) selected.

** 17% of the reviewed papers discuss why this (these) data source(s) is (are) preferred over other relevant data sources, e.g. that the chosen source is representative of total disclosure, or that other sources are hard to obtain.

The bottom row of Table 2 shows whether or not the reasoning behind the choice of data sources consider several (also non-selected) media, e.g. compare characteristics, the extent of use, and/or consequences of including or excluding different media. Freedman and Jaggi (1982:169) is critical to the use of the annual report as the only data source because “total disclosures may not have been examined”, and Niskala and Pretes (1995:459) argues that “[r]estricting a study to annual reports only may give an incomplete view of overall reporting”. Such judgements are not equally common as the first results reported in Table 2, although there has been an increase in the level over the years.

In order to have value, data must be representative of the study’s objective. Therefore, the representativeness of selected data sources relative to total disclosure is an issue that often should be relevant in Table 2. Such methodological considerations might include at least two types of arguments: selection of data source(s) can be based on beliefs – e.g. Freedman and Jaggi (1982) – and/or guided by empirical findings from research – e.g. Niskala and Pretes (1995). The literature study has tried to identify the papers that provide empirical findings on the representativeness of data sources compared to total disclosure. Table 3 reveals how common it is to use these findings in research decision-making. The first number in each cell is the proportion of papers published in that time period that is referring to and discussing the relevant recommendations in the listed paper. Surprisingly this number has not been above 37% in any time period. As from 2005, when diffusion of website disclosures has come far and Table 2 shows that researchers are well aware of the possible methodological challenge, less than 20% of the papers make use of guidance from empirical findings when they choose data source.

Table 3: Citation of articles containing guidance on data source selection

Time period →	No. of papers reviewed	1991-1994	1995-1999	2000-2004	2005-2009	2010-2013	Average
Papers containing guidance on data source selection ↓							
Zeghal and Ahmed (1990)	86***	8%* (23%**)	28% (50%)	16% (37%)	8% (20%)	9% (18%)	14% (30%)
Williams and Pei (1999)	55			0% (5%)	4% (24%)	0% (18%)	1% (16%)
Unerman (2000)	50			29% (43%)	8% (28%)	0% (18%)	12% (30%)
Patten and Crampton (2003)	41			0% (0%)	0% (16%)	0% (27%)	0% (14%)
McMurtrie (2005)	32				0% (0%)	9% (9%)	5% (5%)
Tilt (2008)	15				0% (0%)	0% (0%)	0% (0%)
SUM****	86	8% (23%)	28% (50%)	37% (53%)	16% (48%)	18% (55%)	21% (46%)

* 8% of the reviewed papers in this period have cited Zeghal and Ahmed’s (1990) advice concerning data source selection.
 ** 23% of the reviewed paper in this period have cited Zeghal and Ahmed (1990), irrespective of whether the reference concerns data source selection. This is an illustration of the knowledge of Zeghal and Ahmed (1990).
 *** 86 papers were reviewed in order to examine if they cite Zeghal and Ahmed (1990). These 86 are the reviewed papers that are published at least one year after Zeghal and Ahmed (1990).
 **** The proportion of reviewed papers that cite at least one of the left column articles in this period. Each reviewed paper is counted maximum once per time period, regardless of how many of the left column articles it quoted

Discussing the low use of empirical findings to guide data source selection

Why are the percentages so low? It is reasonable to assume that most CSR disclosure researchers know about at least one of the papers listed in Table 3. Firstly, because Table 3 indicates that the knowledge of the empirical findings is higher than the use – papers are also referring to other parts of the listed papers than the guidance relevant here. Numbers in brackets are the proportions of papers referring to each of the studies listed in Table 3, irrespective of whether or not they refer to

the advice on data source selection. These numbers are in most cases much higher than the first number in the same cell. Secondly, because researchers do not refer to all studies they know about (every time), and thirdly, because some of the papers listed in Table 3, e.g. Unerman (2000), are well known in CSR reporting research. Additionally, influential papers point out the importance of the selection of medium for the validity of studies (Unerman, 2000, Campbell, 2004), and it is frequently discussed (Table 2), so the issue is probably not perceived as insignificant either. But if researchers know about the empirical findings on a significant matter, why not just refer to it – the easiest, most space-saving and powerful way to back the choice of data source(s)?

Two possible explanations are related to the advice of the existing findings. Researchers might perceive the findings to be mixed, so no clear advice appears. The results of some papers indicate that the use of one data source (the annual report) is sufficient [e.g. Niskanen and Nieminen (2001) and Tilt (2008)], while others might seem to promote the use of several sources [e.g. Zéghal and Ahmed (1990); Unerman (2000); Campbell et al. (2003); McMurtrie (2005); and De Villiers and Van Staden (2011)]⁶. Nevertheless, mixed findings should not keep researchers from referring to it. It is important to emphasize even a potential uncertainty (which actually might be an argument in favour of using several data sources), and some findings should be better than confining only to beliefs. Another explanation for the lack of references to existing findings might be that its advice does not suit researchers. The practical advantages of using only the annual report are large compared with using several or other media. Table 1 reveals that the annual report frequently has been the only data source in CSR reporting research, while most of the papers listed in Table 3 (have been interpreted by many to) suggest that several data sources should be used in order to cover total disclosure. Additionally, the literature study showed that only 33% of papers referring to the guidance in the three oldest studies (Zéghal and Ahmed, 1990, Williams and Pei, 1999, Unerman, 2000) – which most papers in the study had the chance to know – follow the advice of at least one of the studies they quote. Campbell et al. (2003:566) uses the annual report as the only data source for the main analysis, despite stating that the “[a]rguments against the selection of only the corporate report are, however, persuasive” because of a reference to the conclusion of Unerman (2000) that the annual report should not be considered a reliable proxy. A potential disparity between the advice from empirical findings and researchers` wish to use a certain data source might explain why so few papers refer to the existing guidance. Still, this does not explain why there has been a change in

⁶ These classifications (mainly stemming from the papers themselves) are disputable, because it depends on how representativeness of total disclosure is measured – as will be discussed below. This is also the reason why the guidance of seven papers is not indicated here.

direction of using several data sources in research. Additionally, it cannot explain why researchers wanting to use only the annual report do not even quote findings in favour of this.

The review performed to clarify the advice concerning selection of data sources, revealed what might be the most important reason why few papers refer to existing findings: The studies have potentially major weaknesses regarding the research design and (current) timeliness of data. There are also large variations in data, analyses and findings, and with that the advice for data source selection, between them. These weaknesses might cause researchers to perceive the empirical evidence as less relevant. A discussion of strengths and challenges in these papers is imperative to improve the research design of the current study. The 15 papers listed in Table 3 are presented in Appendix C.

Strengths and weaknesses in existing empirical findings

All the 15 papers might provide useful input to data source selection, even though they are not equally relevant. This variation in relevance is partly due to the studies' objective. Providing guidance on data source selection from the user-perspective is the objective of one paper only; a few papers provide data on the media selection from the reporting company-perspective; others have a different main objective, but answer more or less a similar research question on the way; while some provide interesting data in the process of addressing other questions. The papers are grouped according to this classification in Table 4. Papers in the three first categories should have an adequate research design but some methodological weaknesses are revealed partly due to development in the state of the art of research at that time.

A basic challenge is the type of data that should be used to measure a data source's representativeness of total disclosure in all relevant media. All the 15 papers in Table 4 confirm that companies use several media to disclose CSR information. This means that a single medium, such as the annual report, constitutes just a part of the total *volume* of CSR disclosure. However, it seems intuitively unlikely that the volume of disclosure in itself an important aspect for *users* of disclosure. Research on use (and impact) of reporting suggests that volume is not important compared to quality and content of disclosure (Hasseldine et al., 2005, Brammer and Pavelin, 2006). This paper argues that *what* is said/reported (the information content of disclosure) is more relevant to users than how many times it is said or how many words, sentences, or pages that is used to say it (volume of disclosure). 10 of the papers in Table 4 use volume of disclosure, of which it is a main measure in five papers. In four of these it is the basis for supporting the need for several data sources in four (Unerman, 2000, Campbell et al., 2003, McMurtrie, 2005, De Villiers and Van Staden, 2011) and the opposite in one (García-Ayuso and Larrinaga, 2003). A claim in this paper is that those conclusions

might be based on non-valid measurement. Volume of disclosure should not be used as a proxy for content of disclosure (which suggests that higher volume means a broader variety of content) either, due to the measures being disconnected and lower precision of analyses – as discussed below. When content of disclosure is the most relevant type of data, why not measure it directly?

Content of disclosure is a better measure than volume. However, how content is operationalized is also relevant. Two papers specify the content only as CSR disclosure (Unerman, 2000, Campbell et al., 2003), while six other specify which general CSR theme it is – e.g. human resources, environment or community (Zéghal and Ahmed, 1990, Clarke and Gibson-Sweet, 1999, Niskanen and Nieminen, 2001, García-Ayuso and Larrinaga, 2003, McMurtrie, 2005, De Villiers and Van Staden, 2011). CSR is a collective term with limited meaning in itself without further specification. Is it valuable for users to know that a company has disclosed something about CSR or even the environment, without knowing more about it? Probably not. At least three of these papers reveal at least whether the information is positive or negative, or monetary, quantitative or narrative (Zéghal and Ahmed, 1990, Niskanen and Nieminen, 2001, García-Ayuso and Larrinaga, 2003). There is a need for more details in order to make data relevant. Some of the 15 papers do this by using a set of content categories (and even subcategories) or other information characteristics as a deepening of each general CSR theme they study. In the same way as for volume of disclosure, the categorisation of information as only CSR disclosure or a general CSR theme (e.g. environment) is unlikely to provide a valid measurement for representativeness of (information content) of data sources. Table 4 reveals the 15 papers containing empirical findings, classified according to both their objective and their type of data used to measure representativeness of data sources relative to total disclosure. Papers in the top left corner are most relevant for this study, based on these two criteria in isolation.

Table 4: Studies classified by their objective and measurement relevant for data source selection

Objective → Measurement ↓	Papers whose objective is to provide advice/ empirical findings on data source selection in research (user’s perspective)	Papers providing empirical evidence on media selection from the reporting company-perspective (and hence provide the same information)	Papers with a different main objective, that answer more or less a similar research question (from the reporting company-perspective) on the way	Papers providing interesting data in the process of answering other questions
Content categories, other information characteristics, and volume of disclosure		<ul style="list-style-type: none"> Tilt (2008) 	<ul style="list-style-type: none"> Patten and Crampton (2003) Williams and Pei (1999) 	<ul style="list-style-type: none"> Frost (2007)
Content categories and other	<ul style="list-style-type: none"> THIS STUDY 			<ul style="list-style-type: none"> Moroney et al. (2011) Aerts et al. (2008)

information characteristics				
Content categories			<ul style="list-style-type: none"> • Branco and Rodrigues (2008) 	
General CSR themes, other information characteristics, and volume of disclosure		<ul style="list-style-type: none"> • Zéghal and Ahmed (1990) 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> • García-Ayuso and Larrinaga (2003)
General CSR themes and other information characteristics				<ul style="list-style-type: none"> • Niskanen and Nieminen (2001)
General CSR themes and volume of disclosure		<ul style="list-style-type: none"> • De Villiers and Van Staden (2011) 	<ul style="list-style-type: none"> • McMurtrie (2005) 	
General CSR themes				<ul style="list-style-type: none"> • Clarke and Gibson-Sweet (1999)
CSR, other information characteristics, and volume of disclosure			<ul style="list-style-type: none"> • Campbell et al. (2003) 	
CSR and volume of disclosure	<ul style="list-style-type: none"> • Unerman (2000) 			

The next methodological issue to be addressed is the low precision level of the chosen analyses. Many of these papers collect data separately for each data source on a company level, then aggregate the data to sample level for each data source separately, so the results for the data sources can be compared on sample level (Clarke and Gibson-Sweet, 1999, Williams and Pei, 1999, Patten and Crampton, 2003, Branco and Rodrigues, 2008). The consequence is that the reader knows that e.g. 50% of the sample has disclosed a type of information in annual reports and ditto on the website, but cannot tell whether it is the same companies that are disclosing in both media while others do not disclose at all (both data sources are, individually, representative of total disclosure), or if each company is using only one data source (both data sources must be selected), or a mix (the representativeness of each data source must be compared to the requirement of that study). Such analyses have a low precision-level, and do hardly provide valid data for the current research question. The same is more or less the case with mean based statistics (De Villiers and Van Staden, 2011). Computed disclosure scores per company, made up of many types of content and other information characteristics, being analysed on a sample level is even less valid for evaluating representativeness (Aerts et al., 2008, Moroney et al., 2011). Frost (2007) has carried out a sample level comparison of totals (or averages) as described above, but has disclosure data for only one data source while controlling for the existence of another. Campbell et al. (2003) use different types of

data for the two data sources, which complicate the analysis. Tilt (2008) does not present much of the results supporting her 1994-conclusion, so it is hard to tell what kind of analysis that is performed. It should be noted that the *opportunity* for within company comparisons is provided by five papers. They have either conducted company level analysis or presented data so the reader can do it (Zéghal and Ahmed, 1990, Unerman, 2000, Niskanen and Nieminen, 2001, Campbell et al., 2003, McMurtrie, 2005). The challenge is the lack of papers that satisfy all the requirements described in this section simultaneously.

A common feature of the five articles using company level comparisons is small sample sizes. Four of the studies include five companies or less. Then it is difficult to know whether the results are generally valid, even within a small population like an industry. (Some of them are not rich on detail and context like a case study either – on this issue). The fifth paper has a clever stratification of the 15 companies sample, so it seems likely that parts of two industries are well represented (Zéghal and Ahmed, 1990).

Another issue of concern is that the sampling criteria of at least seven of the 15 studies lead to underestimation of the representativeness of the information content of annual reports relative to total disclosure from all media. For two studies this is done by selecting only companies that disclose a lot of CSR information and have a CSR policy (McMurtrie, 2005, Tilt, 2008). The likelihood of disclosing unique information in several media is (statistically) higher for companies that disclose much CSR information than for those who hardly disclose anything. This effect is enhanced for annual reports, since it is likely to be among the first media to be used for disclosure, due to regulations, reporting practice, its specific status etc. Moroney et al. (2011) is similar because they selected all companies that have audited or verified environmental reporting as a part of the sample, and do not separate the result between companies according to that characteristic. The same bias arises from selecting only companies that issue separate reports and do not disclose information in annual reports (Niskanen and Nieminen, 2001), with selection of companies with bad environmental performance and environmental crises (De Villiers and Van Staden, 2011), and by selecting companies that are most likely to have reporting requirements due to environmental regulations (Frost, 2007). Tilt (2008) uses a sampling criterion more directly related to data sources as well. By excluding companies that disclose information in less than two media, representativeness of any data source is underestimated. Other papers include only companies that have a website in their sample (Williams and Pei, 1999, Patten and Crampton, 2003, Branco and Rodrigues, 2008). Companies without a website are obviously less likely to have unique information content outside annual reports than those who do. According to Williams and Pei (1999) below 10% of the listed companies in those countries had a website at the time of their study. Hence, the potential

underestimation of the representativeness of annual reports in this population is large. In Campbell et al. (2003) it took at least 2.5 years from the issuance of annual reports to the website study, meaning that a lot of new information reflecting a different reality might have appeared on the websites, and that differences are more likely to exist.

Seven of the 15 papers do not include website disclosures in their dataset (Zéghal and Ahmed, 1990, Clarke and Gibson-Sweet, 1999, Unerman, 2000, Niskanen and Nieminen, 2001, García-Ayuso and Larrinaga, 2003, Frost, 2007, Tilt, 2008). Even if that was not problematic in 1981 (Zéghal and Ahmed, 1990), it is imperative to capture total disclosure today. Tilt (2008) calls for studies on representativeness of data sources that includes website disclosures. The dataset should capture companies' mature use of websites the last decade.

The final issue to be addressed here is the current low timeliness of the data in most of the papers. This is, for most papers, due to the time they were published, and not a choice made by the researchers. Nevertheless, it affects the relevance of the empirical findings. There has been major changes in CSR reporting and the proportion of companies that disclose such information since the 1981/2 data of Zéghal and Ahmed (1990). It is still a relatively young and developing practice. Additionally, the use of media has changed with the diffusion of the internet (and separate reports). Only three of the 15 papers use data that is less than 10 years old (2003) (Branco and Rodrigues, 2008, De Villiers and Van Staden, 2011, Moroney et al., 2011). Since the studies with quite timely data have other weaknesses, more studies with current data are called for.

Historically, two papers are especially important for this research question. Unerman (2000) was the first to clearly emphasize the selection of data source(s) as a methodological issue for researchers to consider, as opposed to focusing on the company perspective concerning the choice of disclosure media (Zéghal and Ahmed, 1990, Tilt, 2008, De Villiers and Van Staden, 2011). It meant a new focus on validity of research, in addition to reliability (Milne and Adler, 1999). However, as reporting practice and research has developed, its current relevance (due to research design and timeliness of data) is limited. Zéghal and Ahmed (1990) was probably the first paper to provide empirical data to consider this issue. The type of data they collected is still relevant, but they did not utilize the data in their analysis, and of course the data are not timely anymore. According to Table 3, these are also the two most frequently used references to empirical findings in discussions concerning selection of data sources.

The review of the existing evidence in the 15 identified studies clearly reveals both the need to and how to improve research (design) on the representativeness of data sources. That is the objective of this paper.

HYPOTHESES

Companies use multiple media to disclose CSR information. Researchers would benefit from knowing what disclosure media that is necessary to access to obtain all the disclosed information content (total disclosure). A cost-benefit approach suggests that it is easier and less costly if the content of a few media, or even a single medium such as the annual report, correspond to total disclosure.

Even if several media are used for reporting purposes, one medium would suffice as data source if it captures all disclosed information content. It would mean that volume of disclosure and information content is disconnected. Tilt (2008) supports this idea, while Zéghal and Ahmed (1990) do not. Since the previous section suggests a cautious use of those findings, it will be addressed empirically in this paper. However, what is interesting here is the reasoning. A basis for rejecting one of these views is statistical probability: With a large volume of disclosure outside the annual report, and the broad variety of possible types of information content to report, it would appear statistically likely that at least some of the information content is unique to sources other than the annual report. Other arguments for why companies possibly disclose different information in different media can also be derived from the papers listed in Appendix C, e.g. regulation (mandatory versus voluntary disclosure) (Patten and Crampton, 2003, Aerts et al., 2008); industry differences (Zéghal and Ahmed, 1990, Patten and Crampton, 2003, Branco and Rodrigues, 2008); and that different stakeholders use different media (Zéghal and Ahmed, 1990). Hence, the representativeness of the annual report relative to total disclosed content is questioned. The last argument will be discussed next, while the rest are addressed through the subsequent general discussion and accompanying hypotheses to be tested.

Zéghal and Ahmed (1990:49/51) claim that companies deliberately disclose partly different information content in the three media examined because they are complementary sources of information with different target publics: The annual report is “directed primarily at the investment community”, “advertisements are aimed at a broader public”, while brochures are targeted at “specific interest groups”. However, such a simplified view of use(rs) are moderated by both researchers (Neu et al., 1998) and standard setters (IASB and FASB, 2006). There is not an unambiguous answer to the question of which media different stakeholders use to get public available information. The most likely answer is that stakeholders usually get information from a large number of media (Rowbottom and Lymer, 2009, Christensen, 2010, De Villiers and Van Staden, 2011, Fallan, 2013), even (indirectly) through communication with colleagues, friends and family, and (news-)media picking up on other media etc. Stakeholders are probably not aware of all sources themselves, or at least they do not deliberately choose all themselves. Then of course the reporting

companies do not know which specific media to use to reach different groups either. (Information specifically targeted to a defined, limited audience is often private information (Solomon and Solomon, 2006), and out of scope of this paper.) Stakeholders get information everywhere. Reporting companies understand that, and disclose similar information in a variety of communication channels.

Luckily, it is easier to say something about what kind of information that is disclosed in different media. Since companies often will choose to disclose the same information content in several media, information content is probably disconnected from volume of disclosure. That is an important argument for the hypotheses below. Nevertheless, the hypotheses are also built on another line of reasoning. This is related to the rejection of a basic, underlying assumption for the statistical argument above: disclosures in different media are not independent observations. An objective of reporting is to provide information of important aspects of the company's performance and status, risks and opportunities, products, the environment in which it operates etc. "To give a "true and fair view" or to "present fairly" a company's financial position and results of operations implies that there is some "economic reality" to be reflected in the accounts" (Zeff, 2012:3). The underlying reality to be described is the same regardless of which communication media that is used. Even though corporate reporting often emphasizes certain (positive) aspects of reality and perhaps leaves others out (Niskanen and Nieminen, 2001, Larrinaga et al., 2002), companies rarely lie or make up stories in their reporting (Evans III et al., 2001). Additionally, the reporting of other companies is known, common reporting practices has developed in industries over time, and companies do in fact often copy last year's reporting: the issues that are generally relevant in disclosure are known. Imitation of own and others reporting is leading to isomorphism (DiMaggio and Powell, 1983) concerning the use of media. Besides, the reporting process incurs costs, and it is cheaper to disclose the same information in several media than to produce tailor-made information for each. Based on this, it is likely that the information content of a company's disclosure in different media is quite similar within each year (and probably also between years).

The annual report is arguably the single data source whose information content is most likely to be representative of total CSR disclosure, if any are, both for mandatory and voluntary disclosures. Firstly, companies in many countries are required to disclose CSR information in documents incorporated in or closely related to the annual report. Secondly, it is natural to disclose voluntary information the same place as well. Thirdly, even information content that is otherwise voluntary are required and expected to be reported in the financial statements, the board of directors' report, and/or the management report if considered material to the company. Fourthly, one purpose of the annual report is to represent the previous year. With that comes naturally a perception of summary, relevant when comparing it with more instant media like press releases, websites/social media.

Fifthly, it should also be forward looking, to capture e.g. risks and opportunities. Sixthly, annual reports are also commonly perceived to have a special status in reporting, both among reporters and users (Tilt, 1994, Unerman, 2000). An aura of importance, credibility, widespread use, clearly defined function and document boundaries etc. keeps up this status even as new media such as websites diffuse. The internet has actually made annual reports more accessible. All these characteristics of annual reports also suggest that its representativeness will increase as the reporting practice matures and assumes qualities of traditional reporting.

The general hypothesis (H) of the paper is that information content is disconnected from volume of disclosure:

Approximately all disclosed information content in all media (TD = total disclosure) is present in the annual report (= AR).

H: AR \approx TD

The strong reasoning supporting the hypothesis suggests that what might potentially appear as mixed findings in previous research are due to the methodological weaknesses of former studies and the low timeliness of the data. Improved research design and more timely data will provide more valid results. A consequence is that several sub-hypotheses should be derived to make the results more robust.

It is necessary to go beneath a general CSR theme level and examine content of disclosure on lower category level to make the analysis meaningful. A robust analysis on CSR disclosure allows for comparison of more than one CSR theme, to consider whether results are indifferent of these specified themes. Some of the existing findings might suggest that companies use disclosure media differently for different CSR themes: e.g. that the annual report is relatively important for working environment/ human resources disclosure (WEHR) (Zéghal and Ahmed, 1990, Branco and Rodrigues, 2008), while environmental disclosure (ENV) is relatively more frequent outside the annual report (Zéghal and Ahmed, 1990, Clarke and Gibson-Sweet, 1999, Williams and Pei, 1999, Branco and Rodrigues, 2008). This is challenged by the above reasoning of the general hypothesis. Therefore, separate hypotheses are formed for these two themes (the selection of which is further described in the methodology section):

Approximately all disclosed environmental and working environment/ human resources information content, respectively, in all media (TD) is present in the annual report (AR).

$H1_{ENV}: AR_{ENV} \approx TD_{ENV}$

$$H1_{WEHR}: AR_{WEHR} \approx TD_{WEHR}$$

Another issue that might affect the use of disclosure medium, and, hence, the selection of data sources, is regulation. Regulations usually concern the content of annual reports. While all disclosure is voluntary on websites, advertisements, press releases etc., most information content is voluntary even in annual reports (Fallan and Fallan, 2009). It is not controversial to claim that the annual report is probably representative of total disclosure for content that is mandatory there and voluntary elsewhere. However, it appears to be believed that the annual report is not representative of total disclosure of information content that is voluntary even in the annual report. "In our view, disclosure that is unique to the web is more likely to be of a voluntary nature" (Aerts et al., 2008:644). Nevertheless, in this present paper it is claimed that the above argumentation for the general hypothesis holds both for mandatory and voluntary information. Information content is made mandatory through law or accounting standards because it is deemed important. The most important information content should be the first to be disclosed in any medium, irrespective of regulations. From this four hypotheses are derived:

Approximately all disclosed *mandatory and voluntary* environmental and working environment/ human resources information content, respectively, in all media (TD) is present also in the annual report (AR).

$$H2_{ENV}: AR_{ENV} (\text{mandatory}) \approx TD_{ENV} (\text{mandatory})$$

$$H2_{WEHR}: AR_{WEHR} (\text{mandatory}) \approx TD_{WEHR} (\text{mandatory})$$

$$H3_{ENV}: AR_{ENV} (\text{voluntary}) \approx TD_{ENV} (\text{voluntary})$$

$$H3_{WEHR}: AR_{WEHR} (\text{voluntary}) \approx TD_{WEHR} (\text{voluntary})$$

Industry differences exist in CSR reporting (Fifka, 2013), though research on this concerns especially the extent of disclosure and not the use of media. Among the findings reviewed in this paper, Tilt (2008) states that companies in environmental sensitive industries are more likely to issue separate reports, but does not experience industry differences for information content between media. Others do potentially experience such differences with varying strength (Zéghal and Ahmed, 1990, Patten and Crampton, 2003, Aerts et al., 2008, Branco and Rodrigues, 2008), though a methodological reservation must be made. The hypothesis here is that the basic argumentation for equality of use of media will hold. It is examined by selecting "extreme" industries.

Approximately all disclosed environmental and working environment/ human resources information content, respectively, in all media (TD) is present in the annual report (AR) for both CSR sensitive and non-sensitive industries.

H4_{ENV}: AR_{ENV} (sensitive industry) \approx TD_{ENV} (sensitive industry)

H4_{WEHR}: AR_{WEHR} (sensitive industry) \approx TD_{WEHR} (sensitive industry)

H5_{ENV}: AR_{ENV} (non-sensitive industry) \approx TD_{ENV} (non-sensitive industry)

H5_{WEHR}: AR_{WEHR} (non-sensitive industry) \approx TD_{WEHR} (non-sensitive industry)

A description of the empirical data and analyses is described next, in order to test these hypotheses and make valid measurements of annual reports' representativeness of the content of total disclosure in all data sources.

METHODOLOGICAL ISSUES

Sample

Answering the hypotheses require data to have some specific characteristics. Two CSR themes must be selected (hypothesis 1), both must be subject to regulation so that both mandatory and voluntary information potentially can occur (hypotheses 2 and 3), and it must be possible to identify (at least) two industry strata (hypotheses 4 and 5).

Norwegian data is particularly relevant for this purpose. The Accounting Act includes reporting regulations of several CSR themes. The law requirements have been in force since 1989 and 1999, which means that reporting practice has had the chance to mature (Fallan and Fallan, 2009). Norway is world leading in internet use, measured as the "percentage of inhabitants using internet"⁷. This is important in this study because if annual reports are representative of total disclosure in Norway – a country where CSR disclosure has relatively long traditions and the use of multiple disclosure media is common – they are likely to be in countries with less use of internet as well. Websites are recognised as the other most important disclosure medium. The consequence is increased robustness for the current analysis, and hence increased probability for valid results.

Since some findings indicate that companies have an inclination to use disclosure media differently for the two CSR themes ENV and WEHR, and both are subject to reporting regulation in Norway, these seem to be good choices for the hypotheses, to increase the robustness of the analysis. By drawing the sample among firms listed on OSE, the law requirements are equal for all

⁷ www.internetworldstats.com/stats4.htm (date: 13-08-2013)

companies in the population. Considerations regarding stratification in two industries (hypotheses 4 and 5) might be different for ENV and WEHR disclosure. Therefore two datasets are used.

Dataset ENV (environmental disclosure) concerns the impact on the natural environment and related risks and opportunities of the company's organizational processes; production processes; or products, in a life-cycle perspective. It is more or less impossible to make a credible continuous scale rating of environmental risk/ impact. However, the risk/impact is perceived to be closely related to industry. In order to examine whether environmental risk matter for the choice of disclosure media, two strata of perceived high and relatively low risk companies are compared (hypotheses 4 and 5). Dataset ENV consists of 63 companies. The 17 companies in "industry" 1 are rated by the Norwegian Climate and Pollution Agency as businesses having the (potentially) most serious environmental risk. Industry 2 consists of 46 randomly selected companies classified by OSE as financials (including equity certificates) and IT according to the Global Industry Classification Standard (GICS). These industries are perceived to have relatively low direct environmental risk (Jose and Lee, 2007).

Dataset WEHR (WEHR disclosure) concerns issues like the atmosphere at the workplace; gender equality; health and safety; sick leave; equal pay, number of employees etc. Different types of work might affect disclosure. Manufacturing or construction work is probably more likely to experience serious accidents and injuries than office work. Under the assumption that the degree of (non-) office jobs varies somewhat with industry, all 39 OSE companies GICS-classified as industrials and all 36 classified as financials (including equity certificates) were included as sensitive and non-sensitive industries (hypotheses 4 and 5) respectively. Dataset WEHR constitutes totally 75 companies.

Selection of media

According to Krippendorff (1980) an essential stage in any content analysis study is deciding which documents are to be analysed. It is a large task to identify and retrieve all media containing CSR disclosure for the two datasets. Hand collection of content analysis data is extremely resource demanding as well. Together this is far beyond the limits of this study. It illustrates a reason why the objective of this paper is important. McMurtrie (2005:133) took the consequences:

"The study was limited to investigating only two companies because as the project evolved it became evident that the volume of disclosure to be examined was large and required a considerable amount of time to be analysed properly."

Instead of a reduction of sample size, like in McMurtrie (2005), it is decided to limit the number of media to be examined per company. Inclusion of the most important disclosure media is assumed to be a good approximation of total disclosure. Especially since websites incorporate many sources like

advertisements, product brochures and press releases, in addition to content published solely on the web. In dataset ENV, the aggregate of disclosure in annual reports, separate CSR reports and on websites is the proxy for total disclosure. In dataset WEHR, total disclosure is approximated by the content of annual reports and websites. Separate CSR reports are probably important for companies that issue them, but relatively few do. Hence, the law of large numbers makes it the medium to be dropped in dataset WEHR due to time and resource restraints. All the selected companies had their own website, and most companies had published the annual report there. For the rest of the companies the annual report was retrieved through the OSE website or by direct contact with the company. There were no missing data for the two samples.

Unit of analysis and operationalization of concepts

In content analysis the meaning of the text is coded from phrase or sentence (Milne and Adler, 1999), and registered in content categories. Qualitative variables are converted into quantitative variables based on qualitative measurement. A large effort is exercised to develop an appropriate method of gathering these data. The pre-defined categorization of information content of environmental disclosure is based on two main principles (Fallan and Fallan, 2009):

- (1) The list of categories should be complete, i.e. all relevant types of information content should be included in one of the categories;
- (2) The categories are mutually exclusive, i.e. one type of information content should only be included in one category and in no other.

The review of research clarified the need for detailed data on information content in order to improve validity. An adequately detailed list of relevant content categories is constructed for both dataset ENV (appendix A) and dataset WEHR (appendix B). The categorisations consider regulations, and are adequate irrespective of industry.

The categorization used in dataset ENV is similar to Fallan and Fallan (2009). It is an adaptation of Ljungdahl (1999), and is based on the thorough validity enhancing work of UNCTC (1991). It is later used in many studies. All the 13 content categories (appendix A) are voluntary disclosure on websites and in separate reports. In annual reports, categories 3, 4, and 13 are required by the Accounting Act, while the remaining categories are voluntary. A detailed description of the categories of dataset ENV is found in Fallan and Fallan (2009).

The content categories of dataset WEHR are developed by Andersen and Kirkeslett (2010)⁸, and a detailed description is found there. It was based on the logic of the reporting requirements in the Norwegian Accounting Act, the accompanying “Norwegian Accounting Standard 16 The board of directors’ report”, and the guidelines of the voluntary reporting standards Social Accountability International 2008 (SA 8000) and Global Reporting Initiative 2006. Validity is further enhanced through the use of a pilot phase to consider whether the categories seemed adequate. Of the 13 content categories (appendix B), categories 1-7 constitute mandatory disclosures in the annual report and voluntary on the web, while the categories 8-13 are voluntary both in annual reports and on websites.

Several measures were implemented to increase the reliability of data, including inter-rater reliability in dataset WEHR. It was decided to make the category variables dichotomous. For each category the value one is assigned if such content is disclosed⁹ and zero if it is not. This leaves less room for subjectivity. For dataset ENV, two coders collected all data together (Stellander and Jørgensen, 2010), while for dataset WEHR, another pair registered data separately (Andersen and Kirkeslett, 2010). The four coders were MSc-students in accounting. Both teams used standardized registration schemes and category descriptions, an example pool of disclosures for each category, discussed and agreed on difficult cases, and training through separate pilot registrations with subsequent comparisons. Coders in dataset ENV trained with an experienced coder. One inter-rater reliability challenge is less relevant in this paper. The analysis only requires that disclosures (in the data sources) within each company are equally registered, not between companies (though that is naturally sought as well). This is secured by having the same person registering all disclosure (in all media) for a company, and by finishing the company before starting on a new one. Another type of reliability is whether all disclosures in all media are discovered and registered. The data sources were printed out and read carefully to identify relevant disclosures. Additionally, electronic searches for key words were performed as a test of completeness.

The time of the website data collection is a potential threat to validity. The annual and separate reports for 2008 were published during the first six months of 2009. Websites for dataset ENV were examined in week two and three in 2010, while the corresponding period for dataset WEHR was February 8th -17th 2010. Ideally, websites should have been examined when annual reports were published in order for the media to reflect the same reality – an assumption that the hypothesis is based on. Companies have had websites for many years, and are mature users. Most of

⁸ An adjustment of mandatory and voluntary categories is done in the current study.

⁹ Irrespective of whether information is one sentence or a whole page, etc.

the content is usually constant over time, not updated from day to day. Few companies experience material CSR crises that would require new disclosure (De Villiers and Van Staden, 2011). Therefore, it is assumed that on average the seven to 13 months delay will not matter that much for which information content categories that are disclosed. Importantly, the consequence if new CSR information content is disclosed on the websites during the delay is that the amount of unique website content increases, and the representativeness of the annual report relative to total disclosure in this study is measured to be lower than it actually is. Hence, the results of the study can be trusted to be at least as strong as they seem.

Data analysis

To get valid results, the disclosure in annual reports and total disclosure must be compared on a micro level – for each information content category for each company – before aggregating the results of all the comparisons to sample level for analysis. In each dataset, for each company and for each data source, the data collection has resulted in the value “1” for each of the 13 types of information content that is disclosed, and “0” for each type of information content that is not. The result is a 13 (content categories) x 63 (companies) matrix for dataset ENV for each data source, and a 13 x 75 matrix for each data source in dataset WEHR. Total disclosure is computed by merging the matrixes for each medium within a dataset in the following way: If, e.g., the top left pane in one or more of the single data source matrixes have the value “1”, then the top left pane in the total disclosure matrix gets the value “1”. If the top left pane in all single data source matrixes equals “0”, then the total disclosure matrix also gets value “0”. After computing the total disclosure matrix, it is possible to compare each pane in the annual report matrix with the corresponding pane in the total disclosure matrix to see if the disclosure in the two matrixes is equal (value “1”) or different (“0”). When aggregating the comparison for each pane, the general result of e.g. the ENV dataset will be a value between zero and $(13 \times 63 =) 819$. If, e.g., 800 of the comparisons resulted in the value “1”, the annual report can be said to cover $(800/819)$ 98% of total disclosure.

RESULTS

The objective of this paper is to examine whether the annual report can be used as the only data source in CSR reporting research. The criterion for this, formulated in the general hypothesis, is that the annual report covers approximately all information content of total disclosure in all used media. In order to examine the general hypothesis, ten more specific hypotheses are derived and empirical results obtained, as summarised in Tables 5 and 6.

Table 5: *The proportion of total ENV disclosure content covered by annual reports*

Dataset ENV	Relatively high environmental risk	Relatively low environmental risk	All companies
Mandatory disclosure	90.2%	97.1%	95.2%
Voluntary disclosure	91.2%	98.5%	96.5%
All disclosure	91.0%	98.2%	96.2%

Table 6: *The proportion of total WEHR disclosure content covered by annual reports*

Dataset WEHR	Industrial companies	Financial companies	All companies
Mandatory disclosure	99.3%	100%	99.6%
Voluntary disclosure	96.6%	98.6%	97.6%
All disclosure	98.0%	99.4%	98.7%

Hypothesis 1 shows that the annual report covers 96% of the environmental content and 99% of WEHR content of total disclosure for these samples respectively. According to hypotheses 2 and 3 these results are more or less independent of whether the content is mandatory or voluntary. The result for WEHR disclosure is also independent of the two industry strata, as seen from hypotheses 4 and 5. There is a modest variation between the industry strata for environmental disclosure, but even for the companies with high environmental risk the annual report covers 91% of the total disclosed content.

DISCUSSION

How should these results be interpreted? The first issue to consider is what percentage of total disclosed content annual reports must cover in order for its representativeness to be perceived as high (and sufficient to be used as the only data source in CSR reporting research)? It depends on the research question, the analysis needed to answer it, and the degree of certainty needed for the wanted conclusion: it is an individual interpretation in each new study. Nevertheless, the “approximately 100%” criterion of this study is so tough that it will suit any study. It is probably much stricter than common requirements such as “an appropriate source of a company’s attitudes towards social reporting” or “totality of reporting intentions” used by e.g. Campbell et al. (2003:566). All the actual percentages in Tables 5 and 6 do approximate the 100% criterion. These results would

undoubtedly have been sufficient for almost all of the studies measuring total disclosure in the literature study described above.

Percentages in Tables 5 and 6 are very high, and there is little variation between or within subsamples. Nevertheless, a few comments should be made about the variation between high and low environmental risk companies, of which annual reports cover 91% and 98% of total disclosure, respectively. This difference might occur because companies from environmentally sensitive industries disclose more types of information content than companies from less sensitive industries, which is true both in this study and generally (Fifka, 2013). The stratum contains the 17 highest risk companies of the population. Then the statistical likelihood of some more unique information content outside the annual report is a bit higher. However, high risk companies are also more inclined to experience environmental events that affect the reality to be reflected in reporting. While there is a risk that the percentages in Tables 5 and 6 are underestimated¹⁰ due to the time lag between the publishing of annual reports and data collection for websites, so they might reflect different realities, the likelihood of understatement is probably largest for the subsamples in hypotheses H4_{ENV} and H4_{WEHR}. Still, none of the above arguments explain why there is no corresponding industry effect concerning WEHR disclosure. There might be a difference in sensitivity of environmental and WEHR issues or disclosures. It is also possible that the current partition between industrial companies and financial institutions does not adequately capture industry sensitivity concerning WEHR issues. However, the 91% result means that the annual report is a valid proxy for total disclosed content of most studies.

The results of Tables 5 and 6 are valid at least when total disclosure is measured using the detailing level chosen in this study (13 content categories) or less. The choice of (these) 13 categories means getting information about the most important aspects of environmental and WEHR issues respectively, but it is still analysis on an overall level. It is a balance between sufficient detail to capture useful information and the ability to apply one set of categories to all companies across industries. Using more categories might also affect reliability and costs of data collection. The current detailing level would have been adequate for the analyses of most of the papers in the above literature study. It should be noticed that studies with a much higher number of content categories (e.g. 100), to enable more detailed analysis, might require more than one data source. This is partly because of the increased statistical probability of unique information in other media, but also because the format of the corporate annual report is to provide information of some importance and

¹⁰ Underestimated results mean that the general hypothesis have even stronger support.

on an aggregate level. The annual report is probably not an equally adequate tool for very detailed disclosure.

The results presented in Table 5 and Table 6 are interpreted as a very high proportion of total disclosed content is present in annual reports, for both CSR themes, for both mandatory and voluntary information content categories, and the different “industries”. This is in line with – and, hence, in support of – the general hypothesis, and has two main implications. Firstly, it means that the annual report is a valid proxy for information content of total disclosure, and can be used as the only data source in environmental and WEHR disclosure research. Secondly, the volume and information content of disclosure is disconnected, and volume of disclosure should not be used as a (single) proxy for information content. The relevance of volume of disclosure depends on the usefulness of the content it is measuring.

The first implication, the current result concerning data source selection, seems to be supported by at least two papers (Niskanen and Nieminen, 2001, Tilt, 2008). However, the results concerning data sources’ representativeness of total disclosed *content* of the 15 papers in Table 4 are so uncertain, due to the methodological issues discussed, that it is not reasonable to make interpretations here. What is clear for the majority of the 15 papers is that the annual report has to be one of the selected data sources (if not the only), due to observed unique content, e.g. monetary, quantitative, economic and negative information.

The second implication concerns which type of data to collect in order to measure total disclosed CSR content. While few papers [except e.g. Fallan and Fallan (2007)] have pinpointed this relationship earlier, the data of a large amount of studies clearly supports the current result that information content and volume of disclosure are disconnected (in different ways) (Hackston and Milne, 1996, Ljungdahl, 1999, Williams, 1999, Williams and Pei, 1999, Adams and Kuasirikun, 2000, Niskanen and Nieminen, 2001, Tilt, 2001, Patten and Crampton, 2003, Tilt, 2008, Fallan and Fallan, 2009, Moroney et al., 2011). Since this result opposes the implicit assumption of the measurement and analyses of e.g. Unerman (2000), Campbell et al. (2003), McMurtrie (2005) and De Villiers and Van Staden (2011), the relevance of the findings of these papers appears to be strongly challenged.

While many papers have looked at reliability issues concerning content analysis (Milne and Adler, 1999), this paper contributes by addressing validity. It is common to claim that the annual report is no longer adequate as the only data source in CSR reporting research, due to the diffusion of the internet and separate reports (Campbell et al., 2006, Clarkson et al., 2011). According to the literature study, such claims are mainly based on beliefs and not empirical findings – at least not recent research. No direct harm is done if several data sources are used in studies. It might add a tiny fraction of extra unique information content. However, indirectly it reduces sample sizes that are

already small due to the extremely resource demanding nature of hand collected content analysis data, and might complicate access to historic and longitudinal data. Decisions of such vital importance for the validity of research should be based on empirical findings. It is also imperative to address material weaknesses identified in existing guidance on data source selection, so researchers can make informed decisions. The results of the study – both for data source selection and which type of data to collect – have important implications for future research designs. Being pragmatic, researchers have to compare benefits and costs of adding more data sources. This findings enlightens researchers' methodological decisions: there is little validity to be gained by adding an extra data source, while large additional resources are needed to hand collected extra content analysis data. Adding data sources might even endanger the possibility of historic and longitudinal studies. Other users of CSR reporting would also benefit from knowing where to look for the information they seek.

The validity of the current study is presumably higher than for existing evidence due to more detailed data for information content, samples and analysis designed to that match the research question, and timely data. Still, limitations in the current study call for more research on issues like industry differences, representativeness of other media, potential systematic differences in representativeness between different types of information content (across media), and the consequences of using a much higher number of content categories or other information characteristics as suggested by Beck et al. (2010). Longitudinal studies could reveal whether the representativeness of annual reports is consistent over time, as the arguments behind the current hypothesis suggest. A general lack of CSR reporting research on unlisted, small and public sector companies should also be addressed regarding the use of media.

CONCLUSIONS

The objective of this study was to examine whether the CSR information content of annual reports is representative of the total disclosure of CSR information content in all used disclosure media. This is important because it will provide guidance to users of information, especially researchers, where to find relevant information content easily and cost effectively. It is also important to update and improve guidance of previous research, because it has weaknesses that make it hard to trust findings of the representativeness of CSR content in different media. The research question is motivated by the literature study of the use of data sources in research. The first contribution of the study is the finding that the annual report has been and still is the most used data source in CSR reporting research, despite the guidance of most previous research that the annual report is not representative of total CSR disclosure. Secondly, it has become more common to use several data sources in this

research, but this has happened without consideration of empirical findings, as these are rarely referred to when explaining the choice of source. The third contribution is that a very high proportion of the totally disclosed CSR information content is present in annual reports, in line with – and hence in support of – the general hypothesis: Approximately all information content concerning environmental and working environment/human resources issues that corporations disclose in different media is also disclosed in annual reports.

The implications of these findings are that the annual report is a good proxy for total disclosure of information content, and can be used as a single data source in environmental and working environment/ human resources reporting research. Robustness tests show that these results are applicable for listed companies, irrespective of industry, and are valid both for mandatory and voluntary disclosure for each of the two CSR information content topics. Moreover, the results are valid for a level of detail consisting of at least 13 information content categories for each of the two CSR themes. Still, the choice of data source(s) should be considered individually, in connection with the research question, analysis and the desired level of certainty with which conclusions should be reported, for each study.

The fact that the annual report is representative of total disclosure of information content from all media also means that the volume of disclosure and the information content is disconnected: Volume of disclosure cannot be used as a proxy of information content and vice versa. It is also likely that the annual report, at the aggregate level of a general purpose report, can be perceived as a summary of CSR reporting. The diffusion of websites and separate reports has not (yet) rendered “the analysis of corporate annual reports, only, as a limited, and less meaningful, research instrument,” as predicted by Campbell et al. (2003:574) and Campbell et al. (2006:102). The same is true for the quote of Unerman (2000), given in the beginning of this paper. In sum, the annual report has been, and should still be, the data source of choice for research on total disclosure of CSR information content.

REFERENCES

- ACCA 2004. Towards Transparency: Progress on global sustainability reporting. London: ACCA.
- ADAMS, C. A. & HARTE, G. 1998. The changing portrayal of the employment of women in British banks` and retail companies` corporate annual reports. *Accounting, Organizations & Society*, Vo. 23, pp. 781-812.
- ADAMS, C. A. & KUASIRIKUN, N. 2000. A comparative analysis of corporate reporting on ethical issues by UK and German chemical and pharmaceutical companies. *The European Accounting Review*, Vol. 9, pp. 53-79.

- AERTS, W., CORMIER, D. & MAGNAN, M. 2008. Corporate environmental disclosure, financial markets and the media: An international perspective. *Ecological Economics*, Vol. 64, pp. 643-659.
- ANDERSEN, L. & KIRKESLETT, A. K. 2010. *Reporting of work environment - A study of differences between annual reports and websites in two industries*. MSc Thesis, Trondheim Business School, Norway (Language: Norwegian).
- BEATTIE, V., DHANANI, A. & JONES, M. J. 2008. Investigating Presentational Change in UK Annual Reports: A Longitudinal Perspective. *Journal of Business Communication*, Vol. 45, pp. 181-222.
- BECK, A. C., CAMPBELL, D. & SHRIVES, P. J. 2010. Content analysis in environmental reporting research: Enrichment and rehearsal of the method in a British-German context. *The British Accounting Review*, Vol. 42, pp. 207-222.
- BRAMMER, S. & PAVELIN, S. 2006. Voluntary Environmental Disclosures by Large UK Companies. *Journal of Business Finance & Accounting*, Vol. 33, pp. 1168-1188.
- BRANCO, M. C. & RODRIGUES, L. L. 2008. Factors Influencing Social Responsibility Disclosure by Portuguese Companies. *Journal of Business Ethics*, Vol. 83, pp. 685-701.
- CAMPBELL, D. 2004. A longitudinal and cross-sectional analysis of environmental disclosure in UK companies - a research note. *The British Accounting Review*, Vol. 36, pp. 107-117.
- CAMPBELL, D., CRAVEN, B. & SHRIVES, P. 2003. Voluntary social reporting in three FTSE sectors: a comment on perception and legitimacy. *Accounting, Auditing & Accountability Journal*, Vol. 16, pp. 558-581.
- CAMPBELL, D., MOORE, G. & SHRIVES, P. 2006. Cross-sectional effects in community disclosure. *Accounting, Auditing & Accountability Journal*, Vol. 19, pp. 96-114.
- CLARKE, J. & GIBSON-SWEET, M. 1999. The use of corporate social disclosures in the management of reputation and legitimacy: a cross sectoral analysis of UK Top 100 Companies. *Business Ethics: A European Review*, Vol. 8, pp. 5-13.
- CLARKSON, P. M., OVERELL, M. B. & CHAPPLE, L. 2011. Environmental Reporting and its Relation to Corporate Environmental Performance. *Abacus*, Vol. 47, pp. 27-60.
- DE VILLIERS, C. & VAN STADEN, C. J. 2011. Where firms choose to disclose voluntary environmental information. *Journal of Accounting and Public Policy*, Vol. 30, pp. 504-525.
- DEEGAN, C., RANKIN, M. & TOBIN, J. 2002. An examination of the corporate social and environmental disclosures of BHP from 1983-1997: a test of legitimacy theory. *Accounting, Auditing & Accountability Journal*, Vol. 15, pp. 312-343.
- DEL BOSCO, B. 2004. Internet use for corporate social disclosures: evidence from Italian listed companies. *Finanza, Marketing e Produzione*, pp. 57-82.
- DEL BOSCO, B. 2011. Communicating corporate social responsibility: a study of the evolution of website disclosures in Italy. Sinergie-Euprera Congress "Corporate Governance and Strategic Communication", Milano, 10-11 November 2011
- DIMAGGIO, P. J. & POWELL, W. W. 1983. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, Vol. 48, pp. 147-160.
- EVANS III, J. H., HANNAN, R. L., KRISHNAN, R. & MOSER, D. V. 2001. Honesty in Managerial Reporting. *The Accounting Review*, Vol. 76, pp. 537-559.
- FALLAN, E. & FALLAN, L. 2007. A longitudinal and cross-sectional analysis of volume and content of corporate environmental disclosure in Norwegian companies: A research note on innovativeness and adoption. TØH-serien 2007:5, Trondheim Business School.
- FALLAN, E. & FALLAN, L. 2009. Voluntarism versus regulation: lessons from public disclosure of environmental performance information in Norwegian companies. *Journal of Accounting & Organizational Change*, Vol. 5, pp. 472-489.
- FIFKA, M. S. 2012. The development and state of research on social and environmental reporting in global comparison. *J Betriebswirtsch*, Vol. 62, pp. 45-84.

- FIFKA, M. S. 2013. Corporate Responsibility Reporting and its Determinants in Comparative Perspective - a Review of the Empirical Literature and a Meta-analysis. *Business Strategy and the Environment*, Vol. 22, pp. 1-35.
- FREEDMAN, M. & JAGGI, B. 1982. Pollution Disclosure, Pollution Performance and Economic Performance. *Omega The International Journal of Management Science*, Vol. 10, pp. 167-176.
- FROST, G. R. 2007. The Introduction of Mandatory Environmental Reporting Guidelines: Australian Evidence. *Abacus*, Vol. 43, pp. 190-216.
- FROST, G. R., JONES, S., LOFTUS, J. & VAN DER LAAN, S. 2005. A Survey of Sustainability Reporting Practices of Australian Reporting Entities. *Australian Accounting Review*, Vol. 15.
- GARCÍA-AYUSO, M. & LARRINAGA, C. 2003. Environmental disclosure in Spain: Corporate characteristics and media exposure. *Spanish Journal of Finance and Accounting*, pp. 184-214.
- GRAY, R., KOUHY, R. & LAVERS, S. 1995. Constructing a research database of social and environmental reporting by UK companies. *Accounting, Auditing & Accountability Journal*, Vol. 8, pp. 78-101.
- HACKSTON, D. & MILNE, M. J. 1996. Some determinants of social and environmental disclosures in New Zealand companies. *Accounting, Auditing & Accountability Journal*, Vol. 9, pp. 77-108.
- HASSELDINE, J., SALAMA, A. I. & TOMS, J. S. 2005. Quantity versus quality: the impact of environmental disclosures on the reputations of UK Plcs. *The British Accounting Review*, Vol. 37, pp. 231-248.
- JOSE, A. & LEE, S.-M. 2007. Environmental Reporting of Global Corporations: A Content Analysis based on Website Disclosures. *Journal of Business Ethics*, Vol. 72, pp. 307-321.
- KPMG 2002. KPMG International Survey of Corporate Responsibility Reporting 2002.
- KPMG 2005. KPMG International Survey of Corporate Responsibility Reporting 2005.
- KPMG 2008. KPMG International Survey of Corporate Responsibility Reporting 2008.
- KPMG 2011. KPMG International Survey of Corporate Responsibility Reporting 2011.
- KRIPPENDORFF, K. 1980. *Content Analysis: An Introduction to Its Methodology*, Sage, London.
- KUASIRIKUN, N. & SHERER, M. 2004. Corporate social accounting disclosure in Thailand. *Accounting, Auditing & Accountability Journal*, Vol. 17, pp. 629-660.
- LARRINAGA, C., CARRASCO, F., CORREA, C., LLENA, F. & MONEVA, J. M. 2002. Accountability and accounting regulation: the case of the Spanish environmental disclosure standard. *The European Accounting Review*, Vol. 11, pp. 723-740.
- LESSEM, R. 1979. Corporate social reporting in action. *Journal of General Management*, Vol. 4, pp. 27-41.
- LJUNGDAHL, F. 1999. The Development of Environmental Reporting in Swedish Listed Companies - Practice, Concepts, Causes. Lund Studies in Economics and Management 50, Lund University, Sweden.
- MCMURTRIE, T. 2005. Factors influencing the publication of social performance information: an Australian case study. *Corporate Social Responsibility & Environmental Management*, Vol. 12, pp. 129-143.
- MILNE, M. J. & ADLER, R. W. 1999. Exploring the reliability of social and environmental disclosures content analysis. *Accounting, Auditing & Accountability Journal*, Vol. 12, pp. 237-256.
- MORONEY, R., WINDSOR, C. & AW, Y. T. 2011. Evidence of assurance enhancing the quality of voluntary environmental disclosures: an empirical analysis. *Accounting & Finance*.
- NEU, D., WARSAME, H. & PEDWELL, K. 1998. Managing public impressions: Environmental disclosures in annual reports. *Accounting, Organizations & Society*, Vol. 23, pp. 265-282.
- NISKALA, M. & PRETES, M. 1995. Environmental reporting in Finland: A note on the use of annual reports. *Accounting, Organizations & Society*, Vol. 20, pp. 457-466.
- NISKANEN, J. & NIEMINEN, T. 2001. The objectivity of corporate environmental reporting: a study of Finnish listed firms' environmental disclosures. *Business Strategy and the Environment*, Vol. 10, pp. 29-37.

- PATTEN, D. M. & CRAMPTON, W. 2003. Legitimacy and the internet: an examination of corporate web page environmental disclosures. In (ed.) *Advances in Environmental Accounting & Management, Vol. 2*, Emerald Group Publishing Limited, pp. 31-57.
- STELLANDER, T. G. & JØRGENSEN, L. K. 2010. *How do Norwegian companies listed on the Norwegian Stock Market report environmental information? A comparative study of annual reports and homepages*. MSc Thesis, Trondheim Business School, Norway (Language: Norwegian)
- TILT, C. A. 1994. The Influence of External Pressure Groups on Corporate Social Disclosure: Some Empirical Evidence. *Accounting, Auditing & Accountability Journal*, Vol. 7, pp. 47-72.
- TILT, C. A. 2001. The content and disclosure of Australian corporate environmental policies. *Accounting, Auditing & Accountability Journal*, Vol. 14, pp. 190-212.
- TILT, C. A. 2008. Environmental disclosure outside the annual report. *Int. J. Management and Decision Making*, Vol. 9, pp. 288-309.
- TINKER, T. & NEIMARK, M. 1987. The role of annual reports in gender and class contradictions at General Motors: 1917-1976. *Accounting, Organizations & Society*, Vol. 12, pp. 71-88.
- TINKER, T., NEIMARK, M. & LEHMAN, C. 1991. Falling down the Hole in the Middle of the Road: Political Quietism in Corporate Social Reporting. *Accounting, Auditing & Accountability Journal*, Vol. 4, pp. 28-54.
- UNCTC 1991. Accounting for environmental protection measures. Report of the Secretary-General, Intergovernmental Working Group of Experts on International Standards of Accounting and Reporting, United Nations Commission on Transnational Corporations Document E/C.10/AC.3/1991/5.
- UNERMAN, J. 2000. Methodological issues: reflections on quantification in corporate social reporting content analysis. *Accounting, Auditing & Accountability Journal*, Vol. 13, pp. 667-680.
- WILLIAMS, S. M. 1999. Voluntary Environmental and Social Accounting Disclosure Practices in the Asia-Pacific Region: An International Empirical Test of Political Economy Theory. *The International Journal of Accounting*, Vol. 34, pp. 209-238.
- WILLIAMS, S. M. & PEI, C.-A. H. W. 1999. Corporate Social Disclosures by Listed Companies on Their Web Sites: An International Comparison. *The International Journal of Accounting*, Vol. 34, pp. 389-419.
- WISEMAN, J. 1982. An evaluation of environmental disclosures made in corporate annual reports. *Accounting, Organizations & Society*, Vol. 7, pp. 53-63.
- ZEFF, S. A. The Objectives of Financial Reporting: A Historical Analysis. ICAEW's Information for Better Markets Conference, November 28, 2012 draft 2012 London.
- ZÉGHAL, D. & AHMED, S. A. 1990. Comparison of Social Responsibility Information Disclosure Media Used by Canadian Firms. *Accounting, Auditing & Accountability Journal*, Vol. 3, pp. 38-53.

APPENDIX A

ENV information content categories

No:	Category:	Regulation:
1	Environmental policy	Voluntary
2	Environmental objectives	Voluntary
3	Environmental impact – process	Mandatory
4	Environmental impact – products	Mandatory
5	Environmental organization	Voluntary
6	Environmental auditing	Voluntary
7	Environmental authorities	Voluntary
8	Environmental events	Voluntary
9	Environmental investments	Voluntary
10	Environmental costs/-revenues	Voluntary
11	Environmental liabilities	Voluntary
12	Definition of environmental concepts/accounting principles	Voluntary
13	No environmental impact	Mandatory

Source: Fallan and Fallan (2009)

APPENDIX B

WEHR information content categories

No:	Category:	Regulation:
1	Sick leave	Mandatory
2	Gender equality	Mandatory
3	Accidents and injuries	Mandatory
4	Implemented measures	Mandatory
5	Planned measures	Mandatory
6	Employee remuneration	Mandatory
7	Number of employees / full-time equivalents	Mandatory
8	Policy for the working environment	Voluntary
9	Objectives for the working environment	Voluntary
10	Organization	Voluntary
11	Survey of the working environment	Voluntary
12	Information and communication	Voluntary
13	Employee remuneration	Voluntary

APPENDIX C

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Zéghal and Ahmed (1990)	CSR disclosure	The largest companies of two industries	Six banks and nine oil companies	Canada		15	30
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1981-1982	1) Annual reports, 2) brochures, and 3) advertisements	Number of words	7 general content categories (environment, energy, fair business practices, human resources, community involvement, products, and other CSR disclosures), each comprising of 1 - 8 subcategories (26 in total).	Monetary, quantitative and narrative information	Data is treated as cross-sectional. Results are presented for each data source pertaining to each company. Only the seven general content categories are used in the analysis.	
<p>The study shows that annual reports and brochures were widely used media for disclosure in the two industries. Although advertisements were used, in general, their usage was rather limited. This is shown both in the volume of disclosure, the number of companies using each media type, and how many content categories the disclosure of each company covers in each medium. According to Zéghal and Ahmed (1990:48) this <i>"indicates that taking into consideration only the social information disclosure made through annual reports gives a somewhat distorted view of a firm's activities in this respect."</i> It is also claimed that companies disclose different information content in the three media types. Zéghal and Ahmed (1990:50) indicate that this <i>"may be interpreted as evidence of the complementary nature of the three media under study as sources of information on corporate social responsibility."</i> However, the content type that is disclosed, and the medium used, appears partly to vary with industry. Disclosures are almost entirely narrative in brochures and advertisements, irrespective of industry. In annual reports, a much higher proportion of the information was quantitative or monetary in both industries. The findings of Zéghal and Ahmed (1990) indicate that the selection of data source matters for CSR reporting research. Both volume and information content as measurement methods suggest that the annual report should not be used as the only medium in order to obtain a picture of total disclosure. Other information characteristics (monetary, quantitative, or narrative information) support the use of annual reports as at least one of the data sources. Unfortunately, their analysis does not consider all relevant aspects, even within the level of detail presented. Cross-tabulating the presented data, shows that all the six banks (100%) disclosed environmental issues and human resources issues respectively in annual reports at least, when the topic was disclosed in other media. Similarly, all nine petroleum companies (100%) provided human resources disclosure in the same manner, while eight out of nine companies (89%) did the same for environmental disclosure. Sadly, the authors only presented data at the general information content category level, and not for subcategories. However, this new data still questions the conclusion to some extent.</p>							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Clarke and Gibson-Sweet (1999)	Voluntary CSR disclosure	The Times 1997 Top 100 companies		UK		92	92
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1995/6	1) Annual reports, and 2) separate CSR reports	NA	Two general information content categories (community involvement and environmental impact).	NA	Results are presented separately for each data source, each measurement method, and both general content categories at the industry level.	
While 73% of the companies disclosed community involvement information in the annual report, only 11% did the same in separate reports. The corresponding adoption rates for environmental disclosures were 70% in the annual report and 46% in separate reports.							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Williams and Pei (1999)	CSR disclosure	Listed companies		Australia, Hong Kong, Malaysia and Singapore	Companies operating their own website	172	172
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1996 = data source 1 and 2; 1997 = source 3	1) Annual reports, 2) separate CSR reports, and 3) websites	Number of sentences	5 general content categories (environment, energy, human resources and management, products and customers, and community), each divided into 7 to 10 subcategories (42 in total).	Monetary, quantitative and narrative information	Results are presented separately for data source 1, while source 2 and 3 are grouped together, and for each type of content analysis data, at the country (not company) level.	
The volume of disclosure is at least as high on websites as in annual reports, and significantly higher for some general content categories in two of the countries. This is largely attributable to more narrative disclosure on the Internet. Disclosures are dominantly narrative, irrespective of content, media, or country. Still, in Australia (and perhaps Singapore) there is relatively more quantitative and monetary disclosure in annual reports than on the Internet for some general content categories. The adoption rate for each of the general content categories does not appear to differ between the two media within each country, except that "products and customers" is more frequently adopted on the Internet. Although average rates do not say much about company level results, Williams and Pei (1999:398) claim that " <i>few companies made corporate social disclosures solely on their Web Sites.</i> " According to Williams and Pei (1999:409), there is not much difference in information content between the media at the subcategory level either, within each country: " <i>The order of ranking of topics was virtually identical regardless of the form of media (annual report or Web Site) considered.</i> " The findings in this paper do not rule out the possibility that it is sufficient to use one medium (annual reports or websites) as data source for studies of information content. A more thorough analysis is needed.							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Unerman (2000)	CSR disclosure	One of the largest companies	One of the world's major energy and chemicals groups	UK		1	100
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1898-1997	1) Annual reports; 2) other docs. in the reporting cycle; 3) regular docs. not apparently for third parties; 4) ad hoc docs.	Number of pages	One general content category (CSR disclosure)		Results are presented separately for each data source for each year, for the single measure (volume of disclosure) and the single company.	
<p><i>"The ... study does ... seek to establish whether CSR is present in a range of corporate reports other than annual reports. If it is not, then exclusive focus on annual reports by other studies might be justified. But if CSR is present in reasonable quantities in corporate reports other than annual reports, this might raise questions over the validity of future studies focusing exclusively on annual reports"</i> (Unerman, 2000:671). Unerman (2000) analyses the total CSR reporting of one large company for a 100-year period. The number of pages containing CSR information is counted (without registering or analysing what type of information that is presented). The study yields two main findings. The first is that a large proportion of volume of CSR information is disclosed outside the annual report. It implies that, by examining only the annual report, there is a risk of underestimating the total volume of CSR disclosure. All studies listed in Table 4a-4c confirm this finding. The second main finding is that the annual report's proportion of total CSR disclosure fluctuates over time. The consequence is that studies measuring the relative change in total volume of disclosure over time, using only annual reports, might yield false results. Hence, Unerman (2000) suggests using several data sources.</p>							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Niskanen and Nieminen (2001)	Environmental disclosure	Companies listed on Helsinki Stock Exchange		Finland	Companies where environmental events have occurred (according to a newspaper), have a separate CSR report, and (relevant here) where negative events are not disclosed in the annual report.	2 (in the analysis relevant here)	2 (in the analysis relevant here)
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1985-1996	1) Annual reports, and 2) separate environmental reports	NA	NA (in the analysis relevant here)	Negative information (in the analysis relevant here)	In the analysis relevant here, results are presented, separately for each data source, as cross-sectional data, at an individual company year level	
<p>Separate environmental reports were not used to disclose information about actual negative environmental events when it was not disclosed in annual reports. Even though no conclusion should be drawn on such scarce data alone, there was no indication of unique information disclosure in separate reports. Niskanen and Nieminen (2001:35) state that <i>"the objectivity of the reporting was not improved compared to the ordinary annual reports."</i></p>							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Campbell et al. (2003)	CSR disclosure	Companies among the largest listed (FTSE 100)	Industries with differing perceived "sinfulness" (tobacco, brewing and retail)	UK		5	5 (websites); 115 (annual reports)
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1975-1997 = Data source 1; 14 November 2000 = source 2	1) Annual reports, and 2) websites	Annual report study: number of words. Website study: ad hoc (pages, paragraphs, etc.), if registered	Annual report study: 3 general content categories (employee welfare, environment, and community). Website study: exploratory/ ad hoc, pre-defined categories not used.	Annual report study: NA. Website study: negative information (ad hoc).	Results presented separately for each company, each year, and for each data source. Volume pertaining to annual report content categories presented only as one total. Results for the two data sources not comparable.	
<p>Campbell et al. (2003) have richer website data than this classification implies; however, the presented annual report analysis limits the readers' ability to compare the data sources. The volume of CSR disclosure indicates varying use of annual reports and websites. The company with consistently lowest annual report disclosure volume certainly had the greatest number of web-based disclosures, while the only company that did not disclose any information on the Internet had as large volume as the last three companies in annual reports. For the rest of the sample, the results were more ambiguous. Campbell et al. (2003:574) conclude: "<i>The rise of the Internet as a communication vehicle renders the analysis of corporate annual reports alone as a limited, and less meaningful, research instrument. This is not to say, of course, that previous studies that adopted an "annual report only" approach to data capture prior to the rise of the Internet are without value.</i>" Another interesting issue is whether or not the reference to the "pragmatic view that [a medium] can be accepted as an appropriate source of a company's attitudes towards social reporting" and "totality of reporting intentions" [italics added] used both by Campbell et al. (2003:566) and several other papers reflects a similar and an equally strong criterion for acceptance of a data source as representative of total disclosure, as is used in this paper (even if information content is considered, instead of Campbell et al.'s (2003) disclosure volume).</p>							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
García-Ayuso and Larrinaga (2003)	Environmental disclosure	Companies listed on the Madrid Stock Exchange	Industrial companies	Spain		Not disclosed for the analysis relevant here	116 (in the analysis relevant here)
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1991-1995	1) Annual reports for the authorities, and 2) annual brochures for general meetings/ shareholders	Number of pages (lines), both in absolute terms and relative to the total number of lines in the report	One general content category is used (environmental disclosure)	Positive and negative information (however negative information was hardly found)	In the analysis relevant here, only cross-sectional results are presented, at a sample level for each data source, and only for the measure "volume of disclosure".	
<p>The results oppose the second finding of Unerman (2000), who noted that the annual report's proportion of total CSR disclosure fluctuates over time, although García-Ayuso and Larrinaga (2003) examine only two rather similar data sources.</p>							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Patten and Crampton (2003)	Environmental disclosure	Companies among the Fortune 500 largest companies	Companies listed as either chemical industry or electrical equipment industry	USA	Companies having an accessible corporate web page	62	62
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1997/8 = data sources 1 and 2; 1998 = source 3	1) Annual reports, 2) separate CSR reports, and 3) websites	Number of sentences	21 content subcategories, 17 of which are positive/neutral disclosures (economic, pollution abatement and other content), and 4 categories under the heading "negative disclosures".	Positive/neutral and negative information	The types of content analysis data are presented separately for data source 1 and as one for sources 2 and 3, on an industry (and partly on a total sample) level.	
	Measured by volume of disclosure, most of the reporting is found on websites (including separate reports etc.) and is almost entirely positive or neutral. Annual report disclosures have a more equal distribution between positive/neutral and negative number of sentences. Shifting the analysis to adoption rates of content categories and other attributes of information shows consistently that more companies disclose the subcategories related to negative information and economic information (defined as positive/neutral) in the annual report than on websites. For other positive/neutral subcategories, the picture is mixed. There are some industry variations in these results. Patten and Crampton (2003:51) state that "... the sample companies, on average, devoted significantly more space to environmental issues on their web pages than in their annual reports However, the concurrent finding that the increased space does not correspond to similar increases in the content of disclosure is interesting." The authors call for a more detailed analysis, including the effect of regulations, noting that the type of information content to be studied might matter for the choice of data source. Anyway, the data clearly shows that the annual report has to be, at least, one of the data sources that are used, since it contains much unique information.						

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
McMurtrie (2005)	CSR disclosure	Large companies	Companies that operate in sensitive industries—one industrial and one mining company	Australia	Companies that publish a large amount of CSR disclosures	2	2
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	Ca. 2001/2	1) Annual reports, 2) separate CSR reports, 3) websites, 4) press releases, and 5) other documents	Number of pages	6 general content categories (environment, energy, community, product, human resources management, and general).	NA	The number of pages of each general content category are presented separately for each data source for both companies	
	McMurtrie (2005:142) concludes that "The examination of CSD [corporate social disclosure] other than that contained in the annual report has shown that, in these [two] cases, the annual report alone would not give a true picture of the CSD profile of the organizations." This is based on the volume of disclosure (websites and separate CSR reports constitute most of the total CSR disclosure in the two companies), as well as the diversity of, and changes in, main disclosed general information content categories in different media and between the companies. However, some patterns appear that are not commented on. For example, ranking the six general content categories in each company according to volume of disclosure results in very similar order, independent of media, in both companies (except for the topic "community" in the mining company). In addition to this within-company result, the ranking is also quite similar between the two companies within annual reports and websites (except "community" for the websites). Hence, the conclusion is not that clear-cut. Research based on more than two companies is needed.						

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Frost (2007)	Environmental disclosure	Companies listed on the Australian Stock Exchange	Companies in environmentally sensitive industries that are most likely to be required to report on environmental performance	Australia		71	284
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1996-1999 (or 1997-2000)	1) Annual reports, and 2) separate CSR reports. Only existence, not content, of separate reports is registered.	Number of words	3 content subcategories (regulations, compliance, and remedial action).	General, specific and negative information	Volume of disclosure presented separately for each year and (non)existence of separate reports. Each content category presented at the total sample level.	
	Frost (2007) claims that the use of separate reports influences the level of environmental disclosure in annual reports. In reality, companies issuing a separate report disclose higher volume of information in annual reports, both before and after the introduction of disclosure regulation; however, the increase is higher for companies not having a separate report. Only annual reports are examined, but the existence of separate reports is controlled for. While the analysis is based on volume of disclosure, it assesses total, mandatory, voluntary and negative disclosure. The picture is the same for all these categories. Annual reports are required as, at least, one of the data sources, since many companies do not issue separate reports. The relevance of this paper for choice of data source is limited, as it does not have data on more than one medium.						

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Aerts et al. (2008)	Environmental disclosure	Listed companies	Non-financial companies from eight industries (facing different environmental risks)	USA, Canada, France, Germany, Belgium and the Netherlands		892	892
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	2002	1) Annual reports, 2) separate CSR reports, and 3) websites	NA	6 content subcategories, each further divided into 3 to 10 sub-subcategories (38 in total).	Monetary, quantitative, specific and general information	Data source 3 presented separately, and sources 1 and 2 as one. Total disclosure scores at the general content and subcategory level analysed by industry, country and continent.	
	"Disclosure scores" for environmental reporting are much higher for "paper" disclosures (annual reports and separate environmental reports combined) than for websites. This applies to the six information content subcategories independent of country and industry, albeit with possible exceptions for environmental management and sustainability. The disclosure score is computed by the use of both information content subcategories and other attributes of information. Comparing the two possible exceptions with the other content categories, Aerts et al. (2008) suggest that regulations might affect disclosure, leaving a more extensive discussion for further research. According to the authors, narrative disclosure completely dominates web-based disclosures, irrespective of country or industry grouping. Due to the way the results are presented, the reader cannot separate information content from other information characteristics. Unfortunately, the same is the case for the two media included in paper-based disclosure. The result shows that annual reports have to be included as data source. It is not clear whether it is sufficient as the only data source.						

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Branco and Rodrigues (2008)	CSR disclosure	Companies listed on the Portuguese Stock Exchange		Portugal	Companies with an accessible website	49	49
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	2003 = data source 1 and 2; 2004 = source 3	1) Annual reports, 2) websites, and 3) press releases	NA	4 general content categories (environment, human resources, products and customers, and community involvement), each divided into 5 to 11 subcategories (30 in total).	NA	Content categories, at both general and subcategory level, are presented separately for data source 1 and are combined for sources 2 and 3, at both industry and total sample level.	
	There is a significantly greater proportion of companies disclosing human resources information content in annual reports than on websites, while there is no such difference for the three other CSR themes studied. The results for each subcategory under the four overall general content categories show the same picture even more clearly. There might be industry variations to these results. It might seem like several data sources are needed to capture total disclosure, in general, but for certain types of information one might suffice.						

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Tilt (2008)	Environmental disclosure	Companies among the top 500 largest, and listed on the Australian Stock Exchange. 1999-sample supplemented by top 200 companies	Various industries, but 50% mining/ chemical industry companies in 1994 and the proportion rose in 1999	Australia	Companies with formal documented corporate environmental policy and disclosing environmental information both in annual reports <u>and</u> other media. 1999-sample supplemented by companies that had separate reports.	36	42-43
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	1994 = data sources 1, 3, 4, and 6; 1994 and 1999 = sources 2 and 5	1) Annual reports, 2) separate CSR reports, 3) brochures, 4) advertisements, 5) press releases, 6) other sources	Number of sentences for some media and number of pages for others; both in absolute terms and relative to all information disclosed (incl. non-environmental)	10 information content subcategories within the general theme environmental disclosure (although the data is presented only for separate reports).	1) Monetary, quantitative and narrative information; 2) general, specific and non-environmental information	Results presented for each data source and specified year separately, at the sample (rather than individual company) level.	
<p>Tilt (2008) explores “whether non-annual report disclosures provide new information or reinforce existing disclosures in the annual report.” This question was previously asked by Zéghal and Ahmed (1990), but is here limited to environmental disclosures. According to Tilt (2008:299/303) “it appears that no new information from that provided in the annual report was disclosed, and that even less detail was provided in every medium other than the separate environmental report” in the 1994 comparison. The data leading to this conclusion is not adequately reported, but the description and work by Tilt (2001) indicate that it is presumably based on information content subcategories (and perhaps other information attributes). The use of separate reports increased from 1994 to 1999, while the use of press items ceased. Tilt (2008:302) states: “The examination of press releases confirmed the suspicion that apart from the use of separate environmental reports and the annual report, little use is made of alternative [non-electronic] disclosure media.” Regarding other information attributes, the 1994 disclosures were almost entirely narrative, while 1999 separate reports contained more quantitative information. The mean volume of disclosure varied significantly between different media, and annual reports constitute just a small part of the total disclosures. While Tilt (2008) does not find any industry differences, environmental sensitive companies were more likely to produce separate reports. In sum, it seemed sufficient to use annual reports as a single data source for studies of environmental content at the subcategory level in the mid 1990s. It is called for similar studies that include websites.</p>							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
De Villiers and Van Staden (2011)	Environmental disclosure	Standard and Poor’s 500 and the largest 3000 US publicly traded firms		US	Companies with bad and not bad environmental performance, that have and have not had an environmental crises.	120	120
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	2005	1) Annual reports, and 2) websites	Number of sentences	One general category: environmental disclosure	NA	Results are presented for each data source at the sample (not individual company) level.	
<p>De Villiers and Van Staden (2011) conclude that voluntary environmental disclosure in annual reports and on websites is complementary. Disclosure in annual reports is relatively more used by companies having long-term bad environmental reputation, while websites are more used by companies experiencing a short-term environmental crisis. It is suggested that both data sources are required to obtain a complete picture of disclosure. However, the analysis relies only on volume of disclosure, not information content. While many control variables are used, the effect is not controlled for industry (within their four strata). To sum up, the advice have to be that both data sources are needed in studies.</p>							

Article	Topic	Size and stock exchange listing	Industry	Country	Other sampling criteria	No. of companies	No. of company years
Moroney et al. (2011)	Environmental disclosure	Companies among the top 500 largest and listed on the Australian Securities Exchange		Australia	Companies with verified disclosure at least once in the period 2003-07, and a matched sample (by time, industry and size) of companies not having verified disclosure.	18-38	148
	Years	Data sources	Type of data: volume of disclosure	Type of data: information content categories	Type of data: other attributes of information	Level of analysis (results and presented data)	
	2003-2007 = data sources 1 and 2; 2008 = source 3	1) Annual reports, 2) separate CSR reports, and 3) websites	NA	7 content subcategories, each further divided into 3 to 10 sub-subcategories (44 in total).	Hard and soft disclosure items	Data sources presented together as one, except for a brief comment, where source 3 is separated, and 1 and 2 combined.	
<p>Moroney et al. (2011) do not present their detailed voluntary environmental disclosure (VED) data separately for each medium, but provide some insight as a digression at the end of the paper: <i>“Environmental disclosures appearing on corporate websites were only included in the scoring of VED for 2007. Most of the information published on corporate websites had been reported in the 2007 annual and/or stand-alone reports. The average VED for 2007 was compared to previous years, and no significant difference was found ($p > 0.10$), indicating that the inclusion of web disclosures in the main analysis did not affect the reported results”</i> (Moroney et al., 2011:21). The authors note lack of unique information content on the websites compared to annual reports and separate reports combined. Unfortunately, the digression does not separate between the two latter media types.</p>							

