Journal of Cleaner Production 196 (2018) 1034–1043

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Review

Success of voluntary sustainability certification schemes – A comprehensive review

Rasmus Tröster*, Michael Hiete

Ulm University, Helmholtzstr. 18, 89081, Ulm, Germany

ARTICLE INFO

Article history:
Received 7 December 2017
Accepted 28 May 2018
Available online 30 May 2018

Keywords:
Certification schemes
Sustainability
Literature review
Contingency analysis
Cross-sectoral

ABSTRACT

Voluntary sustainability certification schemes (CS) like the well-known Forest Stewardship Council are valuable approaches to address sustainability issues in different sectors. So far, a wide range of single topic studies exist but few emphasis has been put on analyses investigating factors considered responsible for shaping different dimensions of CS' success across different sectors and scientific disciplines. This study aims at closing this research gap by conducting a comprehensive systematic literature review of this growing body of research. Based on a database research, 226 records were selected to derive a framework explaining generic factors for the success of CS. To unveil broadly discussed factors and success dimensions within the literature, a contingency analysis was performed based on identified text passages. The relationship between specific factors and certain success dimensions is discussed comprehensively in the literature: The influence of the quality of the requirements and the capacity building measures on the problem solving capacity of CS as well as the influence of the characteristics of the adopting entities and to a lower degree governmental influences on the diffusion of CS are widely discussed. Moreover, the influence of the involvement of stakeholders and the transparency of the CS on the acceptance of CS is a main focus of analysis. A barely studied success dimension is the behavioral change towards more sustainable practices. These results enable and foster an interdisciplinary discussion on certification schemes which is necessary for understanding CS comprehensively and eventually driving practical improvements.

Contents

* Corresponding author.
E-mail addresses: rasmus.troester@gmail.com (R. Tröster), michael.hiete@uni-ulm.de (M. Hiete).

https://doi.org/10.1016/j.jclepro.2018.05.240
0959-6526/

1. Introduction

Negative effects of globalized industrial production are known for a long time and many attempts have been made to tackle these effects—including regulatory approaches, such as national legislation (taxes, bans etc.) and international environmental agreements. As a response to insufficient binding de jure state activities, private approaches have been developed (e.g., Mena et al., 2012; Marx, 2013). International corporations began to voluntarily improve and report their sustainability performance, a concept which was established under the term corporate social responsibility (CSR) (Mena et al., 2012). Moreover, private activities go beyond merely firm activities by involving different private actors and sometimes also public actors (Meidinger, 2003) in the form of sustainable voluntary private certification schemes (CS). Generally, CS are "[…] private in the sense of being formally independent of government but not reducible to industry self-regulation, and regulatory in that it purports to set and enforce standards for [sustainability] performance" (Bartley, 2007, p. 302).

Recently, academia showed a growing interest in analyzing CS and different scientific fields have begun to analyze the success of CS from various perspectives. In political science, authors are interested in the features of CS improving acceptance among various stakeholders, for example the involvement of different stakeholders into the governance of the CS (Marx, 2014) or the transparency of the CS (Auld and Gulbrandsen, 2010). Also, factors relevant for the impact on the ground, i.e. the environmental and social changes were studied, e.g. the consideration of local circumstances (Griscom et al., 2014) or the continuous improvement of the requirements over time (Melo and Wolf, 2007). Despite the great value of these merely single topic studies, few emphasis has been put on cross-sectoral analyses, comprehensively investigating the factors responsible for shaping different dimensions of CS′ success (Mori Junior et al., 2016) across different sectors and scientific disciplines. Thus, the purpose of the present study is to further investigate the reasons for varying success of different CS and thereby further closing the identified research gap. It is not the intention to analyze research dealing with specific types or sector-specifics of CS, for example studies on CS in the forestry sector, but to gain insights into this research field on a comprehensive level. Therefore, a comprehensive systematic literature review of the growing cross-sectoral body of research on the success of CS was conducted aiming to derive generic patterns and explain the barriers and drivers of CS′ success. Accordingly, this study's intention is not to examine different research fields in depth, but to characterize the most discussed factors relevant for the success.

By analyzing CS, this study contributes to the research field of cleaner production—as the main goal of CS is to prevent harmful effects during the production of different economic goods. By merging the scientific discussion comprehensively, this study captures the growing current debate on CS (e.g., Castka and Balzarova, 2018; Salim et al., 2018) and establishes a critical link between the diverse contributions in this research field. This study enables and fosters an interdisciplinary discussion on CS which is necessary for understanding CS comprehensively and eventually pursuing practical improvements.

The study is structured as follows: First, the general characteristics of CS are briefly discussed, followed by a description of the applied systematic literature review and the contingency analysis. Next, the descriptive results of the literature review and the analytical framework developed by the literature review are presented. Then, based on the contingency analysis, the most discussed success factors and success dimension in the literature are presented, before finally depicting the conclusion and limitations of the study.

2. Background: certification schemes

Due to the voluntary and private character, CS need to be distinguished from mandatory legal requirements. To face global (sustainability) problems, policy makers can draw on a variety of options for action. On one hand, they may exercise authority through their own policies. On the other hand, different types of international policy instruments may be applied. These are based, for example, on mutual negotiations or delegation to international institutions like the UN or the WTO (Coglianese, 2000). If countries leave problems unresolved intentionally or unintentionally, private organizations might feel obligated to contribute to solving the problem (Glasbergen, 2013). For example, in the case of forestry, environmental protection organizations regarded the outcome of the Rio de Janeiro Earth Summit as unsatisfying, which led to the development of the private CS organization, Forest Stewardship Council (FSC) (Pattberg, 2005).

2.1. Organizational features of CS

For setting and enforcing sustainable standard requirements, CS need to develop specific organizational features. Based on Bush et al. (2013), Loconto and Busch (2010) and Young (2015) the four main components are described and applied in the following to depict the construct of CS: Standard requirements, standard-setting process, implementation and governance.

Standard requirements are considered to be the most important component of the initiative. They describe the sustainability criteria that are set by the initiative and expected to be implemented on site (Young, 2015). Since CS were developed to tackle sustainability issues, the requirements should deal with at least one and preferably all of the three dimensions of sustainability. The requirements need to be developed by a certain institution. This is achieved by carrying out a *standard-setting process* (second component) which is, depending on the CS, more or less sophisticated. The extent of the standard-setting process is largely determined by the number and heterogeneity of involved actors. It might be a task carried out by the board of the CS, the CS' members or a broad range of external stakeholders. Regularly revisions should be conducted to adjust requirements to changing circumstances. The *implementation* (third component) refers to the application of standard requirements on site. One of the most crucial characteristics of implementation is the audit procedure, which describes the verification and certification on site, i.e. the assessment of compliance with the requirements. The attestation of the auditors in turn, is indicated as accreditation mechanism (International Organization for Standardization, 2005). Note, that the implementation is not

limited to the audits. Further aspects relevant in this dimension are, for example, capacity building measures to support the implementation or grievance mechanisms in the event that participants of the audit do not agree with the result or the conduct of other participants such as with the work of the auditors (Ascoly and Zeldenrust, 2003; Mori Junior et al., 2016).

Governance, the fourth component, addresses ownership and management aspects of the CS (Young, 2015). Most of the discussion in the literature deals with the highest decision-making authority (e.g., Solomon et al., 2006; Stetter and Zangl, 2012), for example a board of directors or, connected with more effort, a general assembly. However, bodies for daily business and financing structures are also discussed within the literature (Mallet, 2007; Wenban-Smith, 2007).

As stated above, CS may differ as a result of the involved stakeholder groups. A distinction is made between multi-stakeholder initiatives (MSI) and for-profit driven initiatives. MSI involve non-profit and for-profit participants and appropriate mechanisms to include all stakeholders in steering the initiative. In for-profit driven initiatives standard development and governance are only carried out by for-profit organizations (ORourke, 2006).

2.2. Success of CS

In contrast to a wide range of business studies, in the field of CS, a measurement of the success based solely on (financial) effectiveness would not satisfy the complexity of circumstances related to the way in which CS function. In the scientific discussion, several success dimensions are analyzed. The potential of a CS to solve or alleviate the sustainability problem that motivated its creation is a core success factor (Rapkin and Braaten, 2009). The positive change on site is seen as a crucial success variable, since it addresses the initial motivation of CS (without effects on site, the main purpose of the CS would not be addressed). As CS do not necessarily aim for addressing all three dimensions of sustainability (environmental, social and economic) or solely address specific issues, it would not be expedient to evaluate the problem-solving capacity comprehensively in all three dimensions. For example, a CS focusing on specific social aspects could perform poorly regarding environmental issues. This does not necessarily indicate an insufficient change on site, since the focus is set on social aspects and in this field the CS might perform excellently.

The acceptance of stakeholders is particularly important, since private CS are voluntary and not state authorized and thus need confirmation by external parties (Overdevest, 2010). In other words, actions of the organization need to be considered desirable, proper, or appropriate (Suchman, 1995) by third parties. Cashore (2002) identified four stakeholder groups relevant for granting authority to CS: governmental organizations representing the state, companies that implement requirements, supply chain actors that put pressure on implementing companies, as well as the civil society (environmental groups, the media, etc.). These stated groups are supposed to be especially important for granting acceptance, since they are linked to the CS either by direct involvement in the system (e.g., standard-setting or implementation) or indirectly by putting pressure on (non) participating companies. For example, acceptance by companies would increase the adoption rate and this is beneficial for the CS. The acceptance of the civil society could potentially lead to a higher pressure on the implementing companies to comply with the requirements of the CS and would eventually lead to a higher adoption rate as well.

3. Methodology

To identify the most discussed success factors and success dimensions, a systematic literature review was conducted. Systematic reviews minimize the bias in identification, selection, synthesis and summary of different studies (Moher et al., 2015). The aim of the review is to identify scientific records, in which the success of CS were analyzed and to derive drivers and barriers. Records from a variety of academic fields applying different methodological approaches were included to gain a comprehensive view of success factors relevant for CS. The data collection and selection is based on the following steps as suggested by Moher et al. (2015): Identification of material, screening of relevant studies and proof of eligibility (Fig. 1).

The first step (identification) includes recordings identified via a database search as well as records identified through cross-referencing. To perform the database search, the Web of Science database was chosen to systematically collect records from 1997 to April 2017. Title, abstract and key words of scientific papers in English language were searched, based on a search string using Boolean operators and search words. The choice of the Web of Science database is based on the comprehensive offer of high quality scientific journals from different major publishers. Furthermore, this database provides various Boolean operators which were very helpful to narrow down the results of the search string. The search string was broken down into three main dimensions. The first dimension contains search words and operators to identify CS. The search-string for searching CS was *certification* NEAR/5 (program* OR initiative* OR system* OR standard* OR scheme* OR approach* OR institution* OR organisation* OR organization*)*. Here, the NEAR operator was very helpful, since CS were named very differently, for example certification standard initiatives or certification non-governmental organizations. The NEAR operator finds records where the terms joined by the operator are within a specified number of words of each other, in this case within five words. The research focus lies on the CS as institution. Thus the combination of the word certification with a second term describing the institution was chosen. Cross-checking has revealed, for example, that searching for the single term *standard* would have been resulted in a sample of records that deal mainly with the characteristics of the requirements of CS. Second, the success dimension was included. To embrace a wide range of records dealing with different success dimension, the following search string was applied: *legitim* OR credib* OR accept* OR impact* OR effective* OR success**. The first three search words aim at identifying records in the field of acceptance of CS. The remaining words include records dealing with the impact of CS on the ground as well as discussing success in a generic way. Third, a search term

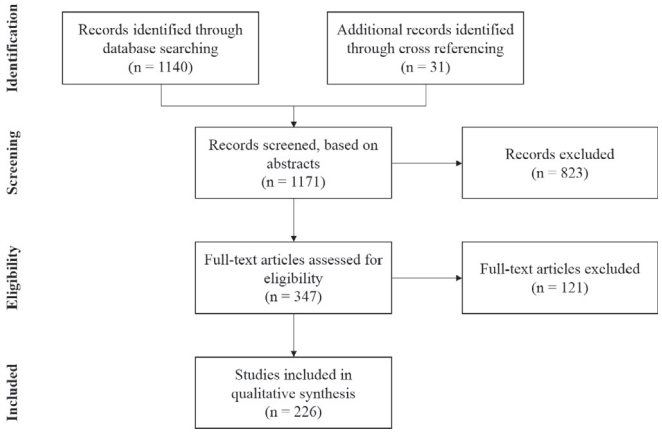

Fig. 1. Flow of information through different phases of the systematic review process (n for number of records).

involving the sustainability dimension was applied: *environment* OR social OR sustainab* OR green OR responsib* OR ecologic**. Finally, four search words inconsistent with the theme were chosen to be excluded from the records: *succession* OR directive OR DNA OR cancer*. All search terms were iteratively adjusted during the entire process. In addition to the database search, additional records were identified through cross referencing. In total *1171 records* were identified for further examination.

In the second step, abstracts were screened to further narrow down the records. The following criteria were applied to include records. The records' main focus lies on existing CS, implying that, for example, records pointing to missing CS in specific sectors or records that just mention but not analyze CS were excluded. Moreover, the records need to clearly focus on sustainability issues, so that for example papers dealing with quality management systems were excluded. Since this study focuses on voluntary and private CS, records that analyze legal systems or systems associated with state authorities were excluded. Finally, the abstracts were proofed whether the records cover the success of CS in their analysis or not. Records not considering success at all or records taking the success of other organizations but not of the CS into account were excluded. Applying these criteria resulted in a sample of *347 records*.

In the third step, an assessment of eligibility based on the full-text articles was conducted. Records that did not refer to drivers and barriers for the success of CS were excluded. This step further reduced the studies to a number of *226 records*. *Text passages* in each of the 226 records that included a statement about the positive influence of specific influences on the success were identified and coded accordingly. This means independent variables (factors relevant for the success) and dependent variables (success dimensions) were identified and used for the development of the analytical framework.

Finally, a contingency analysis was performed. The aim of this step was to identify which relationships (influence of success factors on success dimensions) were analyzed in the scientific discussion on CS' success more broadly. This was quantified by using the frequencies of the coded factors and success dimensions in the text passages. The analysis of contingencies follows Gold et al. (2009) and Sauer and Seuring (2017) who applied the same method, but on the level of records. The advantage of using coded text passages is that a content-related connection of the categories can be ensured, since text passages directly link cause and effect ensuring a content-related connection. In contrast, coding of records might mix aspects that are not necessarily linked on a content-related level (for example, when the factors are stated in completely different sections of a record that have no content-related connection). To illustrate the approach, the following text passage indicates a content-related relationship between context sensitivity (adaption to local circumstances, see below for detailed description) and the diffusion of CS. "For certification measures to attract and sustain large numbers of producers, its implementation has to consider local effects, and then, quickly modify its design, so as to resolve any emerging issues." (Bose et al., 2016, p. 954). Another advantage of the analysis of text passages is that the number of cases is higher, since more than one text passage per record can be included. In this study 512 text passages were identified.[1] Based on contingency tables, chi-square tests of independence were conducted to unveil significant relationships. A non-significant relationship indicates that the frequency of two coded aspects (factors relevant for the success and a success dimension) is equal or near to the expected frequency of these two aspects in the text passages. This would result in a low chi square value. A significant relationship, on the other hand, is characterized by a deviation of the frequency of two coded aspects compared to the expected frequency, leading to a higher chi square value. This indicates that the statistical null hypothesis might be rejected, i.e. independence of the factors cannot be assumed *in the entire population of research dealing with the success of CS*. For this study a significance level of 5% was chosen, indicated by a chi square value above 3.84. Furthermore, a minimum expected count of 5 is necessary to obtain valid results (Everitt, 2000).

The strength of the identified relevant relationships was further examined using the phi coefficient. The phi coefficient varies between −1 and 1 with a phi value of 0 proving independence. The strength of the relationship is measured by the frequency of two aspects coded together in the text passages. Suppose all text passages that are coded with aspect one are also coded with aspect two and all text passages that are *not* coded with aspect one are also *not* coded with aspect two, i.e. the two aspects are only coded together and do not occur separately in other text passages with other aspects. This would result in a perfect positive correlation and a phi value of 1. In the case of the present study this would represent the domination of a specific relationship of a factor relevant for the success and a success dimension in the text passages. A perfect negative correlation ($\varphi = -1$) would occur if all text passages that are coded with aspect one are not coded with aspect two and all text passages that are not coded with aspect one are coded with aspect two, indicating that the two factors do not occur together in the text passages. Note that this statistical approach does not provide information about the absolute frequency of factors coded together in the text passages. Instead, it depicts the relation between the frequency of the actually together coded aspects and the expected frequency of together coded aspects (for a detailed description of the underlying statistical approach see Fleiss et al. (2003). The annex provides further information and an example of the statistical analysis.

4. Results

4.1 Descriptive results

Research dealing with the success of CS gained attention, as the rising number of records over time indicates (Fig. 2). Notably, within the last five years, scientific activities appear to have strongly grown. As this review includes records published within the last 20 years, no statements can be made about records published before 1997.

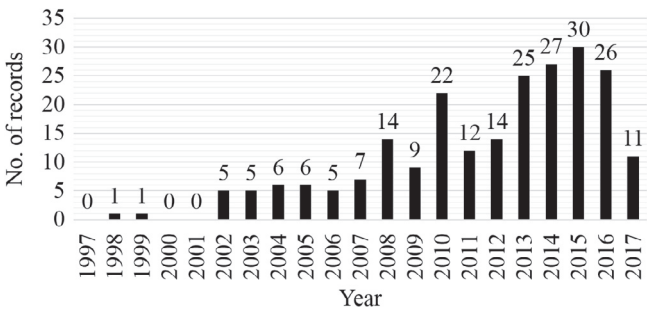

Fig. 2. Temporal distribution of the 226 analyzed records.

[1] For the contingency analysis only text passages were analyzed that contain both factors relevant for the success and success dimensions that were used for the framework. Text passages that contain only relationship between success dimensions or text passages that contain factors that were not applied for the framework were neglected. This reduced the number of records for the contingency analysis to 187.

The overwhelming share of records (151 of 226) addresses challenges in the forestry, agro-food and marine sectors (Table 1). This is not surprising, bearing in mind that the oldest and most widely studied CS stem from these sectors. Accordingly, the most studied CS in this sample are also from these sectors. In the forestry sector the FSC and the Programme for the Endorsement of Forest Certification Schemes (PEFC) are prominent research objects. With concern for the agro-food sector, Fairtrade, different organic certifications (Organic) and the Rainforest Alliance (RA) are extensively studied as is the Marine Stewardship Council (MSC) in the marine sector. Moreover, CS developed for improving the environmental performance of buildings, ISO 14001, (which can be applied cross-sectorally as well as the Leadership in Energy and Environmental Design (LEED)), were often examined. It is worth noting that recently, the number of analyses of CS in the field of tourism has increased, a rarely explored field. Considering the *regional focus* of the records, it is striking that the highest share of studies rely on generic analyses, meaning that the focus is on theoretical considerations. The main geographical foci of studies addressing specific regions are North America, Asia and Europe. Studies considering CS in the southern hemisphere (Africa, Central and South America) are less frequent in the record sample. This finding is especially remarkable, concidering sustainability issues tend to be particularly prevalent in southern developing countries (UNEP, 2012). As indicated above, methodically, a large proportion of the records rely on theoretical considerations, followed by mixed approaches, surveys, document and website analyses, qualitative interviews and database analyses. Only four reviews of the scientific literature were identified, which indicates that so far only relatively little work that grasps comprehensive patterns in the field of CS exists. Tscharntke et al. (2015) review experiences with and conservation impacts of coffee and cocoa certification, Mithoefer et al. (2017) develop hypotheses based on a review of studies in the sector of timber and tree crops and Englund and Berndes (2015) analyze studies dealing with CS that focus on biodiversity to compare the corresponding requirements. As previously mentioned, only the study of Mori Junior et al. (2016) was identified as a comprehensive approach examining drivers and barriers for the success of CS.

4.2. Analytical framework

The analytical framework of this study is divided into two main parts. On one hand, influencing factors for the success were derived inductively, i.e. based on the information gathered from the record sample. On the other hand, the success dimensions were coded deductively based on the classification of Young (1994) and Tikina and Innes (2008).

4.2.1. Factors relevant for the success of certification schemes

The scope of influencing factors is very comprehensive. In order to better structure the broad range of factors, external and internal factors were distinguished. Regarding external factors, the CS is not able to shape them deliberately. The following external factors were identified: characteristics of the adopting entity, governmental influences, interaction between the CS, and end customers' characteristics. Internal factors on the other hand are properties of the CS, which can be directly modified by the management of the CS. The following internal factors were identified: stakeholder involvement, quality of the requirements, capacity building, quality of audits, context sensitivity, continuous improvement, transparency of the CS and communication to customers of the adopting entities. The identified external and internal influencing factors were further categorized and are listed in Fig. 3.

The adopting entity implements the requirements developed by the CS. The type of adopting entity ranges from southern smallholders to large-scale multinational companies. The term

Table 1
Foci of the 226 analyzed records (shares; sums may deviate from one due to rounding).

CS focus		Sectoral focus		Regional focus		Methodological focus	
FSC	0.20	Forestry	0.34	Generic	0.28	Theoretical	0.25
ISO 14001	0.14	Agro-food	0.19	North America	0.16	Mixed	0.24
MSC	0.12	Marine, freshwater products	0.14	Asia	0.15	Survey	0.19
PEFC	0.09	Several sectors	0.14	Europe	0.12	Document and website analyses	0.11
Fairtrade	0.07	Housing, construction	0.06	Several regions	0.11	Qualitative interviews	0.09
Organic	0.07	Tourism	0.03	South America	0.09	Database	0.06
RSPO	0.06	Biomass	0.03	Africa	0.06	Review, meta-analyses	0.02
LEED	0.04	Generic	0.02	Australia	0.02	Modeling and Life Cycle Assessment	0.02
RA	0.03	Animal welfare	0.01	Central America	0.02	Data collection on site	0.01
Other CS	0.17	Other sectors	0.03			Other methods	0.01

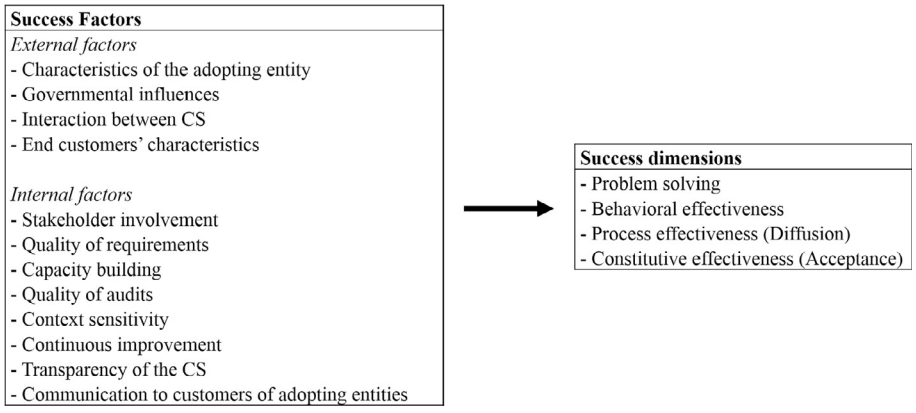

Fig. 3. Analytical framework.

characteristics of the adopting entity is used to describe cognitive aspects such as advantages of the CS as perceived by the adopting entities or the affinity of the adopting entities towards sustainability and CS. Moreover, organizational qualities and pressures by external groups were subsumed under this term. *Governmental influences* describe active supporting measures for CS (e.g., in the form of financial assistance), the configuration and strength of legal frameworks as well as the role of the state as a customer for CS. To consider the relation between CS, the *interaction between different CS* summarizes cooperation of CS in terms of takeovers, mergers, alliances but also competitive relationships between CS. *End customers' characteristics* are especially relevant for CS certifying products for consumers and subsume properties like demographics but also the attitude towards sustainability problems and CS.

The *involvement of stakeholders* refers to the balanced inclusion of different stakeholder groups as well as the participation of specific stakeholders such as scientists into different divisions of the CS. To determine the *quality of the requirements*, the stringency, scope as well as the precision of the requirements were considered. In order to facilitate the implementation of the requirements, CS have the option to provide *capacity building measures*, like group certifications, technical information for best practices as well as trainings for inspectors and for the adopting entities. Beside the quality of the requirements, there is the *quality of the audits* which considers the quality of the auditors, the independence of the auditors and whether they underwent accreditation. Moreover, the frequency and regularity as well as the strictness of the applied assessment methods were included in this category. To adapt to changing local conditions, CS might implement mechanisms that ensure *context sensitivity*, for example by applying different national requirements such or enabling a certain degree of flexibility of the requirements or during the implementation process. Mechanisms that ensure *continuous improvement* prevent stagnation of improvement of adopting entities after certification by shaping the requirements accordingly or the degree of compliance might be tiered. The *transparency of the CS* refers to the disclosure of relevant information on specific sections of the CS, for example transparent decision-making processes in the governance of the CS or the audit results in the course of the implementation process. Lastly, the *communication of information to the customers of the adopting entity* depicts mechanisms to forward sustainability information along the supply chain, like chain-of-custody approaches that ensure the traceability of certified products, namely, in the form of a physical separation.

4.2.2. Success dimensions

Four main dimensions were applied as framework for the coding of the literature: problem solving, behavioral effectiveness, process effectiveness and constitutive effectiveness (Fig. 3). *Problem solving* depicts the ability of the CS to solve or mitigate the problem it was initially developed for. This might be, for example, the improvement of biodiversity conservation in the forest sector (Englund and Berndes, 2015; Mas and Dietsch, 2004).

The *behavioral effectiveness* delineates the behavioral change of adopting entities to a more sustainable manner according to the requirements of the CS (Tikina and Innes, 2008). This is crucial, as only adopting entities with a high sustainability performance might implement the CS' requirements without any further efforts. Then, the CS' contribution to the problem solving on the ground might be very low and lower in comparison to a case where less sustainable entities adopt the CS and improve their sustainability performance (Stetter, 2015), even if the level finally achieved is below that of the high performers.

Attributing a pivotal role to behavioral change, a large-scale improvement of sustainability problems might only be achieved if a large number of adopting entities apply the CS. Thus, the market diffusion of a CS (process effectiveness) is an important success dimension. The presented work subsumes all studies that either deal with the initial implementation or the maintenance of compliance of requirements under the term process effectiveness,

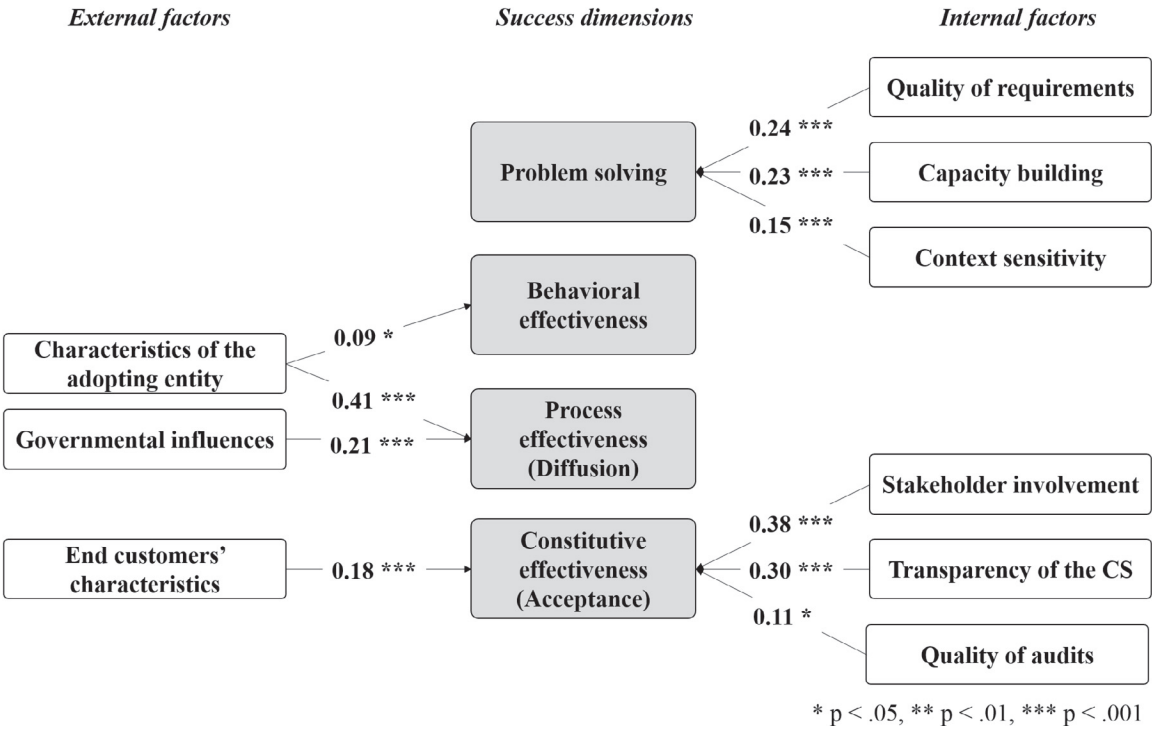

Fig. 4. Results of the contingency analysis (arrows illustrate significant relationships and values the strength, represented by phi coefficients).

since every new participant contributes to the diffusion of the CS. Furthermore, studies explicitly referring to the diffusion, for example measured as market share of the CS, but also studies relating to the certified area of a commodity, were included here.

The fourth success dimension concentrates on the *acceptance* granted by different stakeholder groups towards CS. This is especially important considering CS are voluntary approaches that have no legal obligation and are thus contingent on the support of the stakeholders (Bernstein and Cashore, 2007). Several studies further distinguish this success dimension, for example into legitimacy or credibility (Miller and Bush, 2015) but ultimately these are similar concepts as they both refer to the positive attitude of stakeholders towards CS.

Fig. 3 summarizes the analytical framework. The two main parts (success factors and success dimensions) were applied during the coding process and serve as foundation for the contingency analysis. The arrow illustrates the content-related relationship which is basic for the coding process of each text passage, i.e. success factors and success dimensions were coded together when the text passage makes a statement about the influence of a factor on one of the success dimensions.

4.3. Relationship between factors and success Dimensions

The result of the contingency analysis is illustrated in Fig. 4. Only significant dependencies with a positive phi value are illustrated, indicating broadly discussed success factors and success dimensions in the entire population of research dealing with the success of CS. The phi coefficients vary between 0.09 and 0.41, indicating a low to medium strength of the relationships in statistical terms (Fleiss et al., 2003). A higher phi coefficient indicates a more comprehensive scientific discussion on the influence of a certain success factor on a success dimension. The influence of the characteristics of the adopting entity on the process effectiveness are most analyzed in the scientific discussion ($\varphi = 0.41$), followed by the influence of the stakeholder involvement and the transparency of the CS on the acceptance ($\varphi = 0.38$ and $\varphi = 0.30$).

It is striking that certain factors are rarely discussed in the literature-indicated by not significant or significant but negative phi values (not illustrated in the figure): The interaction between CS (external factor), the continuous improvement and the communication to customers (both internal factors). Although these factors are not considered due to the statistical requirements, this does not necessarily imply that these factors are not studied at all but rather to a lesser degree than statistically expected.

Although the phi coefficients are rather low in statistical terms, the results can serve to identify general and important patterns, as it is intended by this study. In the following sections only widely discussed influencing factors, represented by comparably high phi coefficients are presented. In this regard, the behavioral effectiveness is excluded because the only factor discussed in this context has a very low phi coefficient (0.09).

4.3.1. Discussed factors in the context of the problem solving capacity

The influence of the *quality of the requirements* and the *capacity building measures* on the problem solving are comprehensively discussed in the literature on CS' success. This is indicated by comparably high phi coefficients ($\varphi = 0.24$ and $\varphi = 0.23$). According to the quality of the requirements, stringency plays a key role. Main et al. (2014) specify this condition by noting that the requirements should at least be equivalent to the level of the legal requirements, because otherwise CS would not generate additional value in terms of improving the sustainability performance. Similar important for the impact on the ground is the precision of the principles and indicators. On a practical level, this could refer to the precise differentiation of fish species groups the seafood sector which in turn leads to a higher impact since specific problems of the different fish species groups are addressed (Jonell et al., 2013). Moreover, several authors refer to the importance that a sufficient impact can only be achieved, if all relevant factors of a given sector are considered. For example, for the building and construction sector Wangel et al. (2016) highlight, that in the cases of BREEAM (Building Research Establishment Environmental Assessment Methodology) and LEED important aspects are neglected, such as toxic substances and emissions embodied in buildings and infrastructure which reduces the impact on the ground.

The scientific discussion of the influence of the capacity building measures on the problem solving capacity should rather be interpreted as indirect. For example, Zobel (2007) identifies group certification as an important capacity building measure that helps to reduce the environmental impact of firms in different sectors. Instead of the impact of group certification on problem solving, it can rather be assumed that the group certification enhances adaption which in turn might lead on the long term to an improvement of environmental problems.

4.3.2. Discussed factors in the context of the process effectiveness (diffusion)

Many studies analyze the influence of the *characteristics of the adopting entity* on the diffusion of CS ($\varphi = 0.41$). First, the perceived advantages and disadvantages resulting from implementing CS are recognized as important factors. Economic benefits resulting from the adoption, for example due to internal cost reductions (Bellesi et al., 2005) or price premiums for certified products (e.g. Newton et al., 2015) as well as perceived competitive advantages by entering new markets or increased competitiveness (e.g. Wiengarten et al., 2017), appear to be highly relevant for the adoption and thus diffusion of CS. Moreover, a hoped-for improvement of the reputation based on the implementation of the CS play an important role for the adopting entities (e.g. Crow and Danks, 2010). Second, the organizational quality of the adopting entity is seen as a cause for higher diffusion. Several authors refer to the size of the adopting entity which is associated with higher available resources and institutional capacity (e.g. Peiro-Signes et al., 2014). Greater resources and better structures facilitate the implementation of CS. Related to this, synergy effects with prior certifications (not necessarily in the field of sustainability, possible are also for example quality management systems) enhance the chance of adoption of CS (e.g. Zhu et al., 2013). Third, the affinity of the adopting entity towards sustainability and CS improve the diffusion according to the discussion in the literature as they increase the chance of implementing CS. This refers to the culture of the adopting entity and the employees or, if the adopting entity is a single person, the personal values and norms towards sustainability (Bellesi et al., 2005; Tey et al., 2015). The interest in and the knowledge about CS appears to be crucial, because if knowledge is lacking, the adoption rates could potentially also decline (e.g. Nukpezah et al., 2014). Fourth, authors mention the influence of pressure generated by external groups on the adopting entity to be an important factor for the adoption of CS. One of these groups are customers of the adopting entities (both end consumers and customers from business-to-business relationships, e.g. Holopainen et al., 2015), demanding certified products or sustainable practices and thus generate pressure to comply with CS. An additional important group is civil society. The pressure to adopt CS is not generated via market mechanisms, but by calling attention to negative sustainable practices or by calling for boycotts of products (e.g. Boiral et al., 2017).

The influence of *governmental interventions* on the diffusion are also widely analyzed in the scientific discussion on CS' success

($\varphi = 0.21$). In this context, governmental supporting activities play a role. National subsidy schemes for the preparation and compliance costs (Brandi et al., 2015) or technical assistance for smallholders (Kalfagianni, 2015) facilitate the implementation and therefore increase the adoption rate. Moreover, public organizations can be obliged to purchase certified products or to apply CS to products stemming from public territories, such as forests (Ruzevicius, 2009). High regulatory framework also drive the implementation of CS due to synergy effects (Meidinger, 2006; Stupak et al., 2016), because if high legal standards were already met, the implementation of CS is supposed to be more simple and less costly as organizational structures within the adopting entity are already implemented.

4.3.3. Discussed factors in the context of the constitutive effectiveness (acceptance)

The *involvement of stakeholders* is broadly discussed in the scientific literature as an important factor for the acceptance of CS ($\varphi = 0.38$). The main reason explaining this is a high number of studies mentioning the involvement and equal participation of stakeholders with different interests into the procedures of the CS as condition for the public acceptance of the CS. Several authors refer to this aspect by highlighting the importance of enabling *different* stakeholder groups to participate equally in the CS, for instance, in the form of consensus-based decision-making processes (e.g. Marx, 2014). On the level of the strategic governance, increasing acceptance can be achieved, for example, by giving decision making power to different members in the highest decision-making body (Auld and Gulbrandsen, 2010). On the level of the standard-setting, the same holds true as the involvement of different stakeholders is thought to gain acceptance, since the allegation of green-washing standards might be prevented as environmental groups contribute with higher demands on environmental and social aspects (e.g. Meidinger, 2006). In the course of participatory audits that use techniques such as role-play, different stakeholders are given the opportunity to provide information, for example about working practices (e.g. Schwarzbach and Richardson, 2015) which increase the credibility of the CS and therefore increase the acceptance. By providing valuable resources, *specific* stakeholders increase the acceptance of CS if they were properly integrated. Involving local and directly affected actors like indigenous peoples increases the acceptance because this group tends to lack resources to participate and efforts to involve these actors demonstrates efforts that the problem is taken seriously (Marin-Burgos et al., 2015). Similarly restricted in terms of resources and thus often underrepresented are NGOs. Their reputation in the public, sustainability agenda and lack of commercial interests make this stakeholder group an important factor for the acceptance of CS (Roberge et al., 2011a; von Geibler, 2013). Moreover, by providing profound and independent knowledge experts and scientists enhance the credibility of the CS in terms of scientific robustness (e.g. Eden and Bear, 2010).

Indicated by a phi coefficient of 0.3, the influence of the *transparency of the CS* on the acceptance of CS is also widely discussed in the literature. Transparency is seen to prevent the suspicion of greenwashing activities. The acceptance is fostered by disclosing information in different divisions of the CS, such as decision-making procedures (e.g. Auld and Gulbrandsen, 2010), the standard-setting procedures (Milder et al., 2015), the requirements themselves (Aguilar and Vlosky, 2008) or the assessment procedures as well as the outcomes of those processes, for example assessment result (e.g. Schwarzbach and Richardson, 2015).

4.3.4. Less discussed factors and success dimensions

Factors that are less discussed in the scientific discussion on CS' success are either not statistically significant ($p > 0.05$) or are significant but are indicated by a negative phi value. As mentioned above, it is striking that only the influence of one factor (characteristics of the adopting entity) on the *behavioral effectiveness* is discussed in the scientific discussion to a greater extent. The low phi value of 0.09 indicates that the discussion on factors relevant for the behavioral change is still relatively small compared to the other factors and success dimensions. This low consideration of the behavioral change as important success dimension is surprising, as this success dimension appears to be of great importance. It is argued, that only by modifying practices, noticeable impact on site might be generated (Stetter, 2015). It is assumed that several studies consider behavioral change indirectly in the success dimension *problem solving*. For example, if it is argued that capacity building influences the impact on the ground, a preceding step needs to be the behavioral change initiated by the capacity building measures.

Moreover, the impact of several influencing factors are less analyzed: The *interaction between CS* (external factor), the *continuous improvement* and the *communication to customers of adopting entities* (both internal factors). Nonetheless, few authors do emphasize the importance of these factors for the success. For example, cooperation of and harmonization between different CS are discussed as factors being relevant for the problem solving capacity of the CS, assuming that the stringency of the requirements increases through cooperation (Kalfagianni and Pattberg, 2013). Another example would be the positive impact of continuous improvement on the diffusion of CS, because if the adopting entities have lower hurdles to implement the requirements in the beginning and successively increasing their efforts to improve the sustainability performance, the number of adopting entities implementing and significantly changing their behavior is likely higher (e.g. Englund and Berndes, 2015). Lastly, the *quality of the requirements* might be of relevance for the *diffusion* although this is not identified by contingency analysis, because requirements with a high level of stringency might hinder the diffusion due to implementation efforts (e.g. Garrett et al., 2016). This problem might be approached by applying requirements increasing over time as it is discussed above.

5. Discussion

This study contributes to the research field of cleaner production by analyzing CS, which are important instruments for improving sustainability issues in different sectors. By merging the scientific discussion on success dimensions and influencing factors of CS comprehensively, this study works to capture the growing current debate on CS (e.g. Castka and Balzarova, 2018; Salim et al., 2018) and establishes a critical link between the contributions in this research field which are highly diverse, for example in terms of addressed industries such as the stone industry (Macedo et al., 2018), tourism (Margaryan and Stensland, 2017) or forestry (Espinoza et al., 2012). This study enables and fosters an interdisciplinary discussion on certification schemes which is necessary for understanding CS comprehensively and eventually drive practical improvements. Moreover, by analyzing the role of different institutions in the research on CS, for example the influence of companies, governmental actors and the civil society on the success of CS, this study serves as fundamental guidance for practical improvements addressing relevant actors.

Especially important in the light of future research, are few analyzed factors and success dimensions in the scientific discussion. For example, the success dimension *behavioral effectiveness* is less analyzed although the authors of the present study consider this success dimension to be highly relevant for the impact on site.

Here, future studies could analyze in depth the factors that influence the behavior of adopting entities. As this study's intention was to identify cross-sectoral comprehensive patterns, no analysis was performed differentiating the studies according to the foci or methods (Table 1). In future research this could be addressed by analyzing and comparing relevant success factors of CS in studies with different approaches, for example studies focusing on different sectors (forestry, agro-food etc.) or applying different methods (theoretical versus empirical). It could be possible, for example, that success factors exist that are specific for a particular sector or certain conditions but not for others and therefore do not appear as significant in the contingency analysis. Furthermore, the contingency analysis identifies comprehensively analyzed factors and success dimensions in the scientific discussion and thus gives an indication of the importance according to scientists. The view of practitioners could be different. Moreover, it would be promising to differentiate the analyzed literature sample into theoretical and empirical studies. By focusing on empirical studies it would be possible to get closer to "true" influences of factors. On the other hand, conceptual connections could serve for developing well-founded hypotheses in the field of CS.

As it is often the case with qualitative research, subjective factors play a role. The development of the framework in this study is such a case. Nonetheless, a considerable degree of validity can be ensured, as the success dimensions rely on well-established scientific references. Furthermore, the record sample which is based on the above stated search-string, probably does not cover the whole literature about CS, as the term CS is diversely defined and success dimensions might be differently named. This limitation is a key issue of this study as the findings might change when other studies would be included. Nonetheless, to identify systematic and general patterns as it was intended by this study, the identified record sample was deemed to be sufficient to represent the entire population of research dealing with success of certification schemes. On the one hand a high number (a total of 226) studies was included and a marginal utility regarding the number of papers can be assumed, i.e. the identified importance of factors on the success dimensions in the literature would not substantially change if further studies were included. On the other hand, the included studies are highly diverse, for example in terms of sectoral, methodological and regional focus. This ensures a realistic picture of the broad research on CS. The basic intention of the contingency analysis also reduces the flaw of not covering the entire population of research dealing with this topic. The chi square test allows for making statements about the entire population, i.e. statistically significant values indicate for general patterns that likely also occur in the population. Furthermore, a CS can be seen as a complex structure and success factors are likely not independent from each other. For example, it is likely that some success factors reinforce other factors whereas others might be neutral or in rare cases even negatively affect others. Moreover, it is likely that varying one factor will also induce changes in others (Hiete et al., 2011). Finally, one can expect at least for some factors thresholds to become effective and saturation effects.

The coherences identified by the contingency analysis highly depend on the categorization of the framework; broad categories like governmental influences tend to be more significant as the number of aspects falling under such a category is higher; from this a bias of the coherences might result. Last but not least, the coded text passages were not explicitly differentiated depending on whether or not results or assumptions were presented. This could be differentiated in future studies. Altogether, the present study appears to be an adequate approach to comprehensively grasp the most important discussion points of the growing number of cross-sectoral and diverse studies dealing with the success of CS and stimulating further research.

6. Conclusions

Due to the limited scientific discourse about generic and cross-sectoral success factors for CS, the intention of this study was to comprehensively grasp these factors within the broad and cross-sectoral scientific literature. To achieve this goal, a comprehensive systematic literature review of the growing body of research dealing with the success of CS was conducted and a framework was inductively and deductively derived, based on the relevant text passages. To unveil the most discussed factors and success dimensions, a contingency analysis was applied.

Most of the studies address the forestry, agro-food and marine sector and study well-established CS such as FSC, MSC or LEED. The main geographical foci of studies are specific regions such as North America, Asia and Europe. Methodically, a large proportion of the records rely on theoretical considerations, followed by mixed approaches, surveys, document and website analyses, qualitative interviews and database analyses. Only few reviews of the scientific literature were identified, indicating that so far only relatively few works exist that analyze comprehensive patterns in the field of CS.

Based on the coding process as well as existing classifications an analytical framework was developed: Four success dimensions were derived deductively and four internal and eight external factors were developed inductively. According to the contingency analysis, most discussed influence factors for the problem solving capacity of CS are *the quality of the requirements* and *capacity building measures*. Most discussed factors affecting the diffusion of CS are *characteristics of the adopting entities* and *governmental* influences. The third success dimension *acceptance* is in the literature mainly discussed in connection with the *involvement of stakeholders* and the *transparency of the CS*. Additionally, the success dimension *behavioral effectiveness* is poorly addressed in the scientific discussion as well as certain factors such as the *interaction between CS*, the *continuous improvement* and the *communication to customers of adopting entities*.

By merging the scientific discussion, this study establishes a critical link between the contributions within this highly diverse research field. Therefore, this study fosters the interdisciplinary discussion on CS which is important for understanding the seemingly abstract concept of CS comprehensively and eventually drive practical improvements by providing involved stakeholders with general directions for a supportive behavior.

Acknowledgements

This study is part of the "NamiRo" research project on sustainably-produced minerals (www.namiro-projekt.org/english/). This project is financially supported by the German Federal Ministry of Education and Research (grant no. 01UT140).

Annex

Table 2 provides an example of a contingency table of the influence factor *stakeholder involvement* and the success dimension *acceptance*. The values in bold display the number of coded factors and success dimensions in the text passages. For example, stakeholder involvement and acceptance were coded together 49 times. Contrary, 341 text passages contain none of the two. Based on this information the expected count is calculated using the marginal frequencies and probabilities. For example, 64 text passages contain stakeholder involvement and either acceptance or no acceptance. This is a share of 0.125 of all text passages. For

calculating the expected count this share is multiplied with the marginal frequencies of the acceptance, i.e. the number of text passages that contain acceptance and no acceptance (and both stakeholder involvement or no stakeholder involvement). The rationale of this step is that it is assumed that since the share of *all text passages containing stakeholder involvement* is 12.5%, also text passages containing stakeholder involvement *and* acceptance should be 12.5% as well. The resulting value of 19.5 demonstrates the expected frequency of text passages containing stakeholder involvement and acceptance. The observed value of 49 deviates positively from this value.

Table 2
Example of a contingency table for the influence factor *stakeholder involvement* and the success dimension *acceptance*.

			Acceptance		Marginal frequencies/ probabilities
			0	1	
Stakeholder Involvement	0	**Count**	**341**	**107**	**448 (0.875)**
		Expected Count	311.5	136.5	
	1	**Count**	**15**	**49**	**64 (0.125)**
		Expected Count	44.5	19.5	
		Marginal frequencies	**356**	**156**	**512**

The calculation of the chi square value takes these deviations into account. The squared deviations of the observed values and the expected values in all four cases are divided by the expected value.

$$x^2 = \sum_{j=1}^{k} \frac{\left(h_{bj} - h_{ej}\right)^2}{h_{ej}} \quad (1)$$

Where h_{bj} are the observed values and h_{ej} are the expected values. To test for significance, the resulting chi square value (in this case 73.35) is compared with a value which corresponds to the significance level of the chi square table. In this case the value is 3.84 as a significance level of 0.05 was chosen. If the chi square value exceeds this value, the assumption of independence (Null hypotheses) can be rejected, i.e. a statistically significant correlation is likely. A minimum expected count of 5 is necessary to obtain valid results since otherwise the quotient would increase considerably and bias the result.

To quantify the correlation, the phi value is calculated from square root of the chi square value divided by the total number of text passages (n).

$$\varphi = \sqrt{\frac{x^2}{n}} \quad (2)$$

A perfect positive correlation ($\varphi = 1$) would occur when all text passages that are coded with acceptance are also coded with stakeholder involvement *and* all text passages that are not coded with acceptance are also not coded with stakeholder involvement (Fleiss et al., 2003).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclepro.2018.05.240.

References

Ascoly, N., Zeldenrust, I., 2003. Considering complaint mechanisms. In: An Important Tool for Code Monitoring and Verification. Accessed August 11, 2017.
Bartley, T., 2007. Institutional emergence in an era of globalization. The rise of transnational private regulation of labor and environmental conditions. Am. J. Sociol. 113, 297–351.
Cashore, B., 2002. Legitimacy and the privatization of environmental governance. How non-state market-driven (NSMD) governance systems gain rule-making authority. Governance 15, 503–529.
Castka, P., Balzarova, M.A., 2018. An exploration of interventions in ISO 9001 and ISO 14001 certification context – a multiple case study approach. J. Clean. Prod. 174, 1642–1652.
Coglianese, G., 2000. In: Nye Jr., Joseph S., Donohue, John D. (Eds.), Globalization and the Design of International Institutions. Governance in a Globalizing World, Brookings.
Espinoza, O., Buehlmann, U., Smith, B., 2012. Forest certification and green building standards. Overview and use in the U.S. hardwood industry. J. Clean. Prod. 33, 30–41.
Everitt, B., 2000. The Analysis of Contingency Tables, 2. ed. 1. CRC repr. Chapman & Hall, London.
Fleiss, J.L., Levin, B., Paik, M.C., 2003. Statistical Methods for Rates and Proportions, 3. ed. Wiley-Interscience, Hoboken, NJ.
Gold, S., Seuring, S., Beske, P., 2009. Sustainable supply chain management and inter-organizational resources. A literature review. Corp. Soc. Responsib. Environ. Manag. 16, n/a-n/a.
Hiete, M., Kühlen, A., Schultmann, F., 2011. Analysing the interdependencies between the criteria of sustainable building rating systems. Construct. Manag. Econ. 29, 323–328.
International Organization for Standardization, 2005. Conformity Assessment – Vocabulary and General Principles (ISO/IEC 17000:2004) (ISO/IEC 17000:2004). Accessed February 12, 2016.
Loconto, A., Busch, L., 2010. Standards, techno-economic networks, and playing fields. Performing the global market economy. Rev. Int. Polit. Econ. 17, 507–536.
Macedo, D., Mori Junior, R., Carvalho, L.S., Mizusaki, A.M.P., 2018. Sustainability certification scheme for the dimension stone industry in Brazil. A proposal for an initiative based on the northwest region of Espírito Santo State, Brazil. J. Clean. Prod. 182, 896–909.
Mallet, P., 2007. ISEAL Emerging Initiatives. Module 1: Overview, London. Accessed May 30, 2016.
Margaryan, L., Stensland, S., 2017. Sustainable by nature? The case of (non)adoption of eco-certification among the nature-based tourism companies in Scandinavia. J. Clean. Prod. 162, 559–567.
Meidinger, E., 2003. Forest Certification as a Global Civil Society Regulatory Institution. Social and Political Dimensions of Forest Certification, pp. 265–292.
Mena, S., Palazzo, G., Arnold, D.G., 2012. Input and output legitimacy of multi-stakeholder initiatives. Bus. Ethics Q. 22, 527–556.
Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A., 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst. Rev. 4, 1.
ORourke, D., 2006. Multi-stakeholder regulation. Privatizing or socializing global labor standards? World Dev. 34, 899–918.
Pattberg, P., 2005. What role for private rule-making in global environmental Governance? Analysing the forest stewardship Council (FSC). Int Environ Agreements 5, 175–189.
Rapkin, D., Braaten, D., 2009. Conceptualising hegemonic legitimacy. Rev. Int. Stud. 35, 113.
Salim, H.K., Padfield, R., Hansen, S.B., Mohamad, S.E., Yuzir, A., Syayuti, K., Tham, M.H., Papargyropoulou, E., 2018. Global trends in environmental management system and ISO14001 research. J. Clean. Prod. 170, 645–653.
Sauer, P.C., Seuring, S., 2017. Sustainable supply chain management for minerals. J. Clean. Prod. 151, 235–249.
Solomon, F., Schiavi, P., Horowitz, L., Rouse, A., Rae, M., 2006. Mining Certification Evaluation Project (MCeP). Final Report – January 2006.
Stetter, A., 2015. Global Standards and Certification Schemes. Institutional Design and Successful Sustainability Governance. Dissertation, Munich.
Stetter, A., Zangl, B., 2012. Certifying Natural Resources. A Comparative Study on Best Practice and Future Scenarios for the Certification of Trading Chains in Mineral Production.
Suchman, M.C., 1995. Managing legitimacy. Strategic and institutional approaches. Acad. Manag. Rev. 20, 571–610.
Tikina, A.V., Innes, J.L., 2008. A framework for assessing the effectiveness of forest certification. Can. J. For. Res. 38, 1357–1365.
UNEP, 2012. Global Environment Outlook GEO 5. Environment for the Future We Want. United Nations Environment Programme, Nairobi.
Wenban-Smith, M., 2007. ISEAL Emerging Initiatives. Module 5: Financial Models, London. Accessed May 30, 2016.
Young, O.R., 1994. International Governance. Protecting the Environment in a Stateless Society. Cornell Univ. Press, Ithaca.
Young, S.B., 2015. Responsible sourcing of metals. Certification approaches for conflict minerals and conflict-free metals. Int. J. Life Cycle Assess.