Environmental Management (2009) 44:298–311
DOI 10.1007/s00267-009-9305-9

Development of an Evaluation Methodology for Triple Bottom Line Reports Using International Standards on Reporting

Antonis Skouloudis · Konstantinos Evangelinos · Fotis Kourmousis

Received: 8 April 2008 / Accepted: 5 April 2009 / Published online: 4 June 2009

Abstract The purpose of this article is twofold. First, evaluation scoring systems for triple bottom line (TBL) reports to date are examined and potential methodological weaknesses and problems are highlighted. In this context, a new assessment methodology is presented based explicitly on the most widely acknowledged standard on non-financial reporting worldwide, the Global Reporting Initiative (GRI) guidelines. The set of GRI topics and performance indicators was converted into scoring criteria while the generic scoring devise was set from 0 to 4 points. Secondly, the proposed benchmark tool was applied to the TBL reports published by Greek companies. Results reveal major gaps in reporting practices, stressing the need for the further development of internal systems and processes in order to collect essential non-financial performance data. A critical overview of the structure and rationale of the evaluation tool in conjunction with the Greek case study is discussed while recommendations for future research on the field of this relatively new form of reporting are suggested.

Keywords Global Reporting Initiative · Triple bottom line/non-financial report · Greece · Scoring system · Benchmark

A. Skouloudis (✉) · K. Evangelinos
Department of Environmental Studies, Center for Environmental Policy and Strategic Environmental Management, University of the Aegean, University Hill, Mytilene 81100, Greece
e-mail: skouloudis@env.aegean.gr

F. Kourmousis
Union of Environmental Scientists of Greece, Boulgaroktonou 30, Athens 11472, Greece

Introduction

A new era in corporate reporting is imminent. Many companies around the world have realized that conventional financial reports and related accounting methods alone are not sufficient to provide information on intangible assets and non-financial issues of the organization. Publicly available information on topics such as environmental and social performance, management quality or internal governance transparency is clearly now vital for investors and shareholders in order to make fully accurate decisions (e.g., Holland and Foo 2003; Repetto 2005; Rikhardsson and Holm 2008) as well as for the rest of the stakeholders—customers, suppliers, employees, communities and other social groups—who also expect a higher standard of accountability and demand a more comprehensive depiction of corporate impacts, risks and performance (GRI 2002; Kaptein and Van Tulder 2003; Logsdon and Lewellyn 2000; Rasche and Esser 2006; Woodward and others 1996; WBCSD 2002). In addition, this emerging form of reporting can facilitate corporate planning and decision-making (Adams and McNicholas 2007; Albelda-Perez and others 2007).

The rise of non-financial reporting has also spawned a number of evaluation methods and relevant studies in order to examine the quality and inclusiveness of such disclosure practices. Such studies have added vital information on the evolution and emerging trends in reporting procedures, but have concentrated either on multinational enterprises-organizations or geographical regions with previous experience of non-financial information disclosure. In contrast, there has been almost no debate on organizations and countries with inadequate expertise to draft state-of-the-art reports in terms of content and comprehensiveness. Taking this into account, this study attempts to shed light on methodological weaknesses of previously developed non-financial reporting scoring systems and proposes a new approach, exclusively based on the most acknowledged framework on TBL reporting, the Global Reporting Initiative (GRI) guidelines. The proposed benchmark tool allows the examination of how well an organization's reporting practices stand against the current "de facto" standard on non-financial reporting and, hopefully, it could initiate consensus on the field of TBL reporting assessment.

The rest of the article is structured as follows: First, the concept of TBL reporting is illustrated and, secondly, we draw on the literature of previously developed TBL reporting methodologies and indentify relevant possible drawbacks of methodological nature within such evaluation systems. Third, an overview of GRI guidelines is provided and thereinafter an analysis of the proposed methodological approach is presented. Following on from this, results of the application of the new benchmark tool on Greek reporters are presented. The article concludes with a discussion on the development of our proposed assessment scoring system in conjunction with the Greek results, as well as directions for future research.

Triple Bottom Line Reporting: Where Does It Stand?

Accountability, along with transparency, has been two concepts that have gained prominent importance in corporate business conduct over the last 15 years (Zadek 2001; Accountability 2004; Waddock 2004). This was triggered by negative incidents that focused public attention on specific corporations and sectors (Kolk 2005a) and initiated the publication of the first non-financial reports, which concentrated on health, safety and environmental (HSE) issues, combining employee related and environmental information. Since then, non-financial reporting has evolved to "triple bottom line," a term coined by Elkington (1997), which aims to present the economic, environmental and social performance of an organization in a balanced and comprehensive way. Accountability failures in cases of large corporations have also stimulated governments to enact mandatory requirements to broaden corporate reporting by including non-financial aspects of business activity (Kolk 2003, 2005b; KPMG 2005). Countries in Europe, Australia, Canada, Japan and the United States have undertaken measures to that end. However, only in France are publicly listed companies required to produce stand-alone reports on the economic, social and environmental implications of business activity (Pallenberg and others 2006), while governments of other countries still remain quite reluctant to intervene (Antal and Sobczak 2007). Even though specific elements of non-financial reporting, mostly in the domains of corporate governance and performance, are now required by law, much of it still represents voluntary practice.

Mathews (1995) denotes three major sets of arguments which illustrate why organizations engage in social and environmental accounting and thus explain why they also decide to undertake relevant reporting activities. The first group refers to market-related arguments built on the notion that non-financial disclosures can contribute to increased market performance, since they can potentially "influence share prices and returns, and therefore be in line with the role of modern management" (Mathews 1995, p. 667). The second set of arguments concerns the demonstration of accountability through the establishment and maintenance of organizational legitimacy with specific social groups or the general public. The third and most persuasive group of arguments, according to Mathews, lies in the idea of a social contract between business and society and non-financial reporting, under the principles of morality and (social) justice, is a practical outcome of ethical concerns.

Morhardt and others (2002) stress the following principal incentives to explain why companies use non-financial reporting as a vehicle to discharge their environmental and social accountability:

- Compliance with regulatory requirements and proactive cost-reduction of future, stricter regulations.
- Compliance with industry environmental codes, especially in the case of sanctions for non-compliance.
- Reduction of operating costs.
- Promotion of stakeholder relations.
- The perceived environmental visibility of the firm.
- The notion that reporting on such issues can yield competitive advantages.
- The sense that with active environmental management absent, the organizational legitimacy of the company is questionable.
- The sense of the social responsibilities of doing business and desire to adhere to societal norms.

Over the last few years, the number of organizations publishing information on their economic, environmental and social performance has substantially increased (CSR Network 2003; Kolk 2003). According to a tri-annual international survey on corporate responsibility reporting practices, there has been a steady growth in the number of organizations reporting on environmental and social issues. In 2005, the survey revealed that 52% of the top 250 companies of the Fortune 500 (Global 250) issued separate non-financial reports, compared with 45% in 2002, while the top 100 companies of 16 countries included in the survey and producing such reports increased from 13% to 33% (KPMG 2005). Similarly, the Corporate Register database (www.corporateregister.com) indicates the growth in the amount of published stand-alone non-financial reports;

from less than 30 environmental reports in 1992, the number of TBL reports registered to the database has risen to more than 1800 in 2008. However, the rapid diffusion of this new type of corporate reporting is not uniform throughout the world (ACCA 2004; Context 2006; KPMG 2005). A study prepared by ACCA notes: "…in some countries sustainability reporting has yet to emerge, although in others the pioneering reporters first started 15 years ago. The escalating number of standards, codes, guidelines and regulations developing in different parts of the world has naturally contributed to reporting developments on a regional or country level" (ACCA 2004, p. 4). Nevertheless, GRI has rapidly become the leader among voluntary worldwide TBL reporting systems with more than 1000 organizations, including many of the world's leading brands, declaring their voluntary adoption of the guidelines.

While report content has expanded and moved towards third-party assurance of information provided, increasing the reliability of disclosures, the overall quality of TBL reports is still rather low in terms of comprehensiveness and materiality (Palenberg and others 2006). As reporting surveys reveal, many companies do not address sector-specific key issues and performance indicators. This is further verified by scoring assessments of non-financial reports which rate the information disclosed and allow, to the extent feasible, comparative analyses. The different methodological approaches developed to date to score TBL reports will be discussed in the following section.

Scoring Systems for Evaluating Non-Financial Reports

Previous Studies

Over the last decade, several approaches for judging the information provided within corporate non-financial reports, under scoring variants, have been proposed either by consultancy firms or academic researchers. Such rating systems concentrate on the breadth and depth of topics discussed and allow a convenient comparison between different reporting practices, as well as a ranking of the assessed reports with the aim of distinguishing between better and poorer reporters. Benchmarking TBL reports according to a scoring system can yield potential benefits. It informs stakeholders in a simple but systematic manner about the efforts that have been made by the reporting organizations in order to provide adequate and meaningful information on their operation and impacts. It also assists the companies themselves, as they receive evaluation on the reporting procedures they follow and consequently how well they promote effective stakeholder communication. Moreover, it helps them to identify potential reporting strengths and weaknesses on this relatively new type of reports, but also to compare their reporting performance against their peers. Still, it is an evolving process adapting to the reporting trends that constantly emerge and aiming to spotlight best practices in order to indirectly further raise the bar for better TBL reporting in the future.

Davis-Walling and Batterman (1997) evaluated the environmental reports published in 1994 by the Fortune 50 companies, the 50 largest US firms based on the total 1993 US sales volume. After collecting all the available reports, commonly addressed topics in at least four reports were identified. These topics, 29 in total, were grouped into six categories of scoring criteria: corporate policies and investments, community involvement, employee involvement, regulatory concerns, pollution prevention, and miscellaneous. Ten topics were assigned 0–2 points, while the rest received 0 or 1 points. General statements were awarded with 1 point whereas more specific information received 2 points. Individual topic scores were added together to provide category scores and a total score was reached for each firm. This enabled the comparison of the reports under certain clusters of criteria as well as an analysis of the different sectors that comprised the sample.

SustainAbility in association with UNEP has produced a benchmark tool for the evaluation of corporate reports that has been revised several times since its conception, in order to take into account new developments in reporting. Since 1994, SustainAbility/UNEP has performed several benchmark surveys based on this scoring system evaluating state-of-the-art reports. In its latest form (2006) the assessment methodology consists of 29 criteria, which focus on generic business processes and if/how these take into account sustainability issues. The criteria are grouped into four major sections: governance and strategy, management, presentation of performance, accessibility and assurance. Each of these criteria is scored between 0–4 points: when there is no mention of a certain topic no points are given; 1 point is given when coverage is sketchy; 2 points when it is serious and systematic; 3 points when the presented information is extensive and with no major gaps. The maximum points are assigned when coverage is integrated and evident from the discussion on the specific topic that is linked to general business decision-making and internal core processes.

Deloitte Touche Tohmatsu (2002) developed a sustainability reporting scorecard as a means of self-assessment by organizations in order to evaluate their reports. This tool predominantly judges the ability of a report to communicate effectively and create credibility, rather than the quantity of specific data. Proposed as an assessment method for report writers as well as report users and analysts, the scorecard has been used by Deloitte to evaluate the reporting practices within three major industries: the automotive, the pharmaceutical and the mining and minerals sectors. The scorecard comprises 30 criteria grouped into six clusters (effective communication, relevance of information, management quality and commitments, the organization's sustainable development agenda, performance and credibility), which are assigned a rating between 0 and 4 points, from "no mention" or "very insufficient" to "pace-setting creative new approach" or "best practice". Deloitte gives specific guidance to scorecard users on how to assign points and suggests report assessment to be conducted by more than one person in order to alleviate some of the associated subjectivity.

Clausen and others (2005) implemented a ranking of sustainability reports as a follow up of former IÖW (Institute for Ecological Economy Research) rankings of environmental reports carried out during 1994–2000 in Germany. Their methodological approach was based on 13 main criteria that covered environmental, social and integrated requirements in a sustainability report. Each one of these criteria had a set of sub-criteria, 48 in total, which were evaluated according to the following generic scoring devise: 0, no fulfillment of requirements; 1, poor fulfillment; 3, good fulfillment; 5, exceptional fulfillment. Points awarded for all sub-criteria were added together and divided by the number of sub-criteria. The resulting value was then multiplied by a weighting factor and provided the main criterion's score. In the report assessment, further information was taken into account only when a clear reference or link was given, but this was not the case when general or vague statements were made.

The Roberts Environmental Center of Claremont McKenna College has been conducting research on the field of corporate sustainability reporting for several years now. Report assessment is implemented on the basis of a numerical scoring system: the Pacific Sustainability Index (PSI). The proposed evaluation method assesses the environmental and social disclosures of a non-financial report and grades them according to two systematic questionnaires—a "base" questionnaire that refers to all organizations regardless of the type of business activity, and a sector-specific questionnaire for organizations within the same sector. The selection of criteria–questions is based on the most frequently mentioned issues identified in a large sample of corporate sustainability reports, which were analyzed from 2002 through mid-2005 at the Roberts Environmental Center. A major difference from all the other models mentioned for report evaluation is that PSI also grades the performance of the assessed company; it distinguishes scoring of reporting practices as well as corporate (environmental and social) vision and strategy from the scoring of performance and evaluates them independently. The PSI method has been applied to more than thirty sector surveys while others are still in progress as the Roberts Center evaluates the reporting practices among the largest U.S. and global organizations.

GRI Reporting Guidelines as a Scoring System

Morhardt and others (2002) constructed a scoring system based on the first version of the GRI guidelines (G1) and assessed the environmental reports of 40 of the largest global industrial companies published in 1999. GRI topics were converted to scoring criteria which were assigned a score between 0 and 3 points. When a topic was not mentioned in an assessed report no points were given; the report was awarded with 1 point when the topic was briefly mentioned; 2 points when it was covered in more detail, but covered only selected facilities or used only self-comparison metrics and 3 points when information provided allowed comparison with other companies. Certain items, which were considered by the authors more important in a report, received up to four points, while others, such as the date of the most recent report, only one point. The considerably low results under the GRI G1 scoring system revealed that there was a substantial gap between what the assessed firms thought was appropriate to include in their report and what is hoped for by the Global Reporting Initiative.

The Canadian consulting firm Stratos Inc. performed three benchmark surveys during 2001–2005, examining the practices of corporate TBL reporting in Canada. With slight updates on the methodological approach over time, the Stratos scoring system is primarily based on the GRI reporting guidelines, with added criteria that refer to issues of relevance to Canadian companies, such as the company's management of relations and engagement with indigenous groups. In its latest version (2005) the assessment methodology consisted of 46 criteria, which fell into six major clusters (context and direction; policies, organization and management systems; stakeholder relations; performance information; upstream/downstream influence; quality, credibility and communications) and were assigned a rating of 0–3 points. When no meaningful information was provided on the specific criterion then the assessed report received no points; when coverage was patchy and information was not comprehensive enough the report was awarded with 1 point, while if the report provided "good information on the criterion" or "full coverage" 2 and 3 points were given respectively. The scoring process was conducted by two independent third party assessors and only clearly defined and specific references to other information referred to in reports were taken into account. General references to corporate websites were not cross-checked but only links to specific web pages and online documents were accessed.

The Institute for Sustainable Management (IfSM) at OAS University in Switzerland instigated a comprehensive analysis of the reporting procedures followed by Swiss companies (Daub 2007). It is the second study at a national level after the biennial survey of Stratos. The methodology applied in the Swiss study builds on the GRI guidelines and consists of 33 individual criteria falling into 4 categories; (a) context and coverage, (b) policies, management systems and stakeholder relations, (c) dimensions of performance and (d) transparency and general view. Each criterion received a rating between 0 and 3 points. When no meaningful information was disclosed no points were given to the specific criterion. Patchy statements were assigned 1 point, while "good information on a criterion with one relevant area/indicator not addressed" was awarded with 2 points. Finally, full disclosure to the assessed criterion was assigned the maximum score. The performance cluster of criteria in the IfSM evaluation system has a weighting factor of 2 since this section "…presents the hard facts on the performance of a company in the three sustainability dimensions" as Daub comments (Daub 2007, p. 82). Results revealed that, on average, the assessed companies scored one third of the total points possible while in the performance category of criteria companies attained the lowest average score per reporting category.

Methodological Weaknesses of Previous TBL Reporting Scoring System Methodologies

While all the aforementioned assessment procedures provide fruitful and useful insight on the orientation corporate reporting takes and can contribute to further improvements in disclosure practices, certain weaknesses of a methodological nature can be identified among them, which are discussed below.

Deloitte Touche Tohmatsu Sustainability Reporting Scorecard

The Deloitte Touche Tohmatsu sustainability reporting scorecard was developed during a period when non-financial reporting practices were still experimental to a large extent and there was a long way to go before relevant standards become widely accepted and broadly put into practice. Therefore it focuses mostly on quality and effective communicative elements rather than on quantitative data (indicators) and specific sets of information on business operation and impacts. In addition, there is a level of ambiguity when it comes to scoring the 30 criteria into which individual topics fall. For example, under the "corporate context" criterion a report may include some relevant information (e.g., nature of business and activities, organizational structure, main products/services, major markets served, major sites and operations, legal organization, financial and other key figures, geographical spread and scope of operations) and is awarded with 1 point when there is partial or unclear disclosure and 2 points when the coverage is considered average. But, how can partial and average coverage be distinguished and defined? No further explanation is provided. Furthermore, all scoring criteria are assigned with 3 points when coverage is "state-of-the-art" and 4 points when it is "pace-setting". Such qualification statements imply that reporting organizations as well as report users—stakeholders (and potential assessors of a TBL report) will have rather extensive experience in the field of non-financial reporting in order to distinguish "state-of-the-art" from "pace-setting" disclosures, something which is surely not always the case. Moreover, criteria on topics such as "data quality and accuracy", "management of significant risks and opportunities" have the same relative importance (weight) in the final grade as the "use of tools to support readability" or the way the report is structured. Finally, the Deloitte Touche Tohmatsu approach allows as much information to be provided on a particular topic (such as stakeholder engagement or performance) without significantly altering significantly the importance of that criterion/a in the final result.

SustainAbility/UNEP

The "Global Reporters" assessment methodology has been evaluating the practices of global corporations since 2000. As Daub denotes "these companies have adequate financial means at their disposal to draft detailed reports in terms of content and design. It is not to be assumed that small-to-medium-sized enterprises" (and with limited non-financial reporting awareness) "are in position to use these types of reports as models" (Daub 2007, p. 79). The scoring device, replicable across topics, is applied to evaluate a number of criteria under each topic. However, many of the topics contain several criteria to be assessed, which makes the scoring process a cumbersome and a very complex-complicated task.

For example, the environmental performance criterion on used materials and generated waste is defined by the following factors:

- Types of materials used, and their characteristics (e.g., hazardous, toxic, recycled)
- Content (renewable or non-renewable resources)
- Quantities of materials used
- Percentage of materials that are recycled and/or re-used
- Strategies to reduce the environmental footprint of materials used (e.g., external certification programs), recycling targets and performance
- Total amount of waste by type, and disposal methods
- Quantities of waste produced, transported, imported or exported

- Quantities of special waste—e.g., Basel Convention Annex I, II, III and VIII, national standards (e.g., RCRA in the US)
- Packaging quantity, by weight and type
- Waste and/or packaging recycling rates
- Waste-related hazards or risks

This criterion it is evaluated in terms of:

- Content—breadth of discussion and identification of materials used and their impacts, links to certification standards (e.g., Forest Stewardship Council); work on material efficiency (e.g., Factor Four); links to definitions (e.g., Basel Convention) or work on waste minimization or recycling
- Coverage—of the whole business unit and core products
- Presentation—logical breakdown and analysis of materials use data (e.g. according to business units or regions, by material type or disposition method)
- Interpretation—the company's performance is discussed in terms of past performance and industry performance,
- Implications—discussion of what the data means for the company's future plans and priorities (e.g., commitments to recycling or renewable goals, dematerialization and commitments to waste minimization).

It is evident that there is a lot of information in such a topic, whereas the relevant importance of each parameter is unspecified and unclear, resulting in increased subjectivity in the scoring process.

Pacific Sustainability Index

The Pacific Sustainability Index (PSI) stands out among the other scoring systems developed to date as it rates not only the disclosure practices of an organization but its level of performance as well. Such an integrated approach does provide a useful in-depth analysis of publicly available environmental and social information among different business sectors, but this article indentifies underlying methodological weaknesses. The PSI system does not assess information related to economic performance—one of the pillars of sustainability according to its conventional definition—nor important aspects of internal governance and structure, which can provide essential information among the various stakeholder groups. Moreover, according to the PSI's scoring criteria on performance, when the current performance of an assessed organization is superior to that previously reported, one point is assigned. However, no further explanation is provided in the case of organizations, which are first-time reporters or have just started to monitor their non-financial impact. It is the authors' assumption that such companies would receive no points on such topic(s), an approach that indicates a degree of bias against companies with established reporting mechanisms. In other words, applying the PSI to a sample of organizations that have just started to develop their internal TBL reporting systems may give misleading, underestimated results and conclusions.

Clausen and Others & IfSM Scoring Systems

Clausen and others (2005) and Daub (2007) use weighting factors among the different clusters of criteria of their assessment methodologies. However, in neither of these two evaluation systems is it clarified how these factors were derived. Such an approach is surely meaningful as it distinguishes among statements that most organizations are able to provide and the "hard facts," that is primarily quantitative performance indicators, which require a company to further invest in time and resources. The authors would argue however that a multi-weighted scoring system requires a sound and scientifically justifiable rationale in order to rigorously define the relevant importance of individual topics and components that a TBL report should include.

Stratos Inc

While the Stratos approach builds on the GRI guidelines, special issues of relevance to companies operating in Canada are incorporated in the assessment methodology. Indicators such as the total value of goods and services purchased from indigenous suppliers, the number of employees that have received indigenous awareness training, the number of participants in education and internship programs for indigenous people or the percentage of Alberta's Boreal forest disturbed by corporate operations can be of major importance for organizations operating in the Canadian context but prevent the application of this evaluation system to other regions or countries.

GRI 2002 Guidelines as an Evaluation Tool for TBL Reports

The Global Reporting Initiative Guidelines

The GRI is the result of a joint project between the Coalition for Environmentally Responsible Economies and the United Nations Environmental Program and is the most prominent institution in the TBL reporting context. Nowadays, at least 1000 organizations around the world have declared their use of the GRI proposed framework as the basis for their reports. The GRI Reporting Guidelines organize reporting in terms of economic, environmental, and social performance. This structure has been chosen as it reflects the most widely accepted approach to defining sustainability. Within the broader context of stakeholder engagement, the GRI's mission is to elevate the quality of reporting to a higher level of comparability, consistency, and utility. The purpose of the Guidelines, and the GRI framework as a whole, is to capture an emerging consensus on reporting practices. This provides a point of reference against which reporting organizations and report users can approach the challenge of developing effective and useful reporting practices (GRI 2002).

GRI published the first reporting guidelines (G1) in June 2000 and immediately afterwards, created working groups to assist in revising these guidelines. The second version (G2) of the guidelines was published in 2002 at the beginning of the Johannesburg Summit and gradually they have emerged as the world's de facto standard on non-financial reporting. In October 2006, the latest version of the GRI Guidelines (G3) was launched and replaced the G2 version, but as the GRI denotes, reports based on the 2002 Guidelines will continue to be recognized for two more reporting cycles.

Developing an Evaluation Tool Using GRI 2002 Guidelines as Benchmark

A report assessment methodology was devised using the GRI 2002 Guidelines as a point of reference. Report content according to this framework should cover the following major clusters of topics:

- Vision and strategy: the organization's approach to sustainable development;
- Company profile: an overview of the reporting organization and a brief description of the report's scope;
- Governance structure and management systems: an overview of the internal governing structure, overarching policies and management systems in place to implement the organization's vision for sustainable development and to manage its performance;
- Performance indicators (this section lists quantitative and qualitative indicators for presenting the organization's impacts on the economy, the environment and society). Performance indicators are grouped and structured according to a hierarchy of category, aspect, and indicator (Appendix, Table A1).

In total, the GRI 2002 Guidelines comprise of 141 individual topics falling into the four aforementioned sections. The authors devised a scoring system for each one of these topics and created an index of the GRI scoring criteria, following the structure and rationale of previously proposed non-financial reporting scoring systems. Each topic was allocated a score between 0 and 4 points, with the basic rating qualification scale set as shown in Table 1.

Certain topics in a report, such as the name of the reporting organization or the reporting period, received either 0 or 4 points. Similarly, other topics where the scope for detailed answers was limited were assigned with 0, 2 or

Table 1 Basic rating qualification scale

Points	Rating qualifications/requirements	Example—direct CO_2 emissions
0	The report does not include any information relevant to the specific GRI topic. No coverage	No relevant information is provided in the assessed report
1	The report provides generic or brief statements, without specific information on the organization's approach to the topic	We monitor our CO_2 emissions
2	The report includes valuable information on the topic but there are still major gaps in coverage. The organization identifies the assessed issue, but fails to present it sufficiently	In 2006, the Company's total emissions of CO_2 were equivalent to 800,000 tonnes
3	The provided information is adequate and clear. It is evident that the reporting organization has developed the necessary systems and processes for data collection on the assessed topic and attempts to present it in a consistent manner	Our Head Offices and plants in Greece produced 500,000 tonnes of CO_2, while the rest of our abroad operations result to 300,000 tonnes of CO_2
4	Coverage of the specific issue can be characterized as "full" in the report. It provides the organization's policy, procedures/programs and relevant monitoring results for addressing the issue. The organization meets the GRI requirements	In 2006, the Company's total emissions of CO_2 were equivalent to 800,000 tonnes. Our Head Offices and plants in Greece produced 500,000 tonnes of CO_2, while the rest of our abroad operations result to 300,000 tonnes of CO_2. This is a 5% reduction from last year's emissions. It is our stated commitment to reduce our CO_2 emissions by a targeted 25% by the end of 2008, compared to its 1990 level

Table 2 Number of topics, possible points, and weighting of each section

Section	Topics	Max. possible points	Weight (%)
Vision and strategy	2	8	1
Profile	22	88	16
Governance structure	20	80	14
Economic indicators	13	52	9
Environmental indicators	35	140	25
Social indicators	49	196	35
Labor practices	17	68	12
Human rights	14	56	10
Society	7	28	5
Product responsibility	11	44	8
Total	141	564	100

4 points. The number of topics, the possible points and the contribution of each section to the final grade is presented in the table below (Table 2). It is important to note that the methodology was sent for consultation and review to the Union of Environmental Scientists of Greece, which provided valuable comments for improvement and readjustment of scoring requirements.

Case Study: The Greek Reporters

Sample Identification

The next step was to test the proposed methodology on the Greek TBL reporters. Our initial aim was to gather all the reports with 2005 as the year of reference. Where these were not available, reports from 2004 were included. In order to identify all the Greek reporters a web-based search was conducted. This search initially focused on the GRI database where all the companies that use the Guidelines are invited to register their report. The second search stage concentrated on the firms which are core members of the Hellenic CSR Network and finally the research was expanded to other large Greek companies and multinational corporations operating in the country.

Among the 66 core members of the Network, 55 are industries and service providers while the rest are consultancies and business associations. Out of these 55 companies, only 15 (27%) published a report during 2004–2005. A few of the rest, as subsidiaries of large multinationals, referred to the global report of the parent company (British American Tobacco Hellas, DHL International Hellas, Johnson Diversey Hellas, and Shell Hellas). Only those reports prepared by companies with headquarters in Greece or by those that provided information and data on the company's operations-activities in the country were included in the analysis. For this reason, multinational corporations with subsidiaries in the country were excluded if they only produced a global, corporate-level report, with no breakdown of information at country level.

In total, 19 reports were identified, of which 16 were included in our sample (Appendix, Table A2). Three reports were excluded from the assessment (TIM Hellas, the National Bank of Greece and Halyps Cement) as well as a traditional annual report by the Bank of Cyprus which included a social responsibility section, since they were very brief, provided mostly generic statements and would have scored too low on the assessment thus would not have significantly contributed to the overall conclusions. Out of the 16 reporters, it was the first time seven (44%) had produced a report while the rest had had some reporting experience during the previous years. Moreover, and if assuming that a company predominantly selects the report title in order to communicate the issues it intends to cover, only one report was clearly defined as a "sustainability report", while the rest were mostly entitled as "social" or "CSR" reports.

Results

At a primary analysis level, a skim reading of each report was done in order to identify the major components, the scope and differences in disclosure practices. After the "pre-screening" stage, a detailed examination under the scoring criteria was implemented on scoring sheets. During this secondary stage of assessment, comments and explanations on the points assigned to each topic were noted, providing the referenced page numbers in the assessed reports as well.

The score results ranged from 263/564 (47%) to 66/564 (12%) (Appendix, Table A3). The average score achieved by the sample firms was around 21% of the maximum possible points, while considerable variation in the information provided and different practices in the presentation of issues covered was evident. Three different sub-groups of companies can be derived considering the average score of reports: those who achieved a score above average, those whose score is notably close to the average and a third group of reports which scored considerably lower than the sample's average points (Fig. 1).

Most of the points awarded are for presenting the profile and the governance structure of an organization (average scores 39% and 28% respectively), while significant environmental performance indicators (Fig. 2), issues of human rights and product responsibility are barely mentioned. Moreover, reports focus on labor practices related to health and safety, employee training and education (Fig. 3). In this respect, the GRI social performance indicators

Fig. 1 Ranking of Greek reporters against the GRI guidelines

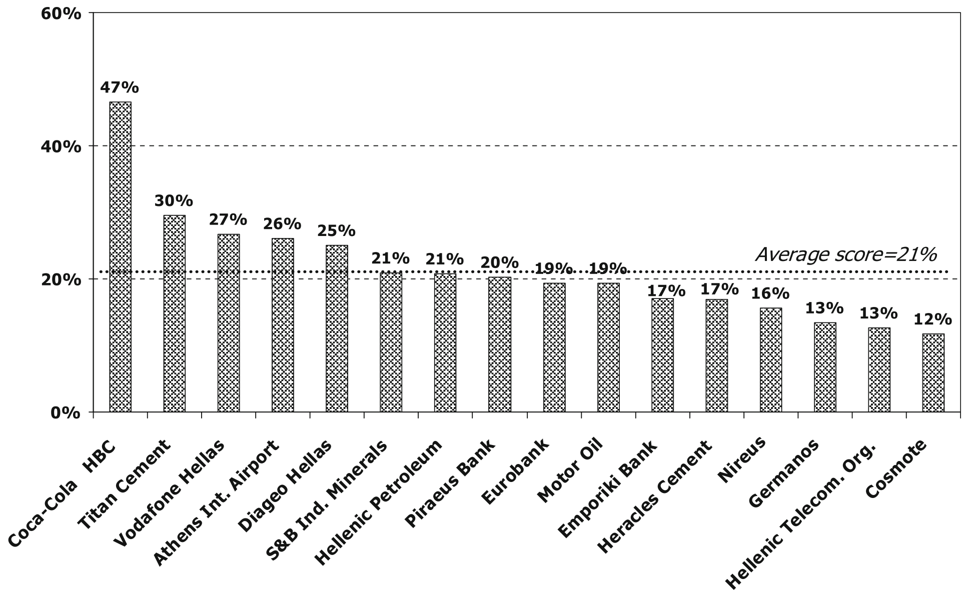
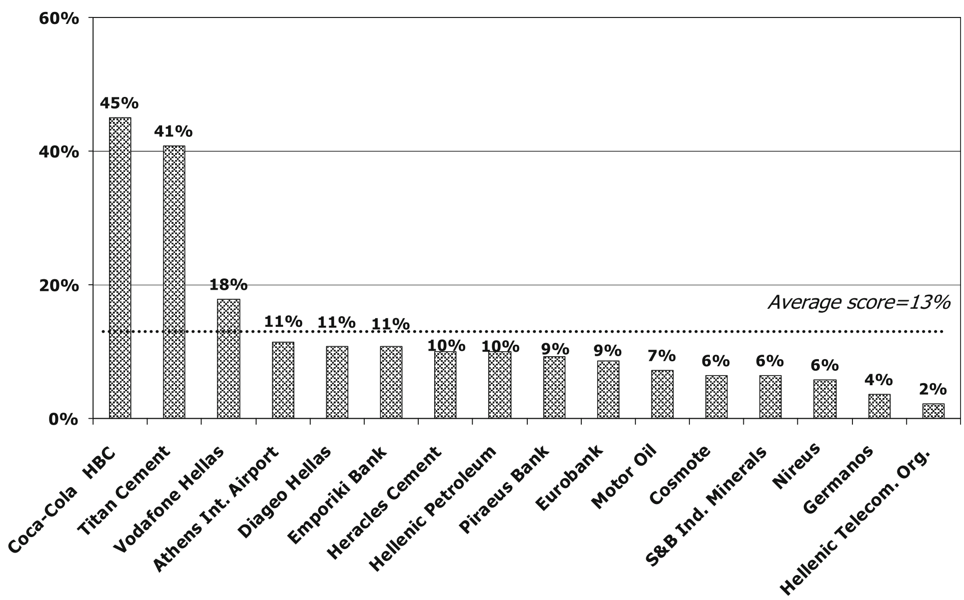


Fig. 2 Environmental performance indicators coverage



concerning the breakdown of the workforce, practices on recording occupational accidents, injury rates, average hours of training per year/employee, are covered by the majority of reporters. Economic performance is poorly presented (average score 23%), probably because most companies are content to leave their financial performance information to their annual financial reports; only turnover/net sales, philanthropy-charitable contributions and donations to community are reported by all the firms analyzed. Considering the environmental dimension of the triple bottom line, all reports describe or simply mention internal initiatives to increase energy efficiency or to use renewable energy sources (which comprise an additional GRI environmental indicator) as well as recycling materials. Many of them provide data on direct energy used and air emissions, but very few cover essential issues, such as total materials used, total amount of waste and significant environmental impacts of principal products and services.

Taking into account the core performance indicators—those that the GRI 2002 Guidelines stress are relevant to most organizations and the information contained therein most relevant to stakeholder groups—ranking of Greek reports changed accordingly (Fig. 4). Titan's report received the second highest overall score, but dropped to 6th place in the core indicators coverage ranking. Even though S&B's report is ranked 6th in the overall results, under the coverage of core indicators it received a score significantly lower than the sample's average. Similarly, while Hellenic Petroleum's report followed S&B's in the overall scores, when considering the core performance

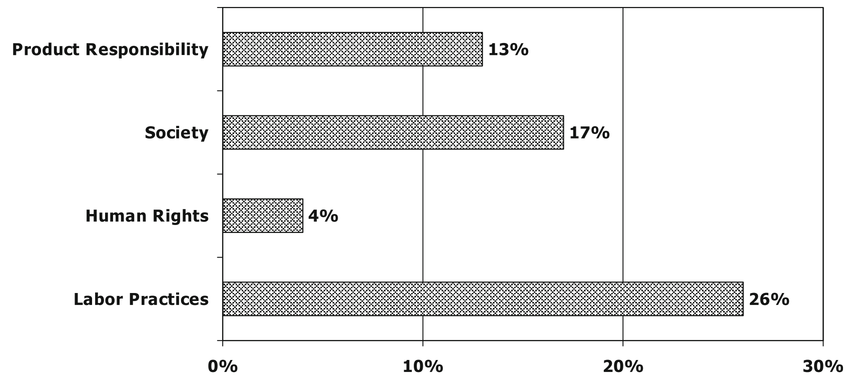


Fig. 3 Average benchmark results in the four clusters of social performance indicators

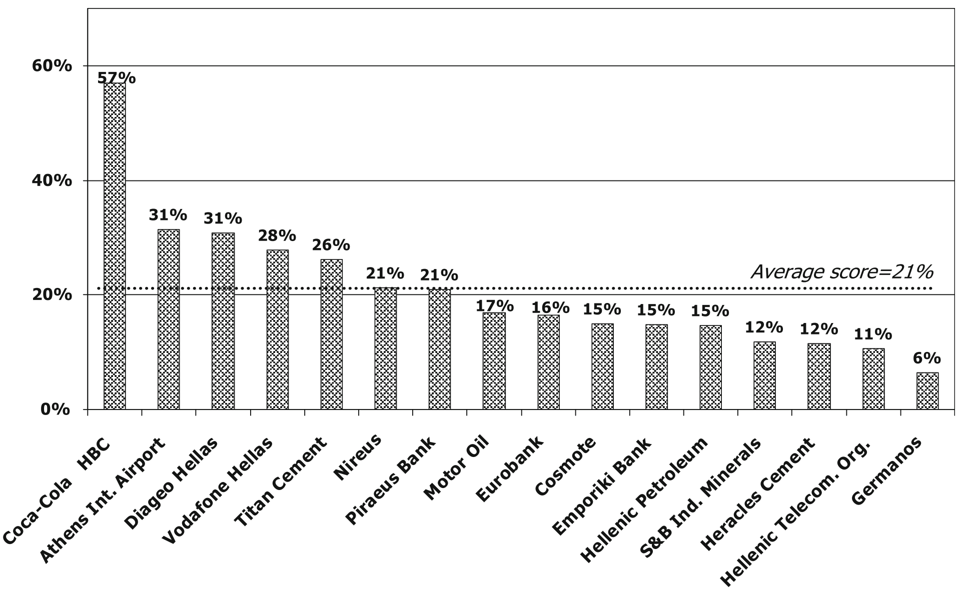


Fig. 4 Ranking of Greek reports based on core performance indicators coverage

information it covered, it is in 12th place. In contrast, Nireus's report is in 13th place in the overall ranking, but in the cluster of core performance criteria moves up to 6th place. Cosmote's report that got the lowest score in the overall ranking, achieved a comparatively higher score under the set of core indicators moving up to 10th place.

Discussion

As denoted in almost all scoring systems for non-financial reports developed so far, the proposed assessment methodology does not directly evaluate corporate performance, but concentrates solely on the reporting practices and the breadth and depth of issues reported. Organizations may have exceptional performance and/or promote governance transparency, but if this is not reflected in their reports they will score low and be ranked accordingly. Morhardt (2001) stresses that "a high score does not necessarily reflect a high level of environmental performance, just the fact that a large number of the topics in the scoring systems was addressed, irrespective of whether the environmental performance was good or bad". Moreover, it is the nature of the benchmark tool that organizations assessed can simply include more topics, superficially discuss them and gain more points than others that choose to include fewer issues and cover them thoroughly. For example in our survey on Greek reporters, Diageo Hellas was ranked higher than others under the cluster of human rights indicators just because it provided only relevant generic statements. Similarly, Emporiki Bank that covered more environmental performance indicators than S&B Industrial Minerals—which was fifth in the overall score results (21%)—was somewhere in the middle of the ranking (scoring a 17% compliance against the GRI assessment methodology).

In this respect, companies can obtain a higher score by addressing more topics on the profile and governance structure and not on essential information on performance, something that was evident in the majority of Greek reports. The changes in the ranking of Greek reports under the core performance set of indicators are indicative that in an evaluation system such as the one currently discussed it is important to closely examine the total scores under specific clusters of criteria or even under specific topics. A company may score well for presenting information on the overview of its operations or its corporate structure rather than its performance during the reporting period, which is probably the most essential part of this kind of report. A high total score does not necessarily mean that a report is inclusive and balanced; this can be confirmed by closely examining the components of this score. Therefore, it is important to scrutinize the total score as overall results can give misleading or superficial conclusions. Two or more organizations may achieve the same score but the "components" of this score may vary considerably. However, a comparatively higher score (like Coca-Cola HBC's 47% in our assessment) *does* mean that the reporting organization discussed more issues and/or provided more detailed information. Still, a 100% compliance result against the GRI Guidelines is clearly theoretical and overestimates the context of an assessed report that receives such a score.

GRI includes a broad set of topics to cover which can not be addressed to the same extent and equally by all organizations. Companies among different sectors have differences in the intensity of their environmental and/or social impact, as denoted by Krut and Munis (1998) and quoted by Morhardt and others (2002). Therefore, it can be useful to apply the GRI evaluation tool to sector studies, examine the coverage of GRI core performance indicators among the sample firms and further analyze the reporting practices of organizations with the same activity under the GRI sector supplement (if available for the sector being assessed), which can be potentially converted to a scoring system as well. As GRI points out, core performance indicators are relevant to most reporting organizations—regardless of the sector they operate in, and are those that interest the majority of stakeholder groups (GRI 2002).

The evaluation tool constructed focuses primarily on the social and environmental performance as defined by the GRI indicators (since they constitute 35% and 25% of the final grade respectively. Such an approach is based on the notion that TBL reports should include a clear discussion of quantifiable statements on the company's impacts and specific results of internal management systems, rather than merely referring to the external commitments of the top management without any strong evidence that the organization is moving in that direction and is implementing relevant policies or effectively communicates the "implementation likelihood" (a concept introduced by Kolk 2004; according to the implementation likelihood, a TBL report can be judged on the likelihood that its contents have indeed been implemented within an organization). Unlike other non-financial reporting assessment methodologies, it is exclusively based on the most prominent reporting framework. It can help readers to evaluate and compare reports, as well as organizations that prepare them to examine the quality and comprehensiveness of the information provided, according to an internationally accepted set of guidelines, which suggest a standardized structure and context of reports.

While the GRI 2002 Guidelines was one of the components in constructing the methodology, the other was the rationale of the scoring systems previously applied in international studies. As pointed out by Deloitte Touche Tohmatsu (2002) "a five-level scoring is the most robust scoring from a statistical point of view." It was noted that in many topics-indicators (that mostly refer to qualitative information) scoring up to 3 points was sequential, but requirements for the maximum points were comparatively more. While the gap between 0, 1, 2, 3 points was to some extent equal, in order to get 4 points, the stipulations were much more demanding. For example, an organization can receive up to 3 points if it discusses its policy and relevant internal procedures for addressing a certain issue, but in order to gain the additional point it has to present its monitoring systems for the implementation of this policy as well as related results of the monitoring process.

The development of such a scoring system also proved to be a dynamic process and directly related to the sample reports. Rating qualifications of many topics–criteria were readjusted several times during the assessment of Greek reports in order to provide a more justifiable scoring system. This was more apparent in the case of quantitative indicators, where some organizations reported on data referring to selected facilities, while a few others included total numbers and/or ratios. Nevertheless, in a report assessment under a numerical system there is always a degree of subjectivity, which can be alleviated by more than one assessor scoring the reports and further examining any strong variations among individual scores. In addition, this kind of assessment does not examine the accuracy and validity of the information, nor does it use cross-references to other corporate publications and information available on the internet. Only in the case where links and references to specific webpages or other publicly available information about the organization assessed are made, are these then included in the scoring process. For example, Coca-Cola HBC and Athens International Airport provided specific links to quantitative data of environmental performance, which are available on the organzations' websites. Similarly, Titan Cement included a link to a publication that presents the company's internal values and code of conduct. While in the assessment of early Greek reporters only one report of each firm was included, it is feasible to examine the evolution of the non-financial reporting practices of a certain sample of companies or even an individual one within a timeline. This can reveal any potential progress in the preparation of such documents and the extent to which organizations learn over the years how to report their TBL performance and promote stakeholder communication and engagement.

The proposed scoring system is constructed under the hypothesis that all GRI topics have the same weight to the final grade; each one of them is scored between 0 and 4 points. This approach was chosen because the authors believe that a multi-weighted scoring system requires a procedure with many complex tasks in order to provide proper weights to each topic on the grounds of sound and scientifically justifiable arguments. Such tasks involve focus groups, engagement with stakeholder groups and with the potential candidates for assessment in order to examine the importance of individual GRI topics and indicators. While this is without doubt a long-term project, the scope of the research would reduce assessors' subjectivity. In our hypothesis, certain topics such as referring to the date of the most recent report were assigned 0 or 4 points. Likewise, another set of topics scored 0, 2 or 4 points, since a five level scaling of the relevant information was not feasible and they were mentioned infrequently or not at all in Greek reports. In future applications of the methodology, given the premise that reports assessed are more comprehensive, it is likely that the scoring devise of these topics will readjust.

Concluding Remarks

As the literature search revealed, most of the previous non-financial reporting assessments have concentrated on the state-of-the-art reports while the methodological approaches to do so have been developed accordingly. Countries (and sectors) which are taking their first steps in TBL reporting are under-represented within these surveys, and in the future more research on new reporting entities would be beneficial. The GRI guidelines as a numerical scoring system proved to be a very thorough and rigorous evaluation tool revealing major gaps in the Greek reports' comprehensiveness; fundamental reporting issues were omitted and significant improvements need to be made.

Research on non-financial accounting and reporting can and should be advanced on its own merit. In this context, the recent study by Archel and others (2008) which provides a great deal of insight of yet unexplored technical developments of TBL reporting should be mentioned along with the vast array of sector studies by the Roberts Environmental Center of Claremont McKenna College. Furthermore, future research needs not only to refine and improve on the aforementioned methodological limitations but also to perform comparisons of reporting practices between different countries, a literature field very thin on the ground. This article provides a useful benchmark tool and a starting point for such an examination. Just as the consensus on reporting practices is gradually gaining broader acceptance, a common approach to assess this new form of "business card" is also advisable.

Acknowledgments Thanks are due to three anonymous referees for their useful comments. The authors are also indebted to Philippa Harris for her proofreading of the English. Any remaining errors on any aspect of the paper are solely the authors' responsibility.

Appendix

See Tables A1, A2 and A3.

Table A1 The hierarchy of GRI performance indicators according to category and aspect

	Category	Aspect
Economic	Direct economic impacts	Customers
		Suppliers
		Employees
		Providers of capital
		Public sector
Environmental	Environmental	Materials
		Energy
		Water
		Biodiversity
		Emissions, effluents, and waste
		Suppliers
		Products and services
		Compliance
		Transport
		Overall
Social	Labor practices	Employment
		Labor/management relations
		Health and safety
		Training and education
		Diversity and opportunity
	Human rights	Strategy and management
		Non-discrimination
		Freedom of association and collective bargaining
		Child labor
		Forced and compulsory labor
		Disciplinary practices
		Security practices
		Indigenous rights
	Society	Community
		Bribery and corruption
		Political contributions
		Competition and pricing
	Product responsibility	Customer health and safety
		Products and services
		Advertising
		Respect for privacy

Source: Global Reporting Initiative Sustainability Reporting Guidelines (2002)

Table A2 List of companies included in the benchmark survey

n/n	Organization name	Sector	Web address
1	Athens International Airport	Transportation	http://www.aia.gr
2	Coca-Cola HBC	Food and beverage	http://www.coca-colahbc.com
3	Cosmote	Telecommunications	http://www.cosmote.gr
4	Diageo Hellas	Food and beverage	http://www.diageo.com
5	Emporiki Bank	Financial services	http://www.emporiki.gr
6	Eurobank	Financial services	http://www.eurobank.gr
7	Germanos	Retail	http://www.germanosgroup.gr
8	Hellenic Telecom. Org. (OTE)	Telecommunications	http://www.ote.gr
9	Hellenic Petroleum	Oil and gas	http://www.hellenic-petroleum.gr
10	Heracles Cement	Construction and materials	http://www.aget.gr
11	Motor-Oil	Oil and gas	http://www.moh.gr
12	Nireus	Food and beverage	http://www.nireus.com
13	Pireos Bank	Financial services	http://www.piraeusbank.gr
14	S&B Industrial Minerals	Basic resources	http://www.sandb.com
15	Titan Cement	Construction and materials	http://www.titan-cement.com
16	Vodafone Hellas	Telecommunications	http://www.vodafone.gr

Table A3 Ranking of Greek reports according to the GRI-based benchmark assessment

n/n	Organization	Report title	Year	Score (max 564 points)	Score (%)
1	Coca-Cola HBC	Social responsibility	2005	263	47
2	Titan Cement	Corporate Social Responsibility & Sustainability	2005	167	30
3	Vodafone Hellas	Corporate responsibility	2004–2005	151	27
4	Athens Int. Airport	Corporate responsibility	2005	147	26
5	Diageo Hellas	Corporate citizenship	2004	141	25
6	S&B Ind. Minerals	Social	2005	118	21
7	Hellenic Petroleum	Social & Environmental	2005	117	21
8	Piraeus Bank	Corporate social responsibility	2005	114	20
9	Eurobank	Corporate social responsibility & sustainability	2005	109	19
10	Motor Oil	Environmental & social	2005	109	19
11	Emporiki Bank	Corporate social responsibility	2005	96	17
12	Heracles Cement	Sustainability	2005	95	17
13	Nireus	Social & environmental	2005	88	16
14	Germanos	Corporate social responsibility	2004	76	13
15	Hellenic Telecom. Org.	Corporate social responsibility	2005	71	13
16	Cosmote	Corporate responsibility	2005	66	12

References

ACCA, CorporateRegister.com (2004) Toward transparency: progress on global sustainability reporting 2004. ACCA, London

Accountability (2004) Strategic challenges for business in the use of corporate responsibility codes, standards, and frameworks. World Business Council for Sustainable Development 48

Adams CA, McNicholas P (2007) Making a difference: Sustainability reporting, accountability and organisational change. Accounting, Auditing and Accountability Journal 20:382–402

Albelda-Perez E, Correa-Ruiz C, Carrasco-Fenech F (2007) Environmental management systems and management accounting practices as engagement tools for Spanish companies. Accounting, Auditing and Accountability Journal 20:403–422

Antal AB, Sobczak A (2007) Corporate social responsibility in France: a mix of national traditions and international influences. Business and Society 46:9–32

Archel P, Fernández M, Larrinaga C (2008) The organizational and operational boundaries of triple bottom line reporting: a survey. Environmental Management 41:106–117

Clausen J, Loew T, Westermann U (2005) Sustainability reporting in Germany: summary of the results and trends of the 2005 ranking. http://www.ranking-nachhaltigkeitsberichte.de. Accessed 10 Oct 2006

Context (2006) Reporting in context: Global corporate responsibility trends 2006. Context, London

CSR Network (2003) Material World: The 2003 benchmark survey of global reporting. CSR Network Limited, CSR Network, Bath

Daub CH (2007) Assessing the quality of sustainability reports: an alternative methodological approach. Journal of Cleaner Production 15:75–85

Davis-Walling P, Batterman SA (1997) Environmental reporting by the fortune 50 firms. Environmental Management 21:865–875

Deloitte Touche Tohmatsu (2002) Deloitte Sustainability Reporting Scorecard. Deloitte Touche Tohmatsu Global Sustainability Group. http://www.deloitte.com/dtt/cda/doc/content/DTT_ERS_FullScorecard_032106.pdf. Accessed 20 October 2006

Elkington J (1997) Cannibals with forks: the triple bottom line of the 21st century. Capstone Publishing, Oxford

Global Reporting Initiative (GRI) (2002) Sustainability reporting guidelines. Global Reporting Initiative, Boston

Holland L, Foo YB (2003) Differences in environmental reporting practices in the UK and US: the legal and regulatory context. British Accounting Review 35:1–18

Kaptein M, van Tulder R (2003) Toward effective stakeholder dialogue. Business and Society Review 108:203–224

Kolk A (2003) Trends in sustainability reporting by the Fortune Global 250. Business Strategy and the Environment 12:279–291

Kolk A (2004) A decade of sustainability reporting: developments and significance. International Journal Environment and Sustainable Development 3:51–64

Kolk A (2005a) Social and environmental accounting. In: Clubb C (ed) The Blackwell encyclopedia of management accounting, 2nd edn. Blackwell, Malden, pp 393–398

Kolk A (2005b) Environmental reporting by multinationals from the Triad: convergence or divergence? Management International Review 45:145–166

KPMG (2005) KPMG international survey of corporate sustainability reporting 2005. KPMG, Holland

Krut R, Munis K (1998) Sustainable industrial development: benchmarking environmental policies and reports. Greener Management International 21:87–98

Logsdon MJ, Lewelynn PG (2000) Expanding accountability to stakeholders: trends and predictions. Business and Society Review 105:419–435

Mathews MR (1995) Social and environmental accounting: a practical demonstration of ethical concern? Journal of Business Ethics 14:663–671

Morhardt JE (2001) Scoring corporate environmental reports for comprehensiveness: a comparison of three systems. Environmental Management 26:881–892

Morhardt JE, Baird S, Freeman K (2002) Scoring Corporate Environmental and Sustainability Reports Using GRI 2000, ISO 14031 and other Criteria. Corporate Social Responsibility and Environmental Management 9:215–233

Palenberg M, Reinicke W, Witte JM (2006) Trends in non-financial reporting. Paper prepared for the United Nations Environment Programmes, Division of Technology, Industry and Economics (DTIE), November 2006, Global Public Policy Institute, Berlin, Germany. www.gppi.net. Accessed 20 August 2008

Rasche A, Esser DE (2006) From stakeholder management to stakeholder accountability. Journal of Business Ethics 65: 251–267

Repetto R (2005) Protecting investors and the environment through financial disclosure. Utilities Policy 13:51–68

Rikhardsson P, Holm C (2008) The effect of environmental information on investment allocation decisions—an experimental study. Business Strategy and the Environment 17:382–397

Roberts Environmental Center (2006) The Pacific Sustainability Index scoring system. http://www.roberts.cmc.edu. Accessed 15 December 2006

STRATOS Inc (2005) Gaining momentum: corporate sustainability reporting in Canada. http://www.stratos-sts.com. Accessed 18 December 2006

UNEP/SustainAbility (2006) Tomorrow's value: the global reporters 2006 survey of corporate sustainability reporting. Sustainability, London

Waddock S (2004) Creating corporate accountability: foundational principles to make corporate citizenship real. Journal of Business Ethics 50:313–327

World Business Council for Sustainable Development (WBCSD) (2002) Sustainable development reporting: striking the balance. World Business Council for Sustainable Development, Geneva

Woodward DG, Edwards P, Birkin F (1996) Organizational legitimacy and stakeholder information provision. British Academy of Management 7:329–347

Zadek S (2001) The civil corporation: the new economy of corporate citizenship. Earthscan publications, London