The current issue and full text archive of this journal is available at
www.emeraldinsight.com/2040-8269.htm

A conceptual framework for the quality evaluation of sustainability reports

Michael Freundlieb, Matthias Gräuler and Frank Teuteberg
Research Group in Accounting and Information Systems, University of Osnabrück, Osnabrück, Germany

Abstract

Purpose – This paper aims to outline a conceptual framework for the quality evaluation of web-based sustainability reports (SRs) aiding managers in determining and evaluating quality criteria for the sustainability report of their company.

Design/methodology/approach – The paper reviews quality criteria in existing standards and guidelines on sustainability reporting and identifies research gaps. A conceptual framework including a multi-method approach for the quality evaluation of SRs is developed and evaluated.

Findings – Existing standards and guidelines on sustainability reporting mainly focus on the content of the reports and neglect common information systems (IS) acceptance criteria such as ease of use and visual appeal. The proposed framework directly involves different stakeholder groups and research methodologies into the quality evaluation process.

Research limitations/implications – The limitations of the research approach offer a number of starting points for future research. The proposed framework needs to be further evaluated by a larger number of test users in a more natural use setting.

Practical implications – The application of a multi-method approach as well as the direct involvement of the stakeholders allows for an in-depth quality evaluation of SRs, enabling reporting companies to meet the readers' demand for information on economic, environmental and social activities of the reporting company. Common acceptance factors from the field of IS should be integrated into existing standards and guidelines on sustainability reporting. Coaching of the users through help functions, wizards, instructional videos or avatars is desirable.

Originality/value – The proposed framework applies innovative technologies such as eye-tracking and software-supported attention analysis. By applying the framework to a set of sample reports, its usefulness and applicability are demonstrated.

Keywords Quality, Reporting, Sustainability, Evaluation, Conceptual framework, Multi-method approach

Paper type Research paper

1. Introduction

Information systems (IS) researchers have identified substantial research gaps regarding the role of IS in the sustainable development of organizations: Melville (2010) suggests to

This work is part of the project "IT-for-Green" (Next Generation CEMIS for Environmental, Energy and Resource Management). The "IT-for-Green"-project is funded by the European regional development fund (Grant Number W/A III 80119242). The authors greatly appreciate the support provided by all involved project partners. They would also like to thank the participants of the experimental evaluation, industry workshops and expert interviews as well as the anonymous reviewers and Marita Imhorst who provided useful insights, support and substantive feedback during the research process.

Management Research Review
Vol. 37 No. 1, 2014
pp. 19-44

2040-8269
DOI 10.1108/MRR-04-2012-0087

further investigate how IS influence beliefs about environmental sustainability, whereas Watson *et al.* (2010) ask what information should be communicated to consumers to advance desirable changes in behaviour. The research questions proposed by Melville and Watson *et al.* share the belief that the communication of sustainability information to and among different stakeholder groups play a key role in the sustainable development of organizations and should thus be further investigated by the IS community.

Sustainability reports (SRs) are "public reports by companies to provide internal and external stakeholders with a picture of the corporate position and activities on economic, environmental and social dimensions" (Heemskerk *et al.*, 2002). Therefore, SRs provide stakeholders with information about a company's efforts to balance its economic, ecological and social goals, which is often referred to as the "triple bottom line" (Elkington, 1998). In this context, sustainability can be defined as "meeting the needs of the present without compromising the ability of future generations to meet their own needs" (Brundtland Commission, 1987). Since the late 1980s, for most of the world's largest companies, reporting on sustainability has become everyday business (Kolk, 2004; Slater, 2008) which appears to be a continuing trend (Bartels *et al.*, 2008; Isenmann *et al.*, 2007; KPMG, 2011) even extending to developing countries (Amran and Haniffa, 2011). Empirical evidence suggests that small and medium sized companies do not report as successfully as large corporations. This may stem from the failure to see the competitive advantages of sustainability reporting and the adoption of a passive attitude regarding environmental improvements or a lack of stakeholder interest (Morhardt, 2010; Stubbs *et al.*, 2012). Scientists and practitioners alike are struggling to find a common and fitting terminology for matters concerning corporate sustainability (van Marrewijk, 2002), therefore, SRs may also be referred to as corporate social responsibility (CSR), corporate social and environmental responsibility (CSER) or corporate citizenship reports.

Studies reveal that SRs have the potential to improve a company's image, influence consumers' buying decisions as well as shareholders' investment decisions, and they play a role in the selection of business-to-business partners (Bartels *et al.*, 2008; Clacher and Hagendorff, 2012; Townsend *et al.*, 2010; Wheeler and Sillanpää, 1998). CSR practices, such as sustainability reporting, serve the purpose of protecting the reporting company's reputation and brands and thereby mitigate the risks to which stakeholders (e.g. investors, employees) are exposed (Welford and Frost, 2006). Thus, by offering a high quality SR to interested parties, decisive competitive advantages can be achieved, making SRs an important instrument of corporate communication and a key factor in the sustainable development as well as the decision making of companies and stakeholders alike. Since the number of companies publishing their SR online is increasing (Isenmann *et al.*, 2007; Slater, 2008), web based SRs offer an interesting field of research for the IS community.

SRs need to fulfill the requirements of many different internal and external stakeholder groups, such as employees, supply chain partners, regulatory bodies, consumers and the general public (El-Gayar and Fritz, 2006; Watson *et al.*, 2010). Taking this into account, the design and implementation of a SR represent a major challenge. Support is provided by a number of standards and guidelines on sustainability reporting. These include quality criteria as guidance for managers who seek to meet the expectations of their companies' stakeholders. However, even when following

these guidelines how can an organization be sure that its SR provides the added value stakeholders expect or even demand?

Acknowledging the relative scarcity of research investigating SRs by employing experimental approaches, this paper follows a design science research approach (Hevner *et al.*, 2004) that develops and evaluates a conceptual framework for the quality evaluation of web based SRs. The framework is able to incorporate multiple experimental setups while focusing on the design criteria demanded by stakeholders. Subsequently, as investigative methods, the conceptual framework is evaluated by conducting an eye tracking study, a software-supported analysis as well as a questionnaire. The objective of this research is not primarily the quality assessment of a SR but the introduction of the conceptual framework. The framework represents a novel approach to quality assessment of SRs and applies to concept, development and operation phases of the design science approach.

The following Section 2 explains the applied research approach and how the results presented later were achieved. Section 3 gives an overview of existing standards and guidelines on sustainability reporting and outlines the main goals and quality criteria mentioned within these guidelines. After reviewing the quality criteria and identifying (research) gaps, we propose a conceptual multi-method framework that directly involves the different internal and external stakeholder groups into the definition and evaluation of quality criteria. In Section 4, the developed framework and its construction is described in more detail. In order to evaluate our framework, it is applied to a number of sample reports. In Section 5, we present the key findings from the eye-tracking and software-supported analyses as well as a short questionnaire. The findings concentrate on the framework itself, the design and implementation of SRs as well as implications for IS researchers and practitioners. In Section 6, we sum up our results and outline the limitations of our research approach as starting points for future research.

2. Research methodology

In the IS discipline, the creation of artefacts, such as the presented conceptual framework, is usually associated with design science research (Hevner *et al.*, 2004; March and Smith, 1995). Following the design science research approach, this paper describes the design, prototypical implementation and evaluation of a conceptual framework for the quality evaluation of SRs. The applied research methodology and process is shown in Figure 1.

To achieve the necessary degree of rigor, we started our research with a literature review, which we explain in Section 3. Using the results from the previous literature review, we derived quality criteria from various sources and took them into consideration while constructing the framework. Prior to the research, three industry workshops with more than 80 participants from science and practice were held and 13 experts were interviewed to capture the current challenges and requirements of sustainability reporting. In this process, it became obvious that certain development stages needed to be revisited as some important points were not included in the first iteration of the development phase. Thereby, our research followed an evolutionary approach on prototyping (Davis, 1992). During the evaluation phase, we conducted two pre-tests to ensure that the used technology worked flawlessly and provided utilizable results. Subsequently, we conducted a software-supported analysis as well

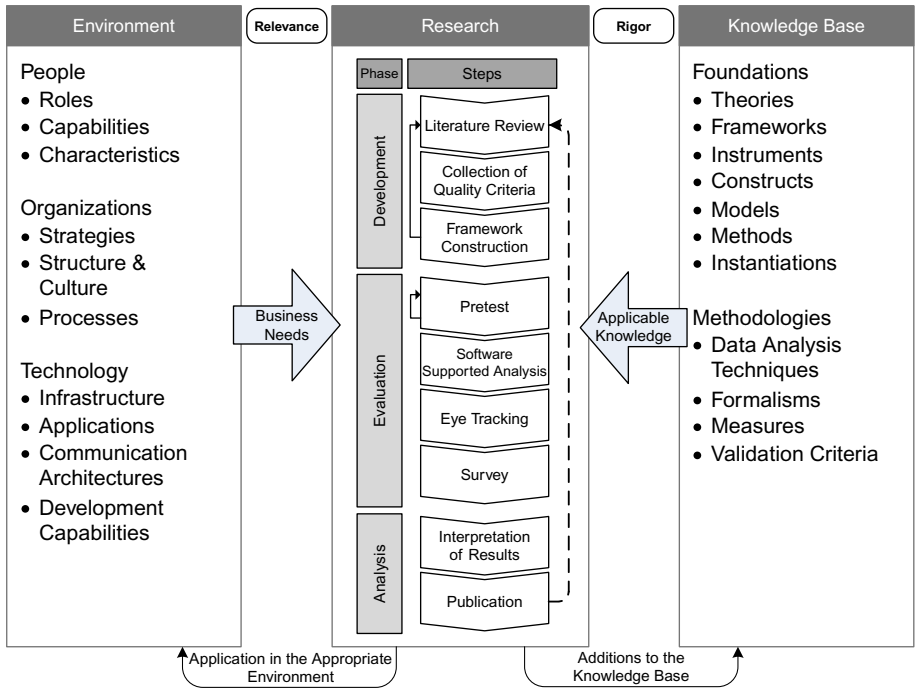


Figure 1.
Research process

Source: Hevner *et al*. (2004)

as an eye-tracking study and a survey on each of the participants. Finally, the results were interpreted and processed for publication.

Our research draws substantially from the existing body of knowledge regarding sustainability reporting, quality evaluation and technology acceptance as well as several IS theories which will be discussed later. Throughout the research process, the applicable knowledge has been extracted and integrated rigorously. The results will then complement the existing body of knowledge. As the resulting conceptual framework can be used for the quality evaluation of any SR, reporting entities are enabled to create SRs adjusted to their stakeholders' requirements, which provides relevance to the research.

3. Review of quality criteria for SRs

In order to identify scientific literature dealing with quality criteria for and/or the evaluation of SRs, we conducted a systematic literature review (Webster and Watson, 2002). Since sustainability reporting is increasingly performed online (Isenmann *et al.*, 2007; Slater, 2008), our literature review focuses on an IS perspective: we searched the top 20 journals of the *Management Information Systems* journal ranking of the Association for Information Systems (2009) as well as the highly ranked and peer-reviewed IS conferences Americas Conference on Information Systems (AMCIS), European Conference on Information Systems (ECIS), Hawaii International Conference on System Sciences (HICSS), International Conference on Information Systems (ICIS)

and Wirtschaftsinformatik (WI) (CORE, 2008) for the search terms "quality" and/or "evaluation" in combination with the terms "sustainability reporting" or "CSR reporting". Our search did not lead to relevant results, which is in line with other researchers who have concluded that the sustainable development and environmental sustainability have not yet been adequately explored by the IS community (Melville, 2010; Watson *et al.*, 2010). Hence, we focused on more practically oriented standards and guidelines for SRs, which are often referred to when rankings or scorings of SRs are performed (Daub, 2007) and searched other databases such as Scopus, ISI Web and Google Scholar. By extending the scope of our literature research, we were able to include works adopting a management IS perspective. Web based SRs are examples of IS artifacts which can be generalized as web sites and reports; therefore, we again extended our search to include research on the usability of web sites and information visualization, as these topics share common ground with the usability and acceptance of SRs.

Table I provides an overview of the standards and guidelines we considered most relevant to our work.

We believe that the development of a conceptual framework for the quality evaluation of SRs requires that the reports' objectives are clearly apparent. Table II gives an overview of the reporting goals stated in the above listed guidelines and norms.

Organization	Background
Global Reporting Initiative (GRI, 2011)	The GRI is a guideline on the content and design of SRs which is based on the "consensus seeking consultation" of different stakeholder groups such as business, labor, non-governmental organizations, investors and accountancy Thirteen quality criteria for SRs are developed and according test procedures are proposed. Also, a reference outline including concrete key performance indicators for SRs is suggested
World Business Council for Sustainable Development (WBCSD, 2003)	The approach of the WBCSD is similar to that of the GRI, but it is of a more general and recommendatory character. The recommendations primarily concern the information needs of the stakeholders as well as the contents of SRs. In order to motivate non-reporters to publish a SR, especially successful companies are referred to and citations from their top management are given
Eco-Management and Audit Scheme (EMAS, 2009)	EMAS is a regulation regarding the environmental management and environmental audits of organizations which goes beyond the requirements of the international ISO 14001 standard. It applies to European companies only. Annex IV of EMAS includes regulations on environmental reporting, especially regarding environmental statements and key performance indicators, as well as availability and accountability
International Organization for Standardization (ISO, 2006)	The ISO standard 14063 suggests five quality principles for environmental management and environmental communication and provides a survey of different channels for environmental communication as, for example, SRs. A reference procedure for environmental communication is described, but there are no concrete suggestions regarding content or key performance indicators
AccountAbility 1000 (AA1000, n.d.)	The AA1000 set of standards is based on three principles: inclusivity, materiality and responsiveness. The series of standards consists of the AA1000 Principles Standard, providing a framework for organizations to handle sustainability issues, the AA1000 Assurance Standard, aiming to help evaluate the extent to which the AA1000 principles are applied in an organization, and the AA1000 Stakeholder Engagement Standard, providing a framework for efficient stakeholder engagement

Table I.
Sustainability reporting standards and guidelines

Table II.
Internal and external goals of sustainability reporting

Organization	Internal goals	External goals
Global Reporting Initiative (GRI, 2011)	Benchmarking and assessing sustainability performance with respect to laws, norms, codes, performance standards, and voluntary initiatives Comparing performance within an organization over time	Provide a balanced and reasonable representation of the sustainability performance of a reporting organization – including both positive and negative contributions Demonstrate how the organization influences and is influenced by expectations about sustainable development Compare performance between different organizations over time
World Business Council for Sustainable Development (WBCSD, 2003)	Showing a balanced and reasonable presentation of an organization's economic, environmental and social performance ("true and fair view", "presented fairly")	
Eco-Management and Audit Scheme (EMAS, 2009)	The active involvement of employees in organizations and appropriate training	Promote continuous improvements in the environmental performance of organizations by the establishment and implementation of environmental management systems by organizations The systematic, objective and periodic evaluation of the performance of such systems The provision of information on environmental performance An open dialogue with the public and other interested parties
International Organization for Standardization (ISO, 2006)	Providing inputs/suggestions for improving the environmental performance of an organization's activities, products and services, and progress toward sustainability Raising the importance and level of environmental awareness to support an environmentally responsible culture and values within the organization	Promoting an organization's environmental credentials, achievements and performance Increasing business support and shareholder confidence Enhancing interested parties' perceptions of the organization Addressing interested parties' concerns and complaints about operational and emergency environmental hazards Improving understanding of interested parties' needs and concerns to foster trust and dialogue Assisting interested parties in understanding an organization's environmental commitments, policies and performance
AccountAbility (AA1000, n.d.)	Aligning the non-financial aspects of sustainability with financial reporting and assurance Leveraging stakeholder engagement	Report, explain and be answerable to stakeholders in regard to sustainability matters to give them sufficient information to understand the sustainability performance of an organization and to make informed decisions Assuring the credibility and quality of sustainability performance and reporting by means of external assurance

We summarized the goals from the standards and guidelines and categorized them into internal (towards the reporting company's internal stakeholders) and external (towards a company's external stakeholders) goals.

In order to constitute an initial set of quality criteria for SRs, Table III gives an overview of the quality criteria included in the standards and guidelines. References to stakeholders are underlined in order to highlight the stakeholder focus inherent in the quality criteria.

Despite the different backgrounds of the standards and guidelines, the contained quality criteria are quite similar. However, the criteria mostly relate to the contents of SRs – many quality criteria and their definitions are analogous to common data quality dimensions (Wang and Strong, 1996). The multitude of references to stakeholders within the quality criteria stresses the stakeholder focus of sustainability reporting. In view of this focus, we found that the following four important aspects are missing in existing guidelines and standards. We would like to address these aspects with our conceptual framework:

(1) Companies are left in doubt about the specific needs of the different stakeholder groups since the existing standards and guidelines do not differentiate between these groups.

(2) Although ease of use and usefulness are common factors in many established acceptance models (DeLone and McLean, 2003; Venkatesh and Bala, 2008; Venkatesh *et al.*, 2003), existing standards and guidelines do not give concrete advice on how to improve the usability.

(3) Research on web sites indicates that the visual appeal has a significant impact on the level of acceptance (Cyr *et al.*, 2009). As nowadays many SRs are web based (Isenmann *et al.*, 2007; Slater, 2008), the visual appeal of these should equally be acknowledged as a decisive acceptance factor.

(4) The existing standards and guidelines do not disclose detailed information on how the included quality criteria were determined. Since many references to stakeholders are included in the definitions of the quality criteria, further assistance on how to involve the stakeholders into the evaluation of these quality criteria needs to be provided.

Due to the mentioned shortcomings in existing standards and guidelines for sustainability reporting, we have extended our search to standards and guidelines in the field of software usability. Table IV provides a survey of the quality criteria for usability stated in the ISO Standard 9241-110, which deals with the ergonomics of human-system interaction (Deutsches Institut für Normung e.V., 2008). Again, we have underlined references to stakeholders within the definitions of the quality criteria.

Literature in the field of web site quality also confirms that quality criteria exceeding the mere content need to be taken into account. Besides usability, these criteria include availability, reliability, adaptability and response time of the system as well as content that is personalized, complete, relevant, easy to understand and secure (DeLone and McLean, 2003). Other researchers also confirm the importance of usability and add site design and media richness to the list of important influence factors on the success of web sites (Palmer, 2002).

Global Reporting Initiative (GRI, 2011)	
Transparency	Complete disclosure of information on the topics and indicators required to reflect impacts and enable stakeholders to make decisions and the processes, procedures and assumptions used to prepare those disclosures
Balanced and reasonable presentation	Balanced and reasonable presentation of the organization's performance considering the organization's purpose and experience and the reasonable expectations and interests of the organization's stakeholders
Materiality	The information in a report should cover topics and indicators that reflect the organization's significant economic, environmental and social impacts or that would substantively influence the assessments and decisions of stakeholders
Stakeholder inclusiveness	The reporting organization should identify its stakeholders and explain in the report how it has responded to their reasonable expectations and interests
Sustainability context	The report should present the organization's performance in the wider context of sustainability
Completeness	Coverage of the material topics and indicators as well as definition of the report boundary should be sufficient to reflect significant economic, environmental and social impacts and enable stakeholders to assess the reporting organization's performance in the reporting period
Comparability	Issues and information should be selected, compiled and reported consistently. Reported information should be presented in a manner that enables stakeholders to analyze changes in the organization's performance over time and could support analysis relative to other organizations
Balance	The report should reflect positive and negative aspects of the organization's performance to enable a reasoned assessment of overall performance
Accuracy	The reported information should be sufficiently accurate and detailed for stakeholders to assess the reporting organization's performance
Timeliness	Reporting occurs on a regular schedule and information is available in time for stakeholders to make informed decisions
Clarity	Information should be made available in a manner that is understandable and accessible to stakeholders using the report
Reliability	Information and processes used in the preparation of a report should be gathered, recorded, compiled, analyzed and disclosed in a way that could be subject to examination and that establishes the quality and materiality of the information
Report boundary	The sustainability report boundary should include the entities over which the reporting organization exercises control or significant influence both in and through its relationships with various entities upstream (e.g. supply chain) and downstream (e.g. distribution and customers)
International Organization for Standardization (ISO, 2006)	
Transparency	Make the processes, procedures, methods, data sources and assumptions used in environmental communication available to all interested parties, taking account of the confidentiality of information as required. Inform interested parties of the reporting company's role in environmental communication
Appropriateness	Make information provided in environmental communication relevant to interested parties using formats, language and media meeting their interests and needs and enabling them to fully participate
Credibility	Conduct environmental communication in an honest and fair manner and provide information that is truthful, accurate, substantive and not misleading to interested parties. Develop information and data using recognized and reproducible methods and indicators

Table III.
Quality criteria for SRs

(*continued*)

Responsiveness	Ensure that environmental communication is open to the needs of interested parties. Respond to the queries and concerns of interested parties in a fully and timely manner. Make interested parties aware of how their queries and concerns have been addressed
Clarity	Ensure that environmental communication approaches and language are understandable to interested parties to minimize ambiguity
Eco-Management and Audit Scheme (EMAS, 2009)	
Relevance	Indicators should be relevant to an organization's significant direct environmental aspects
Understandability	Indicators shall be understandable and unambiguous
Comparability	Indicators shall allow for a year on year comparison to assess the development of the environmental performance of the organization Indicators shall allow for comparison with sector, national or regional benchmarks as appropriate Indicators shall allow for comparison with regulatory requirements as appropriate
Accuracy	Indicators shall give an accurate appraisal of the organization's environmental performance
Clarity	Environmental information shall be presented in a clear and coherent manner
Availability	Anybody interested in the organization's environmental performance can easily and freely be given access to the information
World Business Council for Sustainable Development (WBCSD, 2003)	
Relevance	Information is relevant when it helps to evaluate a company's activities and to confirm or correct past evaluations
Materiality	A company should assess the materiality, i.e. the importance of the information it discloses. Information is to be considered material if its omission or misstatement could influence users when making decisions about their involvement/relations with the company
Reliability	Information is reliable when it is free from material error and bias and faithfully reflects activities and processes
Comparability	Users should be able to compare the sustainable development reports of a company over time in order to identify trends in its sustainable development performance and position
Cost effectiveness	The benefits derived from producing a report should justify the cost
AccountAbility (AA1000, n.d.)	
Inclusivity	The participation of stakeholders in developing and achieving an accountable and strategic response to sustainability
Materiality	Materiality is determining the relevance and significance of an issue to an organization and its stakeholders. A material issue is an issue that will influence the decisions, actions and performance of an organization or its stakeholders. To make good decisions and actions, an organization and its stakeholders need to know what issues are material to the sustainability performance of the organization
Responsiveness	Responsiveness is an organization's response to stakeholder issues that affect its sustainability performance and is realized through decisions, actions and performance, as well as communication with stakeholders

Table III.

In order to fulfill the collected quality criteria, a company needs to know its stakeholders' needs, expectations and capabilities. Thus, concrete quality criteria can only be determined and evaluated by directly involving the stakeholders in the evaluation process. Our conceptual framework, which is presented in the following section, is designed to enable such an interactive process. Nevertheless, existing

	ISO 9241-110 (Deutsches Institut für Normung e.V., 2008)
Suitability for the task	"An interactive system is suitable for the task when it supports the user in the completion of the task, i.e. when the functionality and the dialogue are based on the task characteristics (rather than the technology chosen to perform the task)"
Self-descriptiveness	"A dialogue is self-descriptive to the extent that, at any time, it is obvious to the users which dialogue they are in, where they are within the dialogue, which actions can be taken and how they can be performed"
Conformity with user expectations	"A dialogue conforms with user expectations if it corresponds to predictable contextual needs of the user and to commonly accepted conventions"
Suitability for learning	"A dialogue is suitable for learning when it supports and guides the user in learning to use the system"
Controllability	"A dialogue is controllable when the user is able to initiate and control the direction and pace of the interaction until the point at which the goal has been met"
Error tolerance	"A dialogue is error-tolerant if, despite evident errors in input, the intended result may be achieved with either no, or minimal, corrective action by the user. Error tolerance is achieved by means of error control (damage control), error correction or error management to cope with the errors that occur"
Suitability for individualization	"A dialogue is capable of individualization when users can modify interaction and presentation of information to suit their individual capabilities and needs"

Table IV.
Quality criteria for user dialogues

standards and guidelines on sustainability reporting or the usability of web sites can provide a useful set of basic quality criteria which can be complemented with the help of the proposed framework.

Graphical illustrations assist understanding and working with data (Larkin and Simon, 1987; Norman, 1993), hence the number of visualization techniques is ever increasing, which gives the authors more options to illustrate complex bodies of information (Ware, 2004). This supports the process of decision making (Bertin, 1981) and aids the discovery of unknown aspects of the data (Cleveland, 1985). As SRs are means to transport data and information, knowledge gained in the field of information visualization aids selecting suitable visualization techniques appropriate to individual tasks (Benbasat and Dexter, 1986) and helps to overcome the eight issues of information visualization, which – of course – also apply to information contained in SRs:

(1) what data is relevant to the task;

(2) how to represent the abstract data;

(3) how to present and layout the data to make the result more memorable and important things stand out;

(4) which scale to select and how to deal with multi-dimensionality;

(5) how to rearrange the data to gain new insights from the data;

(6) how to externalize the data so that it is useful to recipients;

(7) what internal model is created by the visualization in the recipients mind; and

(8) how visualization can best be assisted by tools (Spence, 2001).

4. Conceptual framework

Other researchers have already successfully used multi-method approaches for the evaluation of various IT artifacts using constructs, such as trust or social presence

(Cao *et al.*, 2004; Cyr *et al.*, 2009; Maib *et al.*, 2006; Mingers, 2001). At the time of writing and to the knowledge of the authors, there are no previous attempts to investigate the quality of SRs using multi-method approaches. Furthermore, research in other contexts should not be generalized without additional investigation (Cyr *et al.*, 2009). When evaluating IT artifacts, it should be acknowledged that there are many different ways to do so, each with its assets and drawbacks. Using different methods may lead to different or even contradictory results; therefore, it is appropriate to use a variety of methods and to carefully consolidate the respective findings. Bearing this in mind, we designed a conceptual framework including a multi-method approach for the quality evaluation of SRs in order to overcome the shortcomings of existing standards and guidelines outlined in Section 3. Keeping the identified stakeholder focus in mind, our framework directly involves the stakeholders into the evaluation process. An overview of our proposed framework is shown in Figure 2. To help the reader understand the framework, we explain it by using our evaluation as an illustrative example.

The starting point of the conceptual framework is the iterative development of a catalogue of quality criteria. Various research methods, such as literature reviews, empirical surveys, expert interviews and materiality analyses, may be applied to extract relevant criteria from individuals, guidelines or schemes for awards, such as the ACCA sustainability reporting awards. To focus sustainability reporting on the key issues of the reporting organization as well as its stakeholders, it should be based on a thorough materiality analysis (Sullivan, 2011). For the evaluation of our framework, we used the quality criteria outlined in Section 3.

As a next step, we constitute the technological basis for the implementation of the reports and the evaluation. For the latter, we built a set of alternative reports using a data warehouse implemented in an Oracle DBMS and choosing Cognos 8 as the reporting tool. Subsequently, we designed and implemented alternative reports using different color schemes, menu structures and layouts to assess different alternatives in the following multi-method approach (Cyr *et al.*, 2009). Due to limitations of the applied methods (e.g. eye-tracking, software-supported analysis), presenting an entire SR proved impractical, therefore, we only presented excerpts of a SR. We concentrated on web based SRs, which will be of even greater importance in the future (Isenmann *et al.*, 2007), however, the results of this study can directly be transferred to PDF and printed SRs.

The stakeholder theory (Freeman, 1984) has emerged as the dominant paradigm in CSR (McWilliams and Siegel, 2001). It states that companies have relationships with different groups of stakeholders which affect and are affected by the company (Freeman, 1984). Applied to the field of SRs, the theory implies that on the one hand SRs affect the decisions and actions of different stakeholder groups and on the other hand the reports are affected by the requirements and expectations of the stakeholders (El-Gayar and Fritz, 2006). Hence, representatives of the different stakeholder groups, such as customers, investors, employees and suppliers, should be directly involved in the evaluation process (Oesterle and Otto, 2010).

During the evaluation of the framework, the users have to solve tasks by obtaining certain pieces of information from the SR in varying representations. It has been shown that graphical visualizations can assist task performance and efficiency (Larkin and Simon, 1987; Norman, 1993) and that, in accordance to the cognitive fit (CF) theory (Vessey, 1991), the most helpful visualization should be selected on the basis of various task constraints (e.g. desired precision, available time) (Benbasat and Dexter, 1986).

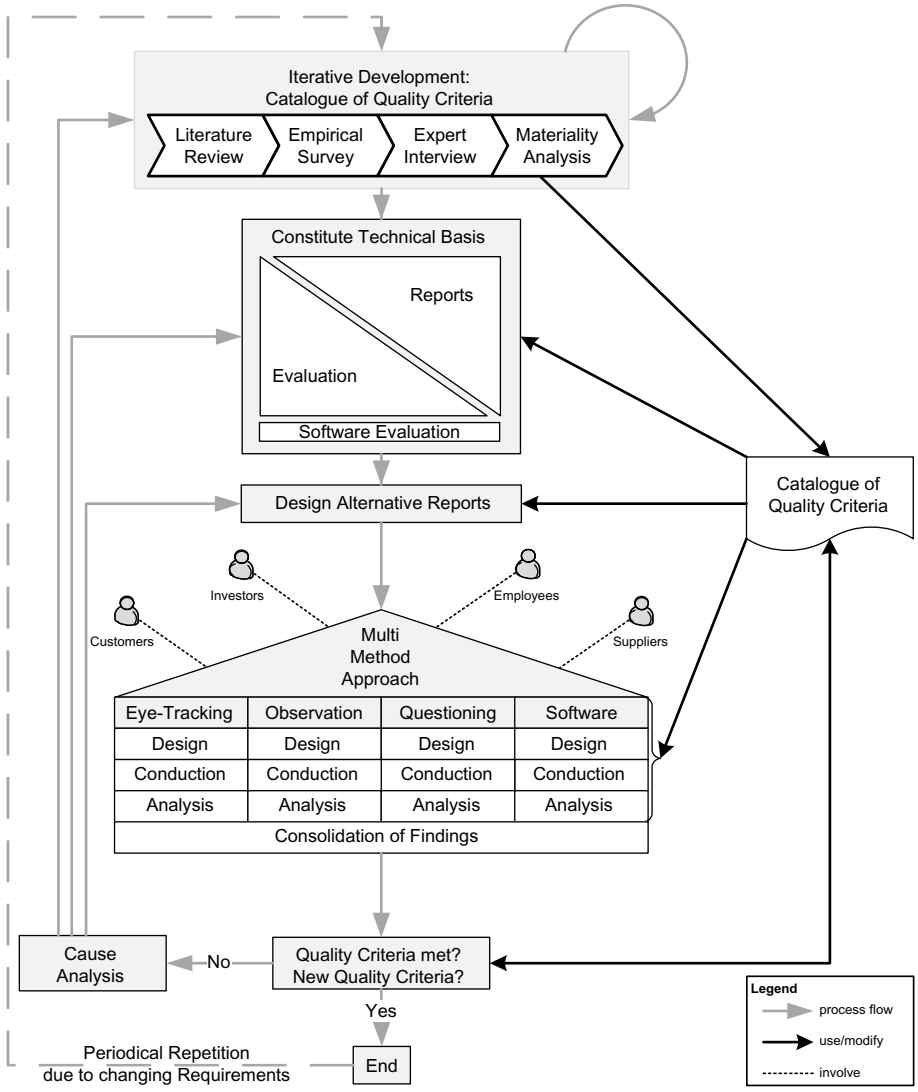


Figure 2.
Conceptual framework for the quality evaluation

Accordingly, the theory of task-technology fit (TTF) also applies to our conceptual framework. TTF asserts IT to be a means for users to complete tasks. The higher the TTF, the better the performance of the users (Goodhue, 1995). By measuring the time needed to complete assigned tasks as well as the number of correctly solved tasks, we evaluate the users' performance while they are working with the reports. Additionally, the expectation confirmation theory (Oliver, 1977, 1980) applies to our framework as the users' expectations, represented by the previously determined quality criteria and the perceived performance of a product (i.e. a SR), lead to (dis-)confirmation of the initial expectations, which has a significant effect on the users' intentions, which may vary by stakeholders.

Research on the acceptance of IT artifacts is frequently based on questionnaires measuring acceptance factors (Venkatesh and Bala, 2008; Venkatesh *et al.*, 2003). When implementing the questionnaire, we could draw from previous experiences and experiments with designing surveys for evaluating the quality and acceptance of SRs (references removed due to blind review), which ultimately led to the research results presented in this paper. During said previous research, we learned valuable lessons with respect to designing questionnaires evaluating environmental matters and experimental setups with sample sizes varying from 30 to 260. This significantly influenced the creation of the conceptual evaluation framework presented in this paper. From experience, we regard questionnaires to be a good method to capture the users' (self-)perception, but more objective methods should be added to achieve a complete picture of the acceptance factors (Venkatesh and Bala, 2008). In order to allow a suitable evaluation of the reports, we apply different research and analyses methods (Cao *et al.*, 2004; Cyr *et al.*, 2009; Maib *et al.*, 2006; Mingers, 2001):

- video recordings from various perspectives (screen as well as face and hands of the test users);
- eye-tracking; and
- software-based attention analysis.

In the evaluation of our conceptual framework, we used an eye-tracker to be able to monitor the users' attention behavior. The test users were asked to retrieve information on a company's environmental performance from the reports and, in addition to the eye-tracking, we video recorded them from various perspectives while working. The cameras captured the screen, the hands and the faces of the test users. This allowed identifying perceived shortcomings of the report alternatives by closely observing the users' behavior (e.g. frowning, sighing, scratching of the head). Afterwards, some of the users were confronted with the recordings and their corresponding behavior in order to further investigate the causes.

In order to back up the eye-tracking results, we performed a software-supported attention analysis using the software EyeQuant (WhiteMatter Labs, 2009), which emulates the attention behavior of users by means of a neural network. Originally designed and trained for the evaluation of web sites, we explored to what extent the software could also aid in evaluating web based SRs. Apart from these innovative research methods, we also used a questionnaire for the evaluation of the reports.

After completion of the design, conduction and analysis of the chosen research methods, we consolidate the findings. In the course of the consolidation process, the results provided by the different research methods are checked for consistency and contradictions are detected and interpreted. Subsequently, the catalogue of quality criteria is reviewed in the light of the research results: Which report alternatives succeeded in meeting the requirements and which failed? Have additional requirements been raised by the involved stakeholders or did some of the requirements turn out to be obsolete?

In case the quality criteria are not sufficiently satisfied, a cause analysis is conducted. This may result in design changes, modifications of the reports' content or the selection of a more suitable technological basis. Alternatively, the catalogue of quality criteria itself may be modified.

When the quality criteria are finally satisfied, the process should be repeated after some time to account for changes in stakeholder requirements and to enable a continuous improvement process.

As mentioned before, there are many ways to evaluate usability and performance of an IT artifact. By adding or disregarding one or more of the proposed research methods, the conceptual framework can easily be modified. In addition to this, it may also be used for the evaluation of other IT artifacts like web sites or software user interfaces, which constitutes a significant strength of our framework. Additionally, the proposed conceptual framework exhibits a high degree of flexibility regarding evaluation methodologies, the degree of stakeholder involvement and number of iterations as well as the number of people involved.

5. Evaluation results and discussion

In order to evaluate our proposed conceptual framework and to gain first insights into the stakeholder-focused quality evaluation, we used the quality criteria collected in Section 3 to design and implement a number of visualizations one may encounter when reading SRs. We later on applied our multi-method framework to the different visualizations. We conducted our initial test with a group of eight users (graduate students and PhD candidates with backgrounds in IS and work experience ranging from zero to five years) who acted as members of different stakeholder groups. We intentionally violated some of the quality criteria, e.g. clarity or suitability for the task, in order to check whether our framework was suitable to discover these violations. The results can be subdivided into findings on the framework itself, findings regarding the design and implementation of SRs and implications for IS theories and future research.

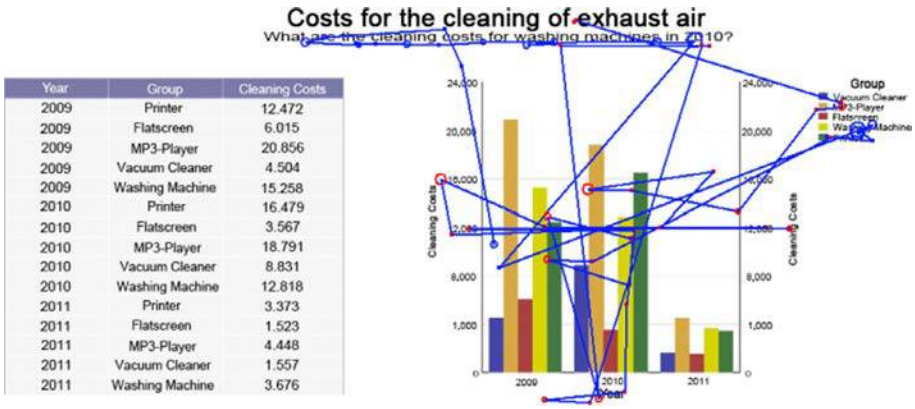
5.1 Findings on the conceptual framework

In this section, we focus on the eye-tracking and software-supported evaluation results since most researchers and practitioners are probably familiar with questionnaires and observations. Additionally, we point out the main benefits and problems we encountered using a multi-method approach.

Although conducting eye-tracking is both time-consuming and expensive, we found that it was well worth the effort. We thereby gained valuable information on what parts of a report the test users actually used to solve certain tasks. Figure 3 shows the eye-tracking results of one test user for one report alternative. The lines indicate saccades (movements of the eyes), whereas the circles indicate fixations (areas the test user focused on for a longer period of time) (Salvucci and Goldberg, 2000).

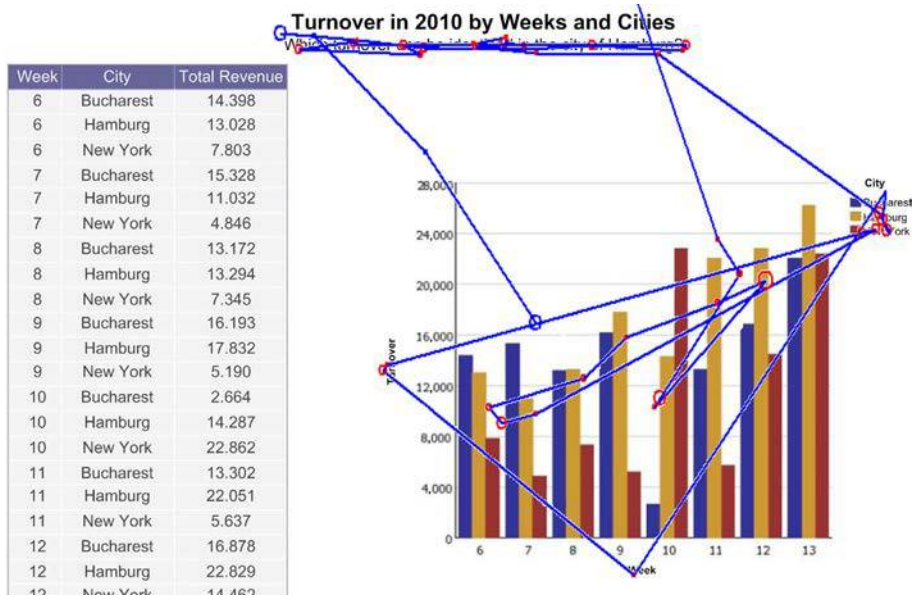
The test user was given the task to state the costs for the cleaning of exhaust air for the product group washing machines in the year 2010. The eye-tracking results reveal that this user selected the bar chart to find the required solution instead of the table in which the exact number is already listed. The test user subsequently failed to complete the task successfully. Obviously, the bar chart violated the quality criterion suitability for the task which was successfully discovered by our conceptual framework.

Figure 4 shows another example of the eye-tracking results. This test user was given the task to identify a revenue trend in a given period. In this case, using the bar chart was appropriate and the test user performed well. When comparing both eye-tracking results, it is noticeable that the test user in Figure 3 performed many unnecessary eye movements since he did not choose the visualization alternative that



Year	Group	Cleaning Costs
2009	Printer	12.472
2009	Flatscreen	6.015
2009	MP3-Player	20.856
2009	Vacuum Cleaner	4.504
2009	Washing Machine	15.258
2010	Printer	16.479
2010	Flatscreen	3.567
2010	MP3-Player	18.791
2010	Vacuum Cleaner	8.831
2010	Washing Machine	12.818
2011	Printer	3.373
2011	Flatscreen	1.523
2011	MP3-Player	4.448
2011	Vacuum Cleaner	1.557
2011	Washing Machine	3.676

Figure 3.
Eye-tracking results exhaust air



Week	City	Total Revenue
6	Bucharest	14.398
6	Hamburg	13.028
6	New York	7.803
7	Bucharest	15.328
7	Hamburg	11.032
7	New York	4.846
8	Bucharest	13.172
8	Hamburg	13.294
8	New York	7.345
9	Bucharest	16.193
9	Hamburg	17.832
9	New York	5.190
10	Bucharest	2.664
10	Hamburg	14.287
10	New York	22.862
11	Bucharest	13.302
11	Hamburg	22.051
11	New York	5.637
12	Bucharest	16.878
12	Hamburg	22.829

Figure 4.
Eye-tracking results turnover

was suitable for the task. Accordingly, the eye-tracking data may also be used to estimate the effort that a test user had to spend to solve a task or in other words, the suitability for the task or CF can be investigated.

As some of our test users were interested in the eye-tracking technology, we presented the visualized data to them and discovered that while viewing the results the test users were able to remember small details and decisions they made. As this information has not been included in the results of our other research methods, we believe that a retrospective confrontation of the test user with their eye-tracking data can lead to further interesting insights. In contrast to other methods to investigate the users' cognitive processes, such as thinking aloud (Van Someren *et al.*, 1994),

this method can be applied retrospectively and thus without distracting the test user during the completion of a task.

The eye-tracking results also revealed which test users felt stressed during the experiment conduction: users who stated that they felt pressured to solve the given task successfully or even within a limited period of time showed more rapid and unstructured eye movements in their eye-tracking data.

Overall, we found eye-tracking to be an interesting method to evaluate SRs. We believe that it may also be useful for other IT artifact evaluations.

In addition to eye-tracking, we conducted a software-supported analysis using the software EyeQuant which is based on a neural network that emulates human attention behavior (WhiteMatter Labs, 2009). It allows generating heat maps and other visual analyses of the presumed user's attention. Figure 5 shows a heatmap that was generated by applying the software to one of our report designs.

Our experiences with EyeQuant were mixed since the software was originally designed to evaluate web sites and not SRs: compared to eye-tracking, the analysis was a lot easier, faster and cheaper to conduct. The software produced quite good results for reports that contained many different elements such as tables, charts, legends, etc. Figure 5 shows an example in which the software performed rather well.

However, for reports that included one large table only, the software was not able to produce comprehensive results. Figure 6 shows an example of such a report.

The mixed experiences with this software-based attention analysis can be explained by the fact that, in contrast to real test users, the software is not able to focus on a certain task. Generally, the software-supported analysis could not replace, but complement the eye-tracking data. For example, Figure 5 shows that the image catches

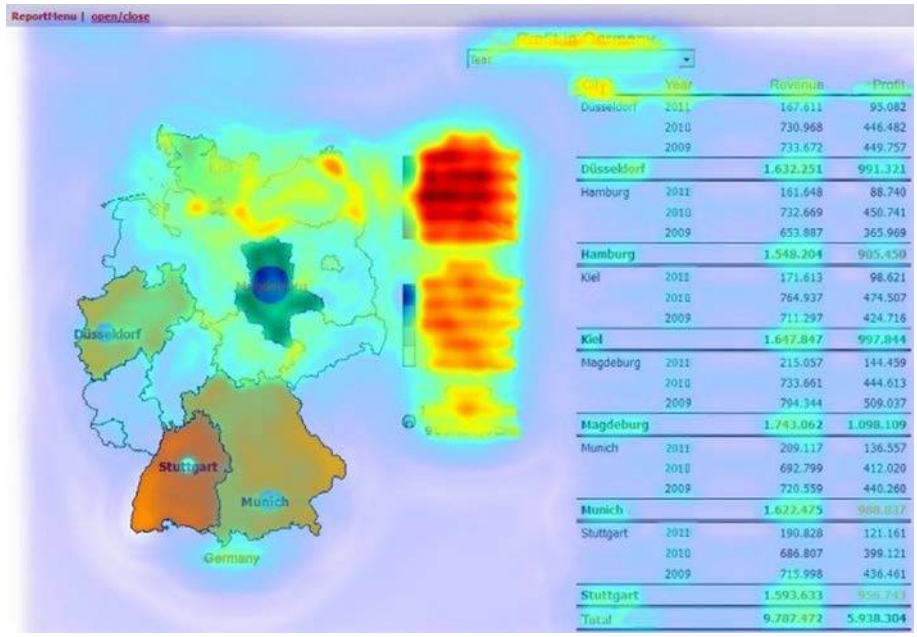
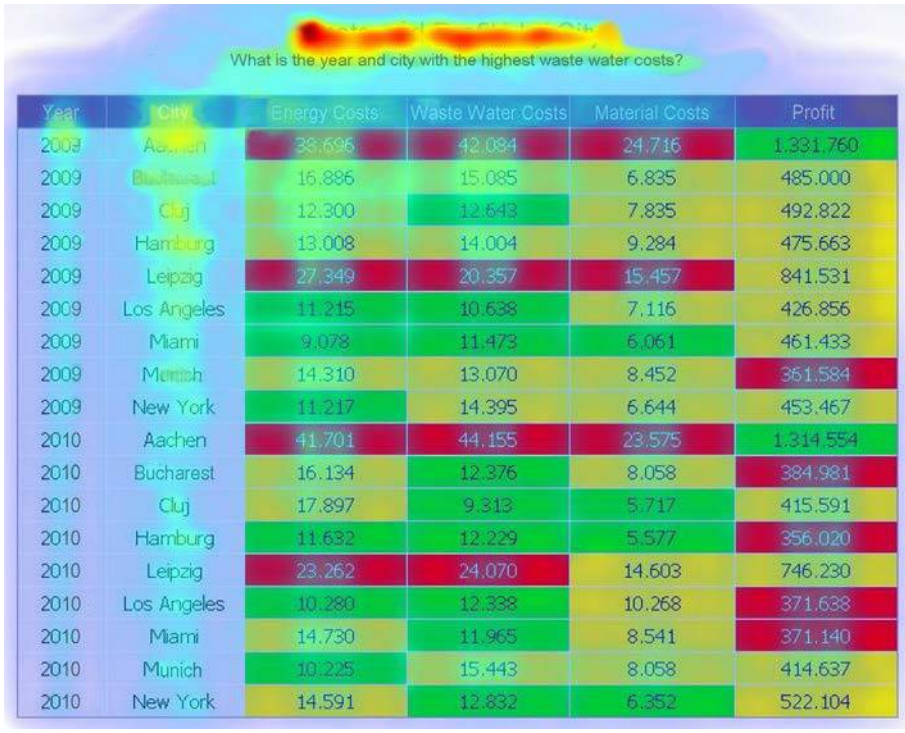


Figure 5.
EyeQuant – different elements



Year	City	Energy Costs	Waste Water Costs	Material Costs	Profit
2009	Aachen	38.696	42.084	24.716	1.331.760
2009	Bucharest	16.886	15.085	6.835	485.000
2009	Cluj	12.300	12.643	7.835	492.822
2009	Hamburg	13.008	14.004	9.284	475.663
2009	Leipzig	27.349	20.357	15.457	841.531
2009	Los Angeles	11.215	10.638	7.116	426.856
2009	Miami	9.078	11.473	6.061	461.433
2009	Munich	14.310	13.070	8.452	361.584
2009	New York	11.217	14.395	6.644	453.467
2010	Aachen	41.701	44.155	23.575	1.314.554
2010	Bucharest	16.134	12.376	8.058	384.981
2010	Cluj	17.897	9.313	5.717	415.591
2010	Hamburg	11.632	12.229	5.577	356.020
2010	Leipzig	23.262	24.070	14.603	746.230
2010	Los Angeles	10.280	12.338	10.268	371.638
2010	Miami	14.730	11.965	8.541	371.140
2010	Munich	10.225	15.443	8.058	414.637
2010	New York	14.591	12.832	6.352	522.104

Figure 6.
EyeQuant – large table

more attention than the table, which explains why in Figure 3 the user preferred the bar chart instead of the more suitable table. Eventually, training the neural network of the software with eye-tracking data from test users reading a SR could further improve the results. However, certain quality criteria, such as suitability for the task, cannot be evaluated using a software-supported analysis. This indicates that depending on the set of identified quality criteria, different methods included in the multi-method approach perform differently. Naturally, a higher number of research methods cover a wider range of quality criteria, but then also the effort to conduct the evaluation and analyze the results increases. Accordingly, there is a trade-off between the comprehensiveness of results and the effort that is required to perform the evaluation.

Overall, the combination of different research and analyses methods proved to be a good approach, which corresponds to findings from existing literature (Cao *et al.*, 2004; Cyr *et al.*, 2009; Maib *et al.*, 2006; Mingers, 2001). The results of each research method provided valuable insights into the stakeholders' view of quality criteria for web based SRs:

- The eye-tracking data allowed us to evaluate which components of a report were used to complete a certain task and which were not. The retrospective eye-tracking provided further valuable insights into the users' cognitive processes.
- The questionnaire allowed us to evaluate the stakeholders' perception of the reports.

- The observation allowed us to see whether the stakeholders' perception was biased.
- The software-supported analysis extended our findings from the eye-tracking in a more time- and cost-effective way.

Naturally, using different research and analyses methods may also lead to contradicting results: for example, some test users perceived a report alternative to be useful and easy to use, although they were not able to solve a given task using the report. As pointed out in Sections 5.3 and 6, these contradictions provide good starting points for future research.

5.2 Findings on the design and implementation of web based SRs

Since the stakeholders did not always choose the part of the report that was suitable for the given task (Figure 3), we consider coaching of the users beneficial and recommend this when designing and implementing SRs. While coaching of internal stakeholders such as employees can easily be provided, coaching of external stakeholders such as customers is certainly not easy to implement. For web based SRs, coaching of external stakeholders can be accomplished by adding a help function, a wizard or an instructional video or avatar demonstrating how to switch between different types of visualization and which kind of visualization is suitable for the corresponding task. In accordance with the quality criteria suitability for the task, the appropriate visualization should be pre-selected where possible.

In view of the fact that we discovered contradictions between the perceived and actual usefulness and ease of use (Section 5.1), we argue that SRs need to address both aspects the users' perception as well as the factual properties since both may contribute to the overall acceptance of a report.

Table V shows an excerpt from our questionnaire and observation results. The eight test users were given two tasks:

(1) For the results on the left hand side of the table, the test users were asked to state an exact number for the cleaning costs of exhaust air for the product group washing machines (Figure 3).

(2) For the results on the right hand side of the table, the test users were asked to identify a revenue trend of a certain company location within a given period (Figure 4).

The two tasks were presented to the test users in random order to mitigate bias or learning effects. After the completion of the tasks, the test users were asked to fill out a questionnaire using a Likert scale ranging from 1 (strong disagree) to 5 (strong agree). The test users were asked whether the table was helpful for the task, whether the chart was helpful for the task and whether the visual appeal of the chart was greater than that of the table.

We extended the questionnaire answers of the test users in Table V by noting whether the chart was perceived to be more helpful than the table. Our observation revealed whether the given task was solved correctly by the test user.

We expected the table to be more helpful for the first task, whereas we expected the chart to be more helpful for the second task. Generally, our expectations have been confirmed by our test users, as the higher mean values of the table for task 1 and of the

Table V. Questionnaire and observation results

Task	Stating a concrete number (cf. report Figure 2)					Identifying a trend (report similar to Figure 2)				
Participant	Table helpful for task	Chart helpful for task	Helpful chart > table?	Task solved?	Visual appeal chart > table?	Tablehelpful for task	Chart helpful for task	Helpful chart > table?	Task solved?	Visual appeal chart > table?
A	3	4	No	No	5	2	5	No	Yes	5
B	5	3	Yes	Yes	3	2	5	No	Yes	3
C	1	4	No	Yes	4	3	2	Yes	No	1
D	5	3	Yes	Yes	4	1	5	No	Yes	4
E	5	1	Yes	Yes	4	1	5	No	Yes	4
F	3	5	No	No	4	3	5	No	Yes	4
G	4	2	Yes	No	1	2	3	No	Yes	4
H	5	1	Yes	Yes	1	2	5	No	Yes	5
Mean	3.875	2.875			3.25	2	4.375			3.75

chart for task 2 indicate, however, the results were not as clear as expected: three out of eight test users found the chart to be more helpful for the first task and one test user was of the opinion that the table was more helpful to solve task 2. Accordingly, out of these four test users highlighted in bold print in Table V, only one was able to correctly solve the given task.

This underlines our recommendations to:

- provide coaching to internal stakeholders;
- offer a help function, wizard, instructional video or avatar to external stakeholders; and
- pre-select the visualization that is suitable for the task where possible.

Surprisingly, the results for the second task were clearer than for the first task even though the report designs were rather similar.

Although the chart was by and large considered to be less helpful for the first task, in both cases the test users considered the chart to be visually more appealing. One could conclude that SRs should therefore generally include charts rather than tables, but as our results also indicate, charts may not always be suitable for the task at hand, confirming previous research (Benbasat and Dexter, 1986).

Figure 7 shows two alternative reports that were designed using different color schemes.

Although the two alternatives use the same layout and include the same information, the stakeholders' experiences differ strongly. While the report on the left uses established traffic light colors (green for low costs/high profits, yellow for medium costs and profits, red for high costs/low profits), the report on the right uses colors with no common connotation (Ware, 2004). Figure 7 shows a boxplot of the time it took the test users to solve the tasks.

It is not surprising, that although the tasks were equally difficult the test users on average needed 72 percent more time to complete the task using the report on the right. The eye-tracking results confirm this finding by showing disoriented, confused test users. Surprisingly, the questionnaire answers found the two reports to be almost equally visually appealing (2.9 for the report on the left compared to 2.75 for the report on the right). The results were more significant regarding the helpfulness for the task (3.3 vs 2.63) and the intuitiveness of the color scheme (3.5 vs 2.5). Some test users

Figure 7.
Color schemes

remarked that the colors in the report on the left hand side were too glaring which might be the reason why the results were not as significant as we expected.

Regarding the color scheme, we conclude the following recommendations:

- color schemes should be intuitive;
- a legend of the color scheme should be included in order to provide orientation; and
- colors should not be too glaring and friendly to the eye.

5.3 Implications

From a theoretical perspective, our findings partly support the stakeholder theory: we identified numerous references to stakeholders within the quality criteria reviewed in Section 3. The sustainability reporting of companies is clearly influenced by the stakeholders; therefore, our conceptual framework directly involves stakeholders into the quality evaluation process and, provided a materiality analysis is performed beforehand, into the process of collecting quality criteria. Since investigating the vice versa relationship (influence of the report on the stakeholders) was not within the focus of our research, we can only partly provide evidence for the applicability of the stakeholder theory.

It is not surprising that different report alternatives had different levels of CF. This was reflected in our results by showing different ratios of correctly solved tasks depending on which report alternative the test users chose (Table V) and how much time it took the users to solve the tasks (Figure 8). This shows that the body of knowledge accumulated by the information and data visualization community cannot be neglected. Surprisingly, when alternatives exhibiting different levels of CF were presented to the test users, the users did not always choose the most suitable alternative. Thus, we suggest to offer coaching to the users and to determine which type of graphical visualization is most task-efficient in order to offer a sensible pre-selection for the visualization.

For the authors of standards and guidelines on sustainability reporting, we conclude that the quality criteria should be extended by common acceptance factors from the IS disciple (DeLone and McLean, 2003; Venkatesh and Bala, 2008; Venkatesh *et al.*, 2003). A generally accepted standard on what environmental information has to be reported in which way, should be established, making it easier for the readers to understand, interpret and compare the presented data. Since many companies publish their SRs in form of a web site (Slater, 2008), it is no longer sufficient to focus on quality criteria regarding the content of static PDF documents (Isenmann *et al.*, 2007, 2011). Also, the guidelines could offer more advice on how to evaluate the stated quality criteria and also information on how the criteria were constituted in the first place.

Researchers have already started to explore the additional possibilities that web based SRs offer in comparison to traditional, paper based reports. Web based reports can, for example, facilitate two-way communication between the reporting company and its stakeholders and provide the possibility to generate custom-tailored reports for different stakeholder groups (Isenmann *et al.*, 2007, 2011) or augment the SR with

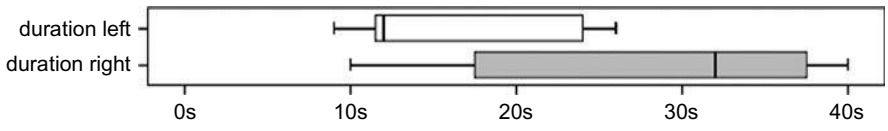


Figure 8.
Color schemes boxplots

additional information relevant to the contents of the report (Freundlieb and Teuteberg, 2012). The application of the presented framework might help to derive additional quality criteria and further evaluate these approaches.

Our research also yields implications for practitioners. As we have shown, in the area of sustainability reporting there is an array of standards and guidelines to be reviewed before the actual work on a SR can begin. As with financial and other reports, project teams in charge of creating a SR must not exclusively consist of environmental experts, also experts from other areas, such as data visualization and web design, should be involved. The proposed conceptual framework can serve as a guide through the process of quality evaluation and improvement of any SR. The framework provides flexibility as to which and how many stakeholders to consult, which and how many evaluation methods to apply and how many iterations to perform. It should be noted that the proposed framework possesses a high degree of scalability and is therefore suitable for a wide range of organizations and project sizes. Hence, it may either be used to improve a SR with only a few iterations within a short period of time or to establish a process of continual improvement.

We criticized that established guidelines do not give concrete advice on how to improve the usability of SRs. The proposed conceptual framework may serve as a basis for a standard procedure for the quality assessment of SRs that may be included in future revisions of said guidelines.

6. Conclusion

This paper contributes to the research on web based SRs by providing an overview of the goals and quality criteria in existing reporting standards and guidelines. Our overview shows that, with the exception of AA1000, existing standards and guidelines are too focused on the content of the reports and, in view of the stakeholder-focused nature of SRs, neglect common IS acceptance criteria which have proven to be valid in many other problem domains. In response to these shortcomings, we present a conceptual multi-method framework that directly involves the different stakeholder groups into the definition of quality criteria and the corresponding evaluation. This paper presents findings on the conceptual framework and the included multi-method approach itself, findings on the design and implementation of SRs as well as theoretical implications. Accordingly, both practitioners and scientists can benefit from the results.

Although we are satisfied with the selected combination of research methods and the conducted evaluation, further research on the combination of different research methods is needed: Which research methods complement each other? How can contradictions between results from different research methods be explained and resolved during the consolidation of findings?

By replacing the research and analyses methods included in the multi-method approach, we believe that our conceptual framework can easily be adapted to suit other problem domains, such as web or application design. We experienced that some research methods, e.g. eye-tracking, are more time-consuming and costly than others, e.g. the software-supported analysis. While we only considered a rather small group of test users, an economical evaluation of different combinations of research and analyses methods could provide interesting insights for researchers and practitioners alike: Which combination of research methods offers the best price performance ratio in a multi-method approach?

There are many scientific contributions to the acceptance of technologies relying on questionnaires to determine acceptance factors, such as ease of use and usefulness. However, the discrepancies we discovered show that there is room for further research. Scientists should further investigate and distinguish between perceived (based on questionnaire) and actual ease of use (observed, measured, eye-tracked) as well as perceived and actual usefulness. Some researchers have already made similar suggestions and started to distinguish between perceived and objective acceptance criteria (Venkatesh and Bala, 2008). Nevertheless, we believe that there is a need for further, continued research on the cause for the differences between perceived and actual acceptance criteria and their impact on the overall acceptance.

Our evaluation results indicate that the visual appeal of a report may have a strong influence on the user's perceived ease of use and usefulness. An empirical investigation of this hypothesis could be another starting point for future research.

Naturally, our research approach suffers from some limitations that may also serve as starting points for future research: the evaluation of our conceptual framework was conducted in an artificial environment (laboratory experiment) possibly biasing the results. However, we believe that the applied multi-method approach reduces this problem to a large extent. Nevertheless, future research can focus on more natural use settings.

Furthermore, our evaluation only involved a small number of rather homogenous test users and we only conducted one iteration of the framework. Ideally, a larger number of users and a more representative user population would improve the findings. Thus, our proposed conceptual framework needs to be further evaluated in the future and constitutes a theoretical construct that still needs to prove its value in practice. Due to the interchangeability of the included research and analyses methods, it can, however, easily be modified by researches and practitioners alike to meet their respective needs. For further evaluation, we invite other researchers to apply and to adopt our conceptual framework to other problem domains (e.g. the design of web sites or software user interfaces) and to incorporate more evaluation methods, such as the methodologies recently used in the field of NeuroIS (Dimoka *et al.*, 2010).

References

Amran, A. and Haniffa, R. (2011), "Evidence in development of sustainability reporting: a case of a developing country", *Business Strategy and the Environment*, Vol. 20 No. 3, pp. 141-156.

Association for Information Systems (2009), "MIS journal rankings", available at: http://ais.affiniscape.com/displaycommon.cfm?an=1&subarticlenbr=432 (accessed March 5, 2012).

Bartels, W., Iansen-Rogers, J. and Kuszewski, J. (2008), *Count Me In – The Readers' Take on Sustainability Reporting*, KPMG, Amstelveen.

Benbasat, I. and Dexter, A.S. (1986), "An investigation of the effectiveness of color and graphical information presentation under varying time constraints", *MIS Quarterly*, Vol. 10 No. 1, pp. 59-83.

Bertin, J. (1981), *Graphics and Graphic Information*, Walter de Gruyter, Berlin.

Brundtland Commission (1987), *Our Common Future*, Oxford University Press, Oxford.

Cao, J., Lin, M., Crews, J., Deokar, A. and Burgoon, J. (2004), "The interaction of research methods for system evaluation and theory testing: a new vision of the benefits of multi-methodological information systems research", *Proceedings of the International Conference on Information Systems* (*ICIS*), *Washington, DC*.

Clacher, I. and Hagendorff, J. (2012), "Do announcements about corporate social responsibility create or destroy shareholder wealth? Evidence from the UK", *Journal of Business Ethics*, Vol. 106 No. 3, pp. 253-266.

Cleveland, W.S. (1985), *The Elements of Graphing Data*, Wadsworth Advanced Books and Software, Monterey, CA.

CORE (2008), "ERA conference rating", available at: http://core.edu.au/index.php/categories/conference%20rankings/1 (accessed April 2, 2012).

Cyr, D., Head, M., Larios, H. and Pan, B. (2009), "Exploring human images in web site design: a multi-method approach", *MIS Quarterly*, Vol. 33 No. 3, pp. 1-32.

Daub, C.-H. (2007), "Assessing the quality of sustainability reporting: an alternative methodological approach", *Journal of Cleaner Production*, Vol. 15 No. 1, pp. 75-85.

Davis, A.M. (1992), "Operational prototyping: a new development approach", *IEEE Software*, Vol. 9 No. 5, pp. 70-78.

DeLone, W.H. and McLean, E.R. (2003), "The DeLone and McLean model of information systems success: a ten-year update", *Journal of Management Information Systems*, Vol. 19 No. 4, pp. 9-30.

Deutsches Institut für Normung e.V. (2008), *Ergonomics of Human-System Interaction – Part 110: Dialogue Principles* (*ISO 9241-110: 2006*) *English Version of DIN EN ISO 9241-110: 2008-09, Displays*, Deutsches Institut für Normung e.V., Berlin.

Dimoka, A., Pavlou, P.A. and Davis, F.D. (2010), "NeuroIS: the potential of cognitive neuroscience for information systems research", *Information Systems Research*, pp. 1-18 (articles in advance).

El-Gayar, O. and Fritz, B.D. (2006), "Environmental management information systems (EMIS) for sustainable development: a conceptual overview", *Communications of the Association for Information Systems*, Vol. 17 No. 1, pp. 756-784.

Elkington, J. (1998), "Partnerships from cannibals with forks – the triple bottom line of 21st-century business", *Environmental Quality Management*, Vol. 8 No. 1, pp. 37-51.

Freeman, R.E. (1984), *Strategic Management: A Stakeholder Approach*, Pitman, Boston, MA.

Freundlieb, M. and Teuteberg, F. (2012), "Augmented sustainability reports – a design science approach", *AIS Transactions on Human-Computer Interaction*, Vol. 4 No. 4.

Goodhue, D.L. (1995), "Understanding user evaluations of information systems", *Management Science*, Vol. 41 No. 12, pp. 1827-1844.

Heemskerk, B., Pistorio, P. and Scicluna, M. (2002), *Sustainable Development Reporting – Striking the Balance, 2002*, Earthprint, Stevenage.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004), "Design science in information systems research", *MIS Quarterly*, Vol. 28 No. 1, pp. 75-105.

Isenmann, R., Bey, C. and Welter, M. (2007), "Online reporting for sustainability issues", *Business Strategy and the Environment*, Vol. 16, pp. 487-501.

Isenmann, R., Gómez, J.M. and Süpke, D. (2011), "Making stakeholder dialogue for sustainability issues happen – benefits, reference architecture and pilot implementation for automated sustainability reporting à la Carte", *Proceedings of the 44th Hawaii International Conference on System Sciences*, IEEE Computer Society, Kauai, HI.

Kolk, A. (2004), "A decade of sustainability reporting: developments and significance", *International Journal of Environment and Sustainable Development*, Vol. 3 No. 1, pp. 51-64.

KPMG (2011), *KPMG International Survey of Corporate Responsibility Reporting 2011*, KPMG, Zürich.

Larkin, J.H. and Simon, H.A. (1987), "Why a diagram is (sometimes) worth 10,000 words", *Cognitive Science*, Vol. 11 No. 1, pp. 65-100.

McWilliams, A. and Siegel, D. (2001), "Corporate social responsibility: a theory of the firm perspective", *The Academy of Management Review*, Vol. 26 No. 1, pp. 117-127.

Maib, J., Hall, R., Collier, H. and Thomas, M. (2006), "A multi-method evaluation of the implementation of a student response system", *Proceedings of the Americas Conference on Information Systems* (*AMCIS*), *Acapulco*.

March, S.T. and Smith, G.F. (1995), "Design and natural science research on information technology", *Decision Support Systems*, Vol. 15 No. 4, pp. 251-266.

Melville, N.P. (2010), "Information systems innovation for environmental sustainability", *MIS Quarterly*, Vol. 34 No. 1, pp. 1-21.

Mingers, J. (2001), "Combining IS research methods: towards a pluralist methodology", *Information Systems Research*, Vol. 12 No. 3, pp. 240-259.

Morhardt, J.E. (2010), "Corporate social responsibility and sustainability reporting on the internet", *Business Strategy and the Environment*, Vol. 19 No. 7, pp. 436-452.

Norman, D.A. (1993), *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*, Addison-Wesley, Boston, MA.

Oesterle, H. and Otto, B. (2010), "Consortium research", *Business & Information Systems Engineering*, Vol. 2 No. 5, pp. 283-293.

Oliver, R. (1977), "Effect of expectation and disconfirmation on postexposure product evaluations: an alternative interpretation", *Journal of Applied Psychology*, Vol. 62 No. 4, pp. 480-486.

Oliver, R. (1980), "A cognitive model of the antecedents and consequences of satisfaction decisions", *Journal of Marketing Research*, Vol. 1 No. 11, pp. 460-469.

Palmer, J.W. (2002), "Web site usability, design, and performance metrics", *Information Systems Research*, Vol. 13 No. 2, pp. 151-167.

Salvucci, D.D. and Goldberg, J.H. (2000), "Identifying fixations and saccades in eye-tracking protocols", *Proceedings of the 2000 Symposium on Eye-Tracking Research and Applications, Palm Beach Gardens, FL*, pp. 71-78.

Slater, A. (2008), *International Survey of Corporate Responsibility Reporting 2008, Corporate Governance*, KPMG, Amstelveen.

Spence, R. (2001), *Information Visualization*, Pearson Education, Harlow.

Stubbs, W., Higgins, C. and Milne, M. (2012), "Why do companies not produce sustainability reports", *Business Strategy and the Environment*, November.

Sullivan, R. (2011), *Valuing Corporate Responsibility: How Do Investors Really Use Corporate Responsibility Information?*, Greenleaf Publishing, Sheffield.

Townsend, S., Bartels, W. and Renaut, J.-P. (2010), *Reporting Change: Readers & Reporters Survey 2010, Change*, Futerra Sustainability Communications, London.

van Marrewijk, M. (2002), "Concepts and definitions of CSR and corporate sustainability: between agency and communion", *Journal of Business Ethics*, Vol. 44 Nos 2/3, pp. 95-105.

Van Someren, M.W., Barnard, Y.F. and Sandberg, J.A.C. (1994), *The Think Aloud Method – A Practical Guide to Modelling Cognitive Processes, Social Science*, Academic Press, London.

Venkatesh, V. and Bala, H. (2008), "Technology acceptance model 3 and a research agenda on interventions", *Decision Sciences*, Vol. 39 No. 2, pp. 273-315.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003), "User acceptance of information technology – toward a unified view", *MIS Quarterly*, Vol. 27 No. 3, pp. 425-478.

Vessey, I. (1991), "Cognitive fit: a theory-based analysis of the graphs versus tables literature", *Decision Sciences*, Vol. 22 No. 2, pp. 219-240.

Wang, R. and Strong, D.M. (1996), "Beyond accuracy: what data quality means to data consumers", *Journal of Management Information Systems*, Vol. 12 No. 4, pp. 5-34.

Ware, C. (2004), *Information Visualization – Perception for Design*, 2nd ed., Elsevier, San Francisco, CA.

Watson, R.T., Boudreau, M.-C. and Chen, A.J. (2010), "Information systems and environmentally sustainable development: energy informatics and new directions for the IS community", *MIS Quarterly*, Vol. 34 No. 1, pp. 23-38.

Webster, J. and Watson, R.T. (2002), "Analyzing the past to prepare for the future: writing a literature review", *MIS Quarterly*, Vol. 26 No. 2, pp. 13-23.

Welford, R. and Frost, S. (2006), "Corporate social responsibility in Asian supply chains", *Corporate Social Responsibility and Environmental Management*, Vol. 13 No. 3, pp. 166-176.

Wheeler, D. and Sillanpää, M. (1998), "Including the stakeholders: the business case", *Long Range Planning*, Vol. 31 No. 2, pp. 201-210.

WhiteMatter Labs (2009), "EyeQuant attention analytics", available at: www.eyequant.com (accessed September 10, 2012).

About the authors

Michael Freundlieb and Matthias Gräuler are Research Assistants in the Research Group in accounting and information systems at the University of Osnabrueck who are predominantly focused on the fields of sustainability reporting and Green IS. Matthias Gräuler is the corresponding author and can be contacted at: matthias.graeuler@uni-osnabrueck.de

Frank Teuteberg, Head of the Research Group in accounting and information systems at the University of Osnabrück, is an Information Systems Researcher with an interest in Green IS, supply chain management and cloud computing.

This article has been cited by:

1. Anushree Poddar, Sapna A. Narula. 2018. Sustainability reporting practices in India: A study of selected conglomerates. *Strategic Change* **27**:6, 543-557. [Crossref]
2. Kieran Usher, Warren Maroun. 2018. A review of biodiversity reporting by the South African seafood industry. *South African Journal of Economic and management Sciences* **21**:1. . [Crossref]
3. MarianoStefania, Stefania Mariano. 2018. Initiating, challenging and improving knowledge in organizational contexts. *Management Research Review* **41**:3, 278-295. [Abstract] [Full Text] [PDF]
4. HuberSandra, Sandra Huber, BassenAlexander, Alexander Bassen. 2018. Towards a sustainability reporting guideline in higher education. *International Journal of Sustainability in Higher Education* **19**:2, 218-232. [Abstract] [Full Text] [PDF]
5. Milena Peršić, Halmi Lahorka. 2018. Exploring the quality of social information disclosed in non-financial reports of Croatian companies. *Economic Research-Ekonomska Istraživanja* **31**:1, 2024-2043. [Crossref]
6. Patrycja Hąbek. 2017. CSR Reporting Practices in Visegrad Group Countries and the Quality of Disclosure. *Sustainability* **9**:12, 2322. [Crossref]