Special Issue: 60th Anniversary

Questioning Impact: A Cross-Disciplinary Review of Certification Standards for Sustainability

Business & Society
2022, Vol. 61(5) 1042–1082

DOI: 10.1177/00076503211056332
journals.sagepub.com/home/bas


Sylke F. Jellema[1], Mirjam D. Werner[1], Andreas Rasche[2], and Joep Cornelissen[1,3]

Abstract

This article provides a review of scholarly approaches to assessing the impact of certification standards for sustainability. While we observe that some theoretical advances have afforded a better understanding of the potential impacts of adopting such standards, we also find that progress has been constrained due to a strong emphasis on assessing impact via linear causal pathways. This linear focus on the net effects for single stakeholders, such as farmers and producers, local communities and ecosystems, falls short of adequately capturing the broader impact of certifications across social and ecological dimensions. Inspired by theories on complex systems thinking, we present a framework based on a systems-based impact logic that better captures and assesses the impacts of certification standards within broader social-ecological systems. Our framework can be used as a heuristic to design impact-related studies and assess the impact of certification standards across disciplinary vantage points and empirical contexts.

Keywords

certification standards, effectiveness, impact, sustainability, systems thinking

[1]Rotterdam School of Management, Erasmus University, The Netherlands
[2]Copenhagen Business School, Denmark
[3]University of Liverpool, UK

Corresponding Author:
Sylke F. Jellema, Business-Society Management, Rotterdam School of Management, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands.
Email: jellema@rsm.nl

The governance, assessment, and assurance of sustainability within businesses and global supply chains is increasingly implemented through the use of certification standards. Such standards reflect "voluntary predefined rules, procedures, and methods to systematically assess, measure, audit and/or communicate the social and environmental behavior and/or performance of firms" (Gilbert et al., 2011, p. 24). Certification standards are designed to improve the lives of workers and encourage the widespread adoption of sustainable business practices. Standards support these aims by defining rules; for example, for the use of environmentally friendly farming practices, together with accompanying price premiums, guaranteed minimum commodity prices, and fair labor conditions.

In recent years, certification standards such as those issued by the Forest Stewardship Council (FSC), the Marine Stewardship Council (MSC), UTZ, and Rainforest Alliance have been adopted by an ever-increasing number of businesses throughout the world. A recent assessment of 40 widespread standards found them to be in use in over 170 countries, with over 10,000 companies applying these standards in their operations (MSI Integrity, 2020). Certification standards have thus moved into the mainstream and now constitute an important basis for improving firms' social and environmental performance. However, the increased uptake and simultaneous expansion of certification standards has also led to a multitude of standards being used, thereby creating a generally foggy field for businesses, stakeholders, and consumers (Reinecke et al., 2012).

This article reviews and critically interrogates the literature on the actual impact of such certification standards. Here, we understand impact quite generally as "a change in an outcome" (Impact Management Project, 2021) that can be attributed to a certification scheme. Such outcomes can be positive or negative, direct or indirect (Clark et al., 2004). We review academic work focused on questions regarding the outcomes that sustainability certifications generate for society, including the ways in which these standards are claimed both to improve the welfare for stakeholders across the globe and to limit the negative impacts of business on the environment. These impact-related questions are central not only to research in the business and society domain but also crucial for informing policy debates and discussions in practice. Indeed, public debate has increasingly turned toward this question of impact in recent years, with exposés of contemporary labor conditions such as the *Washington Post's* article "The children who harvest cocoa" and the BBC documentary "The Real Cost of a Cuppa" openly criticizing certification programs for their lack of broader social and environmental impact. At the time of writing, Greenpeace International (2021) "Destruction: Certified" is but the latest report to conclude that "certification is not a solution to deforestation, forest

degradation and other ecosystem conversion" (p. 1), with Greenpeace criticizing certification schemes for greenwashing products that are known to be connected with ecological destruction and for impeding the adoption of more effective alternative measures.

Our current knowledge about the impact of certification standards remains limited in at least two important ways, however, even in spite of numerous recent scholarly reviews on certification standards (Arton et al., 2020; Auld et al., 2008; de Bakker et al., 2019; Sartor et al., 2016; Tröster & Hiete, 2018). This is primarily because such reviews are often too broad in terms of the issues they discuss, combining insights from certification standards with other types of standards, for example (de Bakker et al., 2019), or combining work on the impact of such standards with other issues such as their creation and dissemination (Tröster & Hiete, 2018). Other reviews fall short in providing broad-based insights as these are geared too much toward specific contexts and perspectives, focusing, for example, on single initiatives or sectors (Arton et al., 2020; Auld et al., 2008; Sartor et al., 2016) or addressing impact only in the context of a specific academic discipline (Dragusanu et al., 2014).

In addressing these limitations, we critically review and problematize the ways in which the impact of certification standards has hitherto been conceptualized and assessed, looking beyond single initiatives or sectors and by adopting a cross-disciplinary angle to compare alternative vantage points, assessment frameworks, and measures of the impact of certification across a range of disciplines and research streams (Grabs et al., 2021). Aiming to provide a critical and comprehensive account of the impact assessment of certification schemes, our overall approach can best be characterized as a problematizing review (Alvesson & Sandberg, 2020) that critically interrogates the existing literature on the impact of certification standards by theoretically synthesizing and problematizing existing conceptualizations and methods of impact assessment. The aim of our approach is to "generate new and 'better' ways of thinking about" impact assessment more generally (Alvesson & Sandberg, 2020, p. 1291) by addressing the following two related research questions: (a) *How have studies researched and conceptualized the impact of certification standards for sustainability?* and (b) *How can we apply the findings of our review and theoretical synthesis of the literature to better conceptualize and capture the impact of such standards?*

We first collected and summarized academic work published between 2010 and 2020 on the impact of certification standards, categorizing and reviewing this body of research according to differences in approaches to impact assessment. In this article, we start out by discussing how previous studies have approached the impact of certification standards on various stakeholder groups, including farmers and producers, local communities, and the corporations that

adopt these standards. From this analysis, we conclude that the predominant impact logic underlying impact assessments is premised on a narrow and linear model of causation that does not align well with the complex ways in which impact is in many instances made and sustained, thereby under-representing the actual impacts realized as a result of certification standards for sustainability.

Informed by the findings of our review, we problematize this narrow but prevalent focus, proposing an alternative framework based on systems thinking that can be applied in future impact studies both within the certification context as well as in researching other topics in the domain of business and society. Our aim with this framework is to offer a heuristic and baseline for impact assessment that is better able to account for the complex causal pathways and for the various emergent outcomes and impacts of sustainability standards. Our study offers three main contributions to the literature. First, we review research on certification standards across disciplines and elaborate the ways in which such standards have been found to impact different stakeholder groups. Second, we identify the most prevalent research approaches to studying the impact of certification and elucidate the theoretical problems associated with these approaches. Third, we draw on systems thinking to propose a framework that addresses these problems and which can be used both to compare and integrate findings across disciplines and contexts as well as to guide further research on societal impact (cf. Alvesson & Sandberg, 2020).

The following section of this article (section "Methods") outlines the methodological steps taken to collect and analyze the material for our review. The Section "Overview of research on the impact of certification standards" discusses the overall findings of our review and details how the impact of certification standards has been conceptualized and examined in previous studies, identifying the theoretical challenges and problems associated with the current body of work on certification and its reliance on impact-related assumptions based on a logic of linear causation. The Section "Assessing impact" outlines a systems-based framework with which to address these challenges and problems in future research, while the final section concludes with a discussion of the implications of our review for researchers and practitioners, including several suggestions for future research.

Methods

Phase 1: Scoping the Review

The scope of our review is intentionally limited to studies that focus on certification standards for sustainability and that address the impact of these standards from multiple disciplinary perspectives. Given that this scope is best understood by comparing it with other reviews in the field, Table 1 gives an

Table 1. A grid to compare literature reviews on sustainability standards.

	Type of sustainability standard										
	Type of standard			*Scope of review*		*Stakeholder orientation*		Thematic focus		Disciplinary orientation	
Study	Principle-based	Certification	Reporting	Single or a set of standards	All standards	Multistakeholder	Single-stakeholder	Broad	Narrow	Single discipline	Across disciplines
This study		X			X	X	X		Impact		X
Auld et al. (2008)		X		Forestry		X	X		Impact		X
Blackman & Rivera (2011)		X		Forestry Fairtrade		X			Impact		X
Dammert & Mohan (2015)		X		Fairtrade		X			Impact	Economics	
de Bakker et al. (2019)	X	X	X		X	X		X			X
Dragusanu et al. (2014)		X		Fairtrade		X		X		Economics	
Oya et al. (2018)		X		Agriculture		X			Impact	Development	
Prieto-Sandoval et al. (2016)	X	X	X		X	X		X			X
Rasche et al. (2013)	X			UNGC		X		X		Management	
Ruggeri et al. (2019)		X		Fairtrade		X		X			X
Sartor et al. (2016)		X		SA8000		X		X			X
Tröster & Hiete (2018)		X			X	X		X			X

overview of the three scoping criteria to compare reviews of the literature on sustainability standards. Below we discuss these three scoping criteria in further detail.

Focus on certification standards. The literature has distinguished three types of sustainability standards (Rasche, 2010; Waddock, 2008): (a) principle-based standards that outline general requirements for improving firms' sustainability performance without monitoring or verifying corporate performance (e.g., the UN Global Compact); (b) certification standards, such as Social Accountability (SA) 8000 that measure and monitor firms' compliance with predefined rules along global supply chains; and (3) reporting standards that outline criteria for harmonizing and benchmarking the disclosure of companies' nonfinancial information (e.g., the Global Reporting Initiative). The focus of this study is exclusively on certification standards. In many cases, certification standards are used to govern buyer-supplier relations in global value chains, as for example in the forestry, coffee, and textile industries. Prominent examples of such standards include Fairtrade, the Roundtable for Sustainable Palm Oil (RSPO) and SA 8000. Accordingly, we exclude principle-based and reporting standards from our review. We also exclude internal firm-level codes of conduct (e.g., Nike's supplier code), as well as specialized certification standards that do not directly focus on sustainability-related issues (e.g., quality certifications such as ISO 9001).

Our study includes all types of certification standards for sustainability-related topics, including both standards that are set up and governed jointly by multiple stakeholder groups, i.e., multistakeholder initiatives (MSIs) and standards that predominantly involve a single-stakeholder group, i.e., single-stakeholder initiatives (SSIs). We do not focus on one particular certification standard (e.g., on studies that review only academic work discussing the FSC) or standards from a specific sector (e.g., those reviewing only academic works that discuss standards from the forestry sector). This broad and inclusive approach is helpful, we contend, in allowing us to include studies that address different initiatives and sectors, while staying focused on the already-heterogeneous context of certification standards (Reinecke et al., 2012).

Focus on impact. The second key element of our review is a direct focus on how the impact of certification standards has been conceptualized and assessed, thus addressing a theoretical and methodological issue that has recently become a central feature of business–society relations (see, for example, Barnett et al., 2020). In line with prior impact-related research, we adopt a rather broad working definition of the impact of certification standards, defining such impact as any positive or negative changes in outcomes

Table 2. An overview of impact definitions in the social and environmental context.

Definition of impact[a]	Source
"Outcome improvement attributed to CSR activities."	Barnett et al. (2020, p. 948)
" . . . the process of providing evidence that *an organization* is providing a real and tangible benefit to the community or the environment."	Grieco et al. (2015, p. 2)
" . . . the portion of the total outcome that happened as a result of the activity of [a] venture, above and beyond what would have happened anyway."	Clark et al. (2004, p. 7)
"*Monitoring* tracks key indicators of progress over the course of a program as a basis on which to evaluate outcomes of the intervention. [. . .] *Impact evaluation* studies whether the changes in well-being are indeed due to the program intervention and not to other factors." (Emphasis in original)	Khandker et al. (2010, p. 7)
"A change in an outcome caused by an organisation. An impact can be positive or negative, intended or unintended."	Impact Management Project (2021)
" . . . the processes of analysing, monitoring and managing the intended and unintended social consequences, both positive and negative, of planned interventions (policies, programs, plans, projects) and any social change processes invoked by those interventions."	International Association for Impact Assessment (2015, p. 1)

Note. CSR = corporate social responsibility.
[a]This includes definitions of the terms "impact," "impact evaluation," "social impact," and "environmental impact" in the context of assessment and evaluation.

for different stakeholder groups that can be attributed to such standards. This understanding of impact is informed by prior scholarly work and definitions of social and environmental impact (see Table 2 for an overview). Although such definitions differ in certain details, they all emphasize that studying impact entails looking at how an intervention (e.g., via the adoption of a standard) produces or causes a change in outcomes (cf., Clark et al., 2004; Impact Management Project, 2021; Khandker et al., 2010). In accordance with this definition, this review thus mostly focuses on: (a) impact through certification standards and (b) impact on different stakeholder groups.

Most reviews of the literature on certification standards have generally adopted a broad thematic approach, discussing scholarly work that focuses on themes as diverse as the input legitimacy of standards, their dissemination and institutionalization, as well as the process of their development (de Bakker et al., 2019; Rasche et al., 2013; Tröster & Hiete, 2018). The few reviews that have focused on impact as a guiding theme have been exclusively concerned with the effects of certifications in a specific commodity sector (Arton et al., 2020; Auld et al., 2008; Blackman & Rivera, 2011). Given that the discourse on certification standards has evolved significantly in recent years, however, a more differentiated perspective is warranted and desirable. We believe our approach of combining a focused review of one area of academic interest (i.e., impact) with a comparison of certification standards across different industries can thus make a meaningful and timely contribution to ongoing debates in the field.

Focus across disciplines. Finally, our review is cross-disciplinary in nature. This is because reviews that focus exclusively on discussing impact in the context of a specific discipline, such as economics (Dragusanu et al., 2014), downplay the cross-disciplinary character of the discourse that has amassed around the potential impacts of certification standards. In this article, by contrast, we extend our scope beyond studies in the management literature so as to increase the likelihood of drawing upon insightful alternative perspectives on certification standards. Political science–oriented studies, for instance, tend to approach the topic from a more macro level, in terms of governance mechanisms or national and international development needs (Nesadurai, 2013), while ecological studies are mostly interested in the environmental effects of certification standards on issues such as resource conservation or the use of hazardous chemicals. Although some extant reviews do contain literature from different disciplines (Tröster & Hiete, 2018), our study differs from such reviews in that it explicitly reflects on the contributions made by different disciplines. In adopting this cross-disciplinary lens, we bridge disciplinary areas and connect different streams of work, thus adopting an approach we feel is more closely aligned with the complexity of the business–society interface.

Phase 2: Data Collection and Search Strategy

Search strategy. We used a Boolean search string to gather the corpus of published articles for this review. This string was chosen after an extensive scoping exercise in which we identified our search terms based on the keywords included in seminal studies within the field (see Table 3). Our key search terms were either synonyms for certification standards and their impact or

Table 3. Search strings used during data collection.

Certification terms	"Eco-label" or "Ecolabel" or "Eco label" or "Sustainability certification" or "Sustainability label" or "Sustainability standard"
	OR
Prominent certification initiatives	"Alliance for Water Stewardship" or "B Corp*" or "Ethical trading Initiative" or "Fair Labor Association" or "Fairtrade" or "Fair Trade" or "Forest Stewardship Council" or "International Sustainability and Carbon Certification" or "Marine Stewardship Council" or "Rainforest Alliance" or "Roundtable on Sustainable Biomaterials" or "Roundtable on Sustainable Palm Oil" or "Round Table on Responsible Soy" or "RSPO" or "SA8000" or "SA 8000" or "Sustainable Forestry Initiative" or "UTZ Certified"
	AND
Impact terms	"Effectiveness" or "Impact" or "Outcome*" or "Performance" or "Result*" or "Value"

Note. RSPO = roundtable for sustainable palm oil; SA = social accountability.

related terms, as well as the names of several prominent certification initiatives frequently covered in the scientific literature. These initiatives included several standards that have a strong focus on social issues (e.g., the Fair Labor Association, Fairtrade, SA8000, and the Ethical Trade Initiative) as well as more environmentally focused criteria (e.g., the RSPO, the FSC, and the MSC). Several searches were conducted to fine-tune the final search protocol. It was only after rephrasing the term "label" to "eco-label," for example, that we were able to narrow down the results from over 25,000 studies on labeling in general to under 1,000 studies on labels for environmental and social impact. For instance, whereas the term "label" can refer to any type of classifying or categorical term, adding the prefix "eco" constrains the search to labels that specifically signal environmental and social attributes (see for instance the Ecolabel Index, a global directory including hundreds of ecolabels focusing on both environmental and/or social criteria). After every search round, moreover, we screened a subsample of articles to assess their relevance to our research aims. With this screening, we also ensured that our final sample included only those studies that focus on certification standards as their core subject, rather than as an empirical context for another topic or research question.

We conducted our searches in the Social Sciences Citation Index of the Web of Science database. The Web of Science database is an independent interdisciplinary database that covers a broad range of journals from different publishers and has frequently been used in systematic literature reviews in

the area of business and society (de Bakker et al., 2019; Tröster & Hiete, 2008). We acknowledge, however, that by relying on a single database, we may not have captured all the important works written on this topic; for instance, the Web of Science does not include newly established journals. Given our aim of conducting a problematizing review of scholarly approaches to assessing impact, however, our intention was not to review every single study out there, but rather to collect a broad and representative sample of studies in the field.

We included articles published between 2010 and 2020, selecting this time window on account of the proliferation of certification standards since the 2010s and the increasing scholarly attention to impact and sustainability throughout this period. We found, for instance, that most standards were set up around the early 2000s, whereas the academic articles on these standards started to proliferate from 2005 onwards. Based on these insights, we decided to limit our focus to articles between 2010 and 2020 for two reasons; (a) to focus on the most recent insights on the impact of certification standards and (b) to keep the number of articles manageable. Our final search was conducted in August 2020, yielding a total of 1,015 studies. We then removed any duplicates resulting in a total of 952 articles.

Selection criteria. To limit our scope to those studies that add directly to our knowledge on the impact of certification, we only included peer-reviewed articles in the selection process. Our first step was thus to filter the sample by removing all non-peer-reviewed publications, including conference proceedings, editorials, commentaries, and workshop descriptions. We also removed 11 literature reviews on certification standards from our sample, though we did use these separately as a source of additional insights (Andorfer & Liebe, 2012; Blackman & Rivera, 2011; Oya et al., 2018).

We further filtered our sample by screening titles and abstracts based on the following two selection criteria: (a) Does the article directly focus on certification standards for sustainability? (or were such standards only mentioned in passing); and (b) Does the article include a clear reference to the impact of certification standards? (or was impact discussed only sporadically or as a secondary concern). Only abstracts that complied with both requirements were included in our final sample. The first author of this article coordinated the overall screening process and coded all abstracts, while the other three co-authors each simultaneously coded a subsample of 40 abstracts. We then compared our samples to check for internal consistency and alignment regarding the inclusion of articles. We discussed any initial differences over which articles to include, ultimately agreeing on a final sample of 203 articles.

Phase 3: Thematic Analysis

In the third phase of our review process, we conducted a thematic analysis with the purpose of grouping any studies with overlapping perspectives on impact. Full-text screening was performed by all four authors, with each coding an equal portion of the sample. The entire data analysis process was iterative, based on a deep reading of the articles in our corpus. We initially started by coding the articles in terms of a number of descriptive characteristics, such as (a) the methods employed in an article (conceptual versus empirical and quantitative, qualitative, or mixed methods); (b) the theories or theoretical perspectives adopted by that article; (c) its disciplinary grounding (based on the journal in which the article appeared); (d) its temporal orientation (e.g., whether it focused on a single point in time, several temporal data points over time, or took the form of a longitudinal analysis); (e) its level of analysis (micro, meso, and macro); (f) the names of the certification standards it discussed (e.g., UTZ); (g) the commodity sector it related to (e.g., forestry); and finally (h) its focus on the stakeholder(s) impacted by the certification standard (i.e., the number of stakeholders and the particular stakeholder groups the article conceives of as being impacted by a certification standard).

After coding the articles based on these descriptive dimensions, we read and reviewed each article in depth to identify its approach to impact. We started by categorizing the articles according to whether they focused on outputs or outcomes, or both. Output here refers to the observable material or immaterial direct effects of certification standards (e.g., metrics on increase in income, increased enrollment in education or the sustainable use of land), while outcomes are understood as the change generated by and through the output, such as changes in general welfare, social capital, or landscape conservation (Mills-Scofield, 2012). We then started to look more deeply at the ways in which scholars have approached and operationalized such impacts. From this deep reading, we concluded that the type of logic scholars use in attributing an output or outcome to certification standards serves to reinforce the way in which their studies conceive what impact is and how it is achieved. In line with this appraisal, we then categorized all the studies in our sample according to the particular impact logic followed by the authors.

This resulted in two overarching categories across all of the reviewed studies: (a) those that followed a linear logic and (b) those that followed a more configurational impact logic. Where the underlying logic of a study was unclear, we coded the article as involving an unspecified causal inference. We labeled studies as following a linear logic when the author(s) applied an impact logic to explain how the adoption or implementation of a certification standard had directly affected one or more output or outcome variables. We

labeled studies as falling into the configurational logic category when the author(s) conceptualized multiple pathways by which impact can be realized, thereby taking account of interconnections between concepts or variables or considering multiple and sometimes different possible outcomes from the adoption of a certification standard. In addition to these two main categories, we also identified a subset of articles (8%) in which it proved difficult to clearly decipher a particular inferential logic and these were placed into the unspecified category (Hilson, 2014; Makita, 2018).

Overview of Research on the Impact of Certification Standards

We begin our findings section by providing a general overview of our data set in terms of some of the descriptive statistics related to the variables discussed above. Our sample includes 203 articles from 92 journals. This large number of different journal outlets highlights the fact that the overall body of work on certification standards has become spread out and rather scattered, a fact further reflected by the diverse range of disciplines that are covered by these journals. The largest share of articles in our data set were from the field of sustainability and environmental sciences (29%), followed by development studies (18%), management studies (18%), economics (16%), and political science (13%). We further found a small number of studies published in sociological and psychological outlets (6%) (Table 4 provides an overview of the journal outlets per discipline). Overall, scholarly interest in the impact of certification standards has been increasing over the last decade, although the precise number of articles published per year varies significantly (e.g., 31 studies were published in 2018 as compared to only nine studies in 2011; see Figure 1). Impact assessment is thus increasingly of relevance to academics studying certification standards, which may in part be due to the fact that many initiatives have matured, thereby making such impact assessments and their problematization necessary and timely.

From the entire body of work of studies reviewed, it is clear that scholars have adopted a range of different methods for assessing the impact of certification standards. The overwhelming majority of articles have studied impact empirically (86%) using quantitative methods (57%), qualitative methods (33%), or mixed-method approaches (11%). The large proportion of studies that focus on quantitative methods reflects the dominant use of surveys or large-scale panel data to gauge impact. At the same time, however, it is noteworthy that one third of the reviewed studies used qualitative methods, which

Table 4. Journals per Discipline.

Management Studies	Sustainability and Environmental Sciences	Economics
Academy of Management review	*Ambio*	*Agricultural Economics*
Agribusiness	*Biological Conservation*	*Applied Economic Perspectives and Policy*
Annals of Operations Research	*Ecology and Society*	*Applied Economics*
Business and Society	*Ecosystem Services*	*Australian Journal of Agriculture and Resource Economics*
Business Strategy and the Environment	*Environmental Planning A-economy and Space*	*Bulletin of Indonesian Economic Studies*
California Management Review	*Environmental Development and Sustainability*	*Ecological Economics*
Consumption Markets and Culture	*Environmental Research Letters*	*European Review of Agricultural Economics*
Corporate Social Responsibility and Environmental Management	*Environmental Science and Policy*	*Feminist Economics*
European Journal of Marketing	*Environmental Science and Pollution Research*	*Forest Policy and Economics*
International Journal of Advertising	*Environmental Values*	*International Journal of Production Economics*
International Journal of Life Cycle Assessment	*Global Environmental Change*	*Journal of Agricultural Economics*
International Journal of Logistics Management	*International Journal of Sustainable Development and World Ecology*	*Journal of Development Economics*
International Journal of Operations and Production Management	*Journal of Agriculture and Environmental Ethics*	*Journal of Economic Inequality*
Journal of Business Ethics	*Journal of Cleaner Production*	*Journal of Environmental Economics and Management*
Journal of Business Venturing	*Nature Sustainability*	*Journal of the Association of Environmental and Resource Economist*
Journal of Macromarketing	*Regional Environmental Change*	*Marine Resource Economics*
Journal of Management Studies	*Society and Natural Resources*	*Review of Economics and Statistics*
Journal of Retailing and Consumer Services	*Sustainability*	
Manufacturing and Service Operations Management	*Sustainable Development*	
Management Science	*Wileys Interdisciplinary Reviews—Climate Change*	
Organization and Environment		
Production and Operations Management		
Production Planning and Control		
Supply Chain Management		

(continued)

Table 4. (continued)

Political Science, Policy and Governance	Development Studies	Sociology and Psychology
Annals of the American Academy of Political and Social Science	*Canadian Journal of Development Studies*	*Agriculture and Human Values*
Australian Journal of Political Science	*Development and Change*	*Decision Sciences*
Development Policy Review	*Globalizations*	*Journal of Happiness Studies*
Food Policy	*Journal of Agrarian Change*	*Rural Sociology*
Global Environmental Politics	*Journal of Development Effectiveness*	*Sociological Perspectives*
Global Policy	*Journal of Development Studies*	*Sociologica Ruralis*
Journal of Environmental Policy and Planning	*Journal of Environment and Development*	
Journal of Public Policy and Planning	*Journal of International Development*	
Land Use Policy	*Journal of Peasant Studies*	
Law and Policy	*Third World Quarterly*	
Marine Policy	*World Development*	
New Political Economy		
Regulation and Governance		
Sustainability Accounting management and Policy Journal		

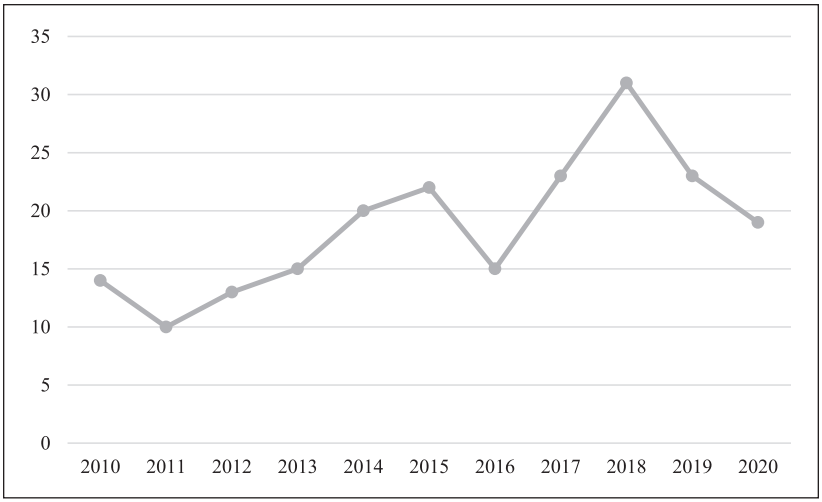

Figure 1. Number of publications (2010–2020).

in the context of our corpus usually involves case-based data as a way of providing more contextualized insights into micro-level settings.

In terms of scope, impact was mostly studied in the context of certification standards related to a limited number of commodity sectors, with research on the impact of coffee (27%), forestry (13%), and fishery (10%) certifications being the most prominent. Such a scope is not surprising, since most research has focused on a select set of certification standards, with Fairtrade (32%) being the most frequently studied. In addition, the bulk of the work to date has focused exclusively on MSIs (83%), perhaps reflecting the relative newness of SSIs.

Finally, we found that most of the impact-related work we reviewed had a decidedly micro focus, with half of the studies conducted in the context of one specific geographic location (50%) and the majority focused on one specific moment in time (66%). Indeed, the bulk of work on impact to date has comprised studies of specific communities at a given point in time, such as farmers in the Ugandan coffee sector (Latynskiy & Berger, 2017), fishers in the Filipino tuna fishery sector (Tolentino-Zondervan et al., 2016), and foresters in Slovakia (Paluš et al., 2018).

Assessing Impact

We synthesize extant findings on the impacts of certification standards by zooming in on the following two dominant approaches to impact assessment

that we identified in the literature: (a) a predominantly linear net effects approach and (b) an alternative configurational approach. The *linear approach* involves a form of hypothesizing by which researchers theoretically and methodologically interrelate certification standards with impact-related outputs and/or outcomes that are suggestive of a causal connection. This approach aligns with hypothesis-testing and regression-based methods that "parse social reality into fixed entities with variable qualities" (Abbott, 1992, p. 428) and that model the way in which these variables determine each other, depicting a "general linear reality" (Abbott, 1988). We found this linear approach to be central to those empirical studies conducted in field settings that make use of econometric techniques in a controlled and quasi-experimental manner to assess the effects or impact of a certification standard as a form of "intervention" in a local system (Simon, 1952).

Unlike linear approaches that focus more on bivariate relations (Furnari et al., 2020), scholars who adopt a *configurational approach* explicitly aim to account for "multifaceted interdependencies" between the adoption of certification standards and outputs and outcomes for different stakeholder groups. With this approach, researchers intentionally complicate matters by analyzing alternate configurations of interactions through which impacts can be realized. For example, scholars may anchor their assessments on a particular impact-related output or outcome and then search for the multiple ways in which such outputs or outcomes could be realized. Alternatively, they may start from the details and dynamics of a particular certification standard and then trace the various ways in which it could impact different stakeholder groups differently (Grabs et al., 2021). Both the linear as well as the configurational approach paint a mixed picture of certification standards' impact on relevant stakeholder groups. While there is some evidence for social and environmental gains, there is also plenty of evidence of factors that impede the creation of positive impact. We discuss the linear and configurational impact approaches as well as their most important findings in more detail below. See also Table 5 for a summary of the two approaches and their most relevant findings.

The Linear or Net Effects Approach to Assessing Impact

The linear net effect logic envisages an impact scenario that is quite limited in scope. Scholars following this logic tend to focus on a particular output or outcome for a specific stakeholder group, often measuring these outputs and outcomes by proxies. In this way, they aim to establish the direct effect of the adoption of a certification standard on the given output or outcome while

Table 5. Overview of the Main Impact Logics Discussing Certification Standards.

Logic	Linear approaches	Configurational approaches
Percent of articles out of overall dataset	• 70%	• 22%
Dominant disciplines	• Sustainability and environmental sciences (32%) • Economics (21%) • Management (19%) • Development studies (17%)	• Development studies (24%) • Political science (24%) • Sustainability and environmental sciences (18%)
Percent of empirical articles	• 89%	• 80%
Methods used within empirical articles	• Quantitative (75%) • Qualitative (13%) • Mixed (12%)	• Quantitative (11%) • Qualitative (83%) • Mixed (6%)
Temporal orientation	• Singe point in time (68%) • Several temporal data points over time (18%) • Longitudinal analysis (14%)	• Singe point in time (58%) • Several temporal data points over time (20%) • Longitudinal analysis (22%)
Level of analysis	• Micro (53%) • Meso (28%) • Macro (19%)	• Micro (38%) • Meso (27%) • Macro (36%)
Main stakeholder focus	• Farmers and producers (56%) • Local communities and environment (15%) • Corporations (14%) • Consumers (12%)	• Global value chains (60%) • Local market systems (40%)
Number of stakeholder groups taken into account	• One stakeholder group (75%) • Two stakeholder groups (23%)	• One stakeholder group (51%) • Two stakeholder groups (42%)

(continued)

Table 5. (continued)

Logic	Linear approaches	Configurational approaches
Focus of research	• Economic aspects of welfare and productivity impact, based on output measures (e.g., income and yield). • Social aspects of impact (e.g., inequality and working conditions; educational opportunities and distribution of benefits across communities). • Environmental aspects of impact (e.g., use of chemicals; resource conservation and biodiversity). • Indirect impact through increased green purchasing behavior and firm position (e.g., revenue, profit and prices).	• Combination of social, economic, and environmental outcomes in within particular geographical or commodity markets (e.g., local market opportunities; power distribution across supply chains; and public discourse).
Main findings	• Positive impacts: higher wages; more job satisfaction; better labor conditions; access to global markets, education and healthcare; the use of eco-friendly farming practices; increased land and biodiversity conservation; increased awareness of product quality and attributes and willingness to pay premium prices and enhanced marketing possibilities (e.g., firms' reputation). • Negative impacts: longer working hours; increased production costs; often unequal distribution of benefits and power; inflation of label meaning and consumer confusion.	• Positive impacts: potential for market transformation in terms of public norms; transparency of information and supply chain structures and collaborations. • Negative impacts: misalignment of standards' criteria with local socio-political context; perpetuated wealth power imbalances.

trying to control for any outputs or outcomes that would have transpired irrespective of that standard.

The majority of studies in our review (70%) adopted this linear type of approach to conceptualizing and studying the impact of certification standards. In general, the studies within this cluster are drawn from the fields of sustainability and environmental sciences (32%), economics (21%), and management studies (19%). The great majority are predominantly empirical (89%) and employ quantitative methods (75%), such as surveys, experiments, and econometric modeling on archival, panel or household, consumer or corporate data. Scholars following this type of impact logic often conduct their studies within one particular geographic context (53%) at a single point in time (68%), with the great majority assessing the impact of certification on one particular stakeholder group (75%).

To assess impact, scholars applying a linear impact logic often rely on standardized proxies. For example, in one particularly dominant strand in the literature that studies the impact of certification standards on local farmers and producers, scholars have assessed the economic impact of certificates on individual farmers and producers using data on hourly wages (Krumbiegel et al., 2018), household disposable income (Parvathi & Waibel, 2016), and overall farm profitability (Beuchelt & Zeller, 2011). Other studies have assessed the social impact on local communities. In a study of the impact of forest certifications, for example, Doremus (2019) has investigated the distribution of certification benefits in terms of increased material wealth in a local community by distinguishing between indigenous and nonindigenous households. Analyzing the welfare effects of organic and Fairtrade standards, Meemken and colleagues (2017) used panel data from coffee producers in Uganda to study the direct effects of these two certification standards on household expenditures, child education, and nutrition. Linear studies with an environmental focus have similarly assessed impact through such metrics related to gross tree cover loss (Garrett et al., 2016), pesticide usage (Ho et al., 2018), and carbon density (Kitayama et al., 2018).

Overall, these studies reveal a mixed picture in terms of the impact of certification standards. Scholars have, for instance, found evidence for positive economic and social outcomes in terms of enhanced farmer income, local community building, and access to education and healthcare (Akoyi et al., 2020; Chiputwa & Qaim, 2016; Krumbiegel et al., 2018; Mook & Overdevest, 2018) as well as for positive environmental outcomes through reduced chemical inputs and the adoption of environmentally friendly management practices in local contexts (Blackman & Naranjo, 2012; Dwivedi et al., 2018). However, such scholarship has also identified factors that can hinder a net-positive impact from certification standards. For instance,

research has found that complying with certification standards may result in longer working hours (Becchetti et al., 2012), increased production costs (Latynskiy & Berger, 2017), and unequal welfare improvements due to either hierarchies among farmers (Meemken et al., 2019) or to seasonal employment (Cramer et al., 2017). Studies have further shown that the targeted beneficiaries of certification standards are often located at the bottom of the socio-economic pyramid and that in spite of higher wages or product yields arising from these standards they are often still found to be living below the general poverty line (Beuchelt & Zeller, 2011) and dependent on fluctuating commodity market prices (Parvathi & Waibel, 2016). Some studies have further found that the benefits arising from certification standards may be distributed unequally across communities (Doremus, 2019), while others have reported a lack of evidence of any environmental progress at all in certain local contexts (Blackman et al., 2018).

In addition to the strong focus in this literature on impact effects for producers and farmers, our review identified a second prominent stream of literature in the linear tradition that focuses on the impact of certification standards on consumers and corporations. At the corporate level, scholars have studied correlations between the adoption of certification standards and a firm's stock price (Bouslah et al., 2010) or have conceptualized impact as being tantamount to changes in firm revenues and sales (Joo et al., 2010). Studies adopting a consumer perspective typically focus primarily on the direct effects of certification standards on consumer awareness, green purchasing behavior, and marketing performance. For example, Song and colleagues (2019) have measured the effects of eco-labels on consumer perceptions through a survey of product attributes and environmental attitudes and concerns. In another example, Campbell and colleagues (2015) conducted an experimental study to measure consumers' acceptance of price increases due to a product's fair trade credentials.

In this strand of scholarship, studies have generally found that certification standards can provide positive financial outcomes for corporations, for instance in terms of enhanced perceived firm legitimacy (Feng et al., 2020) or brand recognition (Joo et al., 2010) and can convince some consumers to pay a premium for ethically approved products (Fonner & Sylvia, 2015; Song et al., 2019). These outcomes have also been found to be moderated by other factors such as the business strategy around the adoption of certification (Tey et al., 2020), the cultural features of the corporation's home country (Orzes et al., 2017), product prices (Campbell et al., 2015; Michal et al., 2019), characteristics of the brand itself (Konopka et al., 2019), and consumer characteristics (Herédia-Colaço et al., 2019). Finally, a number of scholars have argued that certification standards are failing to convey to consumers what certification

entails and what standards have been adhered to (Goossens et al., 2017; Gutierrez & Thornton, 2014; Heyes et al., 2020). This information deficit, they argue, may pose significant challenges to the impact of certification standards on consumer decision-making as it complicates comparability across labels, as well as the ability to assess their true sustainability credentials.

In summary, the linear approach assumes a limited scenario whereby the adoption of a certification standard directly leads to consequent changes for stakeholders. In doing so, these studies paint a fragmented and mixed picture in terms of the actual impact generated by certification standards. On the one hand, studies show that different stakeholder groups such as farmers and local communities, benefit from the adoption of standards in some, but not all, cases. On the other hand, there is also ample evidence that certification standards may, in fact, yield negative effects too (e.g., a loss of labor productivity or an unequal distribution of benefits among actors). Methodologically, such scholarship follows and gives form to this logic by using controlled (quasi) experimental research designs and econometric techniques. Although this notion of a direct net effect may have a certain appeal, it also invariably highlights the limitations of this approach. The question that arises in this regard is whether the adoption of a certification standard can be presumed to be both necessary and sufficient for the output or outcome to be realized, especially against the background of other potential conditions and causes that may influence the livelihood and welfare of farmers or producers. One key limitation of this approach, then, is that it rests on rather strong assumptions, as is borne out by the fact that scholarship in this stream has invariably led to mixed and even conflicting findings about the impact of certification for particular stakeholder groups.

The Configurational Approach to Assessing Impact

A second group of studies in our review (22%) incorporated what we have termed a configurational approach to impact assessment that actively takes account of the multifaceted nature of the adoption of certification standards and their varied and wide-ranging impacts. Scholars following this more configurational approach consider impact as the result of a conjunction of multiple factors and conditions and of the interactions between these factors, aiming to identify "why or how multiple explanatory factors combine into configurations that bring about an outcome of interest" (Furnari et al., 2020, p. 9).

In general, most studies within this configurational cluster are from the fields of development studies (24%), political science (24%), and sustainability and environmental sciences (18%), with the great majority being empirically oriented (80%) and applying mainly qualitative methods (83%).

Scholars adopting this approach have typically studied the impacts of certification standards either in a particular local context (38%) or across geographical settings (36%). Most of these studies have been conducted at a single point in time (58%), with a few studies taking a longitudinal approach (22%).

One strand of literature we identified in our review has assessed the impact of certification standards on the way business is done within a particular industry or socio-political context (McCarthy, 2012; Swartz et al., 2017; Vos & Boelens, 2014). Studies within this cluster generally assess impact as a combination of economic, environmental, and social outcomes that certification standards may produce in a particular context. For example, Swartz and colleagues (2017) have identified connections between local socio-cultural consumer attitudes and the evolution of three eco-labels in the Japanese seafood market, while Vos and Boelens (2014) have studied the effect of certification standards on local water communities, finding that such standards can reinforce existing political and market power inequalities between companies, local communities, and governments due to a misalignment between the certifications and the socio-political contexts in which they operate. In another example, Barrios and colleagues (2016) identified four types of longitudinal certification mechanisms that enabled certified coffee growers in supporting the transition of the Colombian war economy into a peace economy (namely empowerment, communication, community building, and regulation).

A second strand of studies within the configurational approach looks at the role of global value chains (Dare, 2018; Pichler, 2013; Vagneron & Roquigny, 2011), where the adoption of certification standards is likely to impact how these chains function in terms of vertical trade patterns, global power structures, and the overall market environment. Accordingly, a number of these studies have researched how the introduction of certification schemes affects power relations in global value chains (Dare, 2018), investigating how standards have influenced stakeholder's authority (e.g., in terms of their ability to coerce others) and how they have reinforced dependency relationships. Impact is understood here not as an effect on (or for) a single stakeholder but is conceived rather as an emergent outcome of how an entire system of stakeholders within a single global value chain interact with each other. In a study on forest certification and local politics, for example, Dare (2018) approaches certification standards as complex and dynamic boundary-spanning regimes, theorizing how different certification standards enhance competing "boundaries" within global value chains, with embedded local interests and contested institutional claims exerting conflicting pressures that can undermine the overall effectiveness of these standards. In another example, Pichler (2013) uses a critical state and hegemony theory perspective to argue how the palm oil value chain includes exclusion mechanisms that enhance power

asymmetries and marginalize those actors at the start of the value chain, the farmers and producers of palm oil.

A third and final cluster of studies in the configurational category also focuses on global value chains but further highlights how certifications can change the very nature of commodity markets (Naylor, 2018; Ponte, 2012; Raynolds, 2014). Most scholarly contributions in this stream are concerned with understanding how the introduction of standards endeavors to combine an efficiency-driven market rationality with a value-driven rationality because some standards ultimately challenge which norms should govern market relationships. Introducing sustainability certificates into global value chains can therefore be seen as an attempt to regulate supply and demand for commodities that align with social and environmental norms. In one example of applying such an approach, Shorette (2014) has studied the temporal dimension of public norms and argued that sustainable conditions of production are only valuable in a specific normative-regulatory context that is tied to more general shifting global norms related to equality, human rights, and environmental protection.

While configurational studies, like linear studies, also paint a mixed picture of the impacts of certification standards, they do so in a much more holistic and comparative manner than those following a linear approach—for example, by actively comparing and contrasting the extent to which the introduction of certification standards alters existing dependency relationships in global value chains. Reflecting on the positive findings, studies that have focused on social relations in specific local contexts have found that standards create positive effects for communities and strengthen social ties. For instance, Naylor (2018, p. 1041) finds in her research of coffee farmers in Chiapas, Mexico that "there appears to be a socially-minded community economy between coffee growing cooperatives and coffee roasters as they attempt to build relationships and work together to think about how to build communities that can live well."

On the contrary, several other studies emphasize that once we focus on global value chains, the introduction of certification standards do not fundamentally challenge global power structures. Vagneron and Roquigny's (2011) comparative study of the adoption of a fair trade standard in a conventional banana value chain is particularly insightful in this context. The authors show that producers actually gain less, relatively speaking, when participating in fair trade value chains; for while producers can obtain a higher price for their products, they also "remain pressurized between powerful European retailers and large Dominican plantations" (Vagneron & Roquigny, 2011, p. 336). One problem identified by this and other configurational studies is that certified global value chains generally include the same stakeholders as conventional

value chains and despite their efforts to promote more direct relations between these stakeholders, such as farmers and consumers for example, certification standards do not seem to fundamentally challenge global power relations. Indeed, certification standards may even serve to perpetuate existing power structures within communities due to a failure to take sufficient account of local dependencies and underlying social structures (McCarthy, 2012; Vos & Boelens, 2014). In their study of sustainability standards and the "water question" for example, Vos and Boelens (2014) found that existing political and market power inequalities between companies, local communities, and governments were actually reinforced due to a misalignment between the certification standards and the local socio-political context.

Findings depicting such mixed pictures of the impacts of certification standards are typical of configurational approaches that consider the extent to which conditions and outcomes are interconnected or may even be nested in one another. While the configurational approach is thus more holistic than the linear view in comparing and contrasting different theoretical perspectives and different assumptions about the impacts of certification in context (Grabs et al., 2021), this approach also has its limits. The focus of studies in this stream often remains comparative in nature, for example, providing multi-sided perspectives but not always tracing in detail the ways in which complex interactions and interdependencies produce different impacts for different stakeholder groups over time.

Discussion

The Dominant Logic of Impact Assessment

In summary, the vast majority of studies in our review approach impact through a linear lens, hypothesizing and testing the direct effects of a certification standard on a particular output or outcome for a specific stakeholder group. Indeed, this focus on singular cause-effect relationships is widespread in the business and society domain (Barnett et al., 2020; Blackman & Rivera, 2011). In a recent review of 6,254 articles published on the impact of corporate social responsibility (CSR) since 1973, for example, Barnett and colleagues (2020) have argued for the need to adopt a logic framework that is consistent with such a linear approach so as to "enable studies to better determine causation, rather than just identify correlation" (Barnett et al., 2020, p. 939), further proposing that the business and society field as a whole should be reconceived as a "science of design" in which researchers manipulate one (or more) CSR factors as putative causes and eliminate any confounding explanations through controls in the design of their experiments or through

statistical randomization (randomized controlled trials). As our review has shown, this field has long been dominated by research designs based on linear hypothesis-testing that attempt to identify causal effects as the direct results of the introduction of certification standards. Given the limited and mixed findings of these studies, we caution against the exclusive adoption of such a limited research approach insofar as it leaves the field in general at risk of developing a rather narrow and ultimately inconclusive view of the impact of certification standards.

The Limitations of a Linear Impact Logic

While studies applying a linear impact logic have a number of methodological strengths and can yield in-depth insights into the local effects of certification standards, such logic also comes with significant limitations when deployed as the overarching structure for charting the causal processes and relationships that lead to the societal impact of certification standards.

One such limitation is that these studies often limit themselves to identifying the local effects of a single output proxy as opposed to charting the broader impacts of certification standards across spatial and temporal scales. This schematic reduction consists of settling on an "extremely simple representation of causality" (Talmy, 1988, p. 92) that marks few distinctions, abstracts away any particulars and contingencies of the context, and brackets a simple autonomous scenario without any consideration of causal precursors or consequences beyond the simplified scene depicted. Construing causality through such a delimited schema obscures the complexities of local contexts, confining these contexts within a specific format focused on charting the direct effects of a certification standard on a selected output or outcome while blanking out other conditions and effects. This screening off almost inevitably increases the likelihood of creating an ambiguous and incomplete picture of impact when comparing the findings of studies across contexts and impact measures. The reliance of such linear studies on quantifiable proxies and thus on measuring specific quantifiable outputs, such as wage differentials, rather than the wider socio-economic outcomes of adopting a standard, such as net welfare increases for farmers in a particular setting or local system, further prevents these studies from assessing the broader and more systemic impacts of certification across contexts and across different social and ecological dimensions. Certification standards yield different outcomes for farmers and producers in the Ethiopian coffee sector compared to Indian tea farmers, for example, and these mixed findings emphasize the importance of accounting for such contextual differences and of adopting a more complex approach when assessing impacts and how such impacts are generated in practice.

A second limitation of adopting a linear logic-based approach is that in presuming a predetermined role for certain causes as root causes, the approach (see Abbott, 1992; Fiss, 2011) rests, as mentioned, on strong assumptions about both the necessity and sufficiency of a particular cause, which we argue is inherently problematic in view of the complexity of many business and society domains. Such strong assumptions result in research prone both to underestimate the multifarious ways in which impact can actually be achieved and to overrate the importance of the direct contribution of a presumed single cause such as the adoption of a certification standard as compared to other important conditions or processes in a particular setting. The economic proxies used to assess impacts in the linear studies above, for instance, may in this sense be better interpreted quite differently, for example, as indicating incremental steps in improving economic welfare. For example, certification standards can increase output variables such as product prices, which can lead to increased hourly wages which in turn enhances output variables such as overall household income. As another example, a reduction in environmental degradation is not likely to be exclusively an effect of compliance with certification standards but rather the result of an interplay of various conditions and factors—including education levels, land usage, and local environmental resilience—that have broader spill-over effects and trade-offs across the social-ecological system.

Seeking to address the one-sidedness of this approach, a recent review by Grabs and colleagues (2021) has proposed an alternative template-based analysis that involves researchers initially anchoring their studies in certain outcomes but then systematically varying their vantage points (across stakeholders, over time, and across levels) and theoretical assumptions to better capture the multifaceted nature of impact. Other studies have similarly critiqued the econometric focus of linear impact assessments on bivariate relations (Furnari et al., 2020), thus making the case for more processual and configurational approaches to account for the interdependencies between processes and the various impacts these give rise to at different levels of analysis. Such approaches are better able to capture, both in theory and practice, the multifaceted nature of the impact of certification standards and the complex causal processes through which such impacts are created. Given that certification standards are embedded in wider political, economic, and social systems, future studies may thus view their impacts in a broader context in which stakeholders engage interactively, creating a multitude of causal pathways throughout the system, rather than viewing such impacts as the direct effect of standards on specific aspects such as environmental degradation or general welfare. In our configurational cluster, for example, we encountered studies emphasizing the importance of contextual nuance in assessing impact,

such as the effect of socio-cultural characteristics (Swartz et al., 2017) or socio-political underlying structures (Vos & Boelens, 2014) on the potential outputs and outcomes of certification standards.

While a configurational approach goes some way toward addressing these issues, our review also flagged the limitation of this approach in terms of its scope. One limitation of configurational studies, for example, is that they often include variables exclusively referring to one type of outcome such as the economic outcomes of adopting certification standards. In addition, few of the studies we have reviewed incorporate temporal aspects in conceptualizing impact. As a consequence, the potential long-term interaction effects between social, economic, and ecological outcomes—intended as well as unintended—remain under-explored along with any other unidentified factors that have the potential to alter local and global systems as a whole.

A Systems-Based Impact Logic

To complement the linear and configurational approaches to assessing the impact of certification standards, there is a need for studies that follow a different inferential logic for thinking about impact. Whereas existing extant impact research tends to focus on univariate singular causal relationships, we encourage the use of a more holistic perspective on interdependencies between outputs and outcomes and a whole range of possible and realized effects within and between social and ecological systems. In other words, we advocate a logic for thinking about impact that incorporates the value of "think[ing] holistically [. . .] to understand causally relevant conditions as intersections of forces and events" (Ragin, 2009, p. 109).

We argue that such a more holistic approach, one that is capable of accounting for interactions across multiple levels and processes, enables scholars to identify research questions that move beyond single linear relationships, singular and local output proxies, and presumed root causes (such as between certification standards and farmer's hourly wages). Such a system-based approach would better enable scholars to ascertain the actual processes and mechanisms that drive the impact of a certification standard across systems and for the different stakeholder groups involved and, thus, not only complements specific knowledge on local net effects but also the ongoing development of more complex accounts of impact. Resting on a variety of different methods and analytical tools—ranging from system-related scoping tools that determine the social and ecological elements to be included into a study (Sitas et al., 2021) to network analysis that unpacks how elements are connected with each other and whether dependencies, trade-offs, and synergies exist between them (Maciejewski & Baggio, 2021)—such approaches

may support scholars studying the impact of certification standards to conceptualize impact from a more holistic perspective and across spatial and temporal scales (Grewatsch et al., 2021). To give this further shape, we turn next to recent advances in social-ecological research on impact that are based on complex systems accounts of causal mechanisms (Rocha et al., 2018).

The starting premise of this complex systems perspective is that the impacts generated by certification standards relate to and are embedded within social-ecological systems. We can understand such systems as including "social (human) and ecological (biophysical) subsystems in mutual interactions" (Harrington et al., 2010, p. 2773). If we accept that the impacts of certifications are manifested as changes in outcomes, therefore, we first need to acknowledge that these outcomes are embedded in the specific social-ecological systems in which the target beneficiaries operate. These systems consist of the interdependent behaviors of both human and nonhuman entities, such as farmers and the communities around them as well as other species and fauna that share the same environment. The entities in larger social-ecological systems are organized into the following three subsystems: (a) social systems (e.g., cooperatives, families, communities); (b) ecological systems (e.g., plantations with bushes and water holes); and (c) economic systems (e.g., markets). From our survey of the literature, we find there has been an insufficient integration of such systems thinking within empirical research designs aimed at assessing the impacts of certification.

These three subsystems are presented in Figure 2 as part of a basic schematic framework for studying the impact of certification standards from a systems perspective. This framework assumes that interventions resulting from the adoption of a certification standard can cause changes in outputs and outcomes related to all three subsystems. We have included a number of exemplary outputs and outcomes within the three subsystems. The systems approach has a number of characteristics that make for a more realistic assessment of the impact of certification standards, above all because this approach assumes that impact is the result of interaction effects related to entities both *within* and *among* the different subsystems. Accordingly, this approach focuses on the connections between entities to identify causal dependencies and interactions. These dependencies may involve one-way dependencies (e.g., a "domino cascade," with one system triggering a chain reaction in others) as well as two-way interactions between subsystems (e.g., a "hidden feedback cascade") (Rocha et al., 2018). In other words, the primary focus is on the underlying interactions within and between subsystems and the tipping points, spillover effects, and feedback loops that create extended processes and impacts. Even if such cascading effects cannot always be directly observed in an empirical context, they leave hypothetical

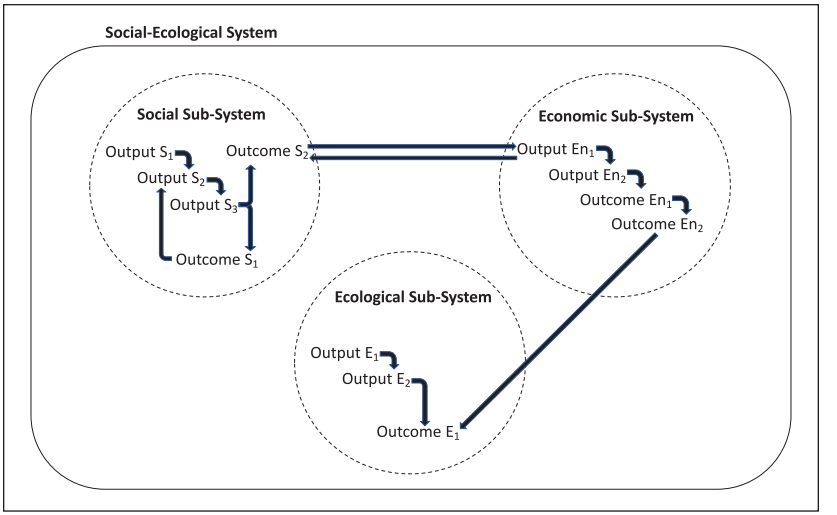

Figure 2. A systems-based framework to discuss the impact of certification standards.

traces of interacting processes and effects that can be studied and verified (Hedström & Swedberg, 1998).

Such a systems view on impact has become widespread within the natural sciences (Steffen et al., 2015) precisely because it allows scholars to see the bigger picture (Polasky et al., 2020) rather than over-focusing on single cause–effect relations. We believe that future scholarly work on the impact of certification standards can make fruitful use of such insights, particularly in light of our finding that current scholarly work typically provides insights into single causal relations but does so in ways that obscure broader context. For instance, while we have ample evidence that certifications can improve farmers' yields and hence also increase their incomes (Mitiku et al., 2017), we know little about how these effects interact with other outcomes such as better education, improved healthcare for families, or improved environmental protection.

Figure 3 shows some examples of systems relations that are relevant to consider when studying the impact of certifications, in this case within the social-ecological system of farming. The diagram illustrates the interrelated effects of adopting standards such as organic certifications that typically require farmers to use less fertilizers and how their doing so creates a cascade effect because it affects nitrogen pollution and hence also the availability of

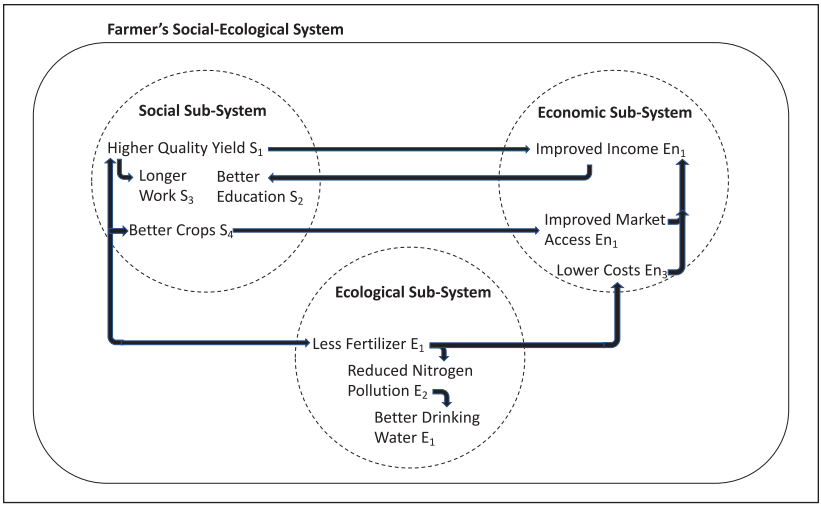

Figure 3. Exemplary adoption of the systems-based framework.

clean drinking water. Using less fertilizer also has economic impacts, moreover, because it reduces farmers' expenditure on chemicals (Liu et al., 2018). While reducing fertilizer use often yields better quality crops that improve income and market access, however, this farming method also requires longer working hours from farmers (Vanderhaegen et al., 2018).

This example of the complex effects that arise from what might seem the simple intervention of reducing chemical fertilizers shows that in discussing the impact of certification standards, we need to consider multiple cascading effects, cross-system effects, and trade-offs between subsystems. As an example of such trade-offs, while certification standards can improve both education (through higher family income) and health (through better water quality), they can also offset some of these gains due to decreased labor productivity and the longer working hours required by more environmentally sustainable farming methods. Focusing on such systems dynamics when assessing the impacts of certification standards should enable researchers to attain a more accurate understanding of the social-ecological nature of the overall system in which the work of certification standards is embedded. In this respect, our proposed approach accords with Seager's (2008, p. 447) statement that "the locus of study in sustainability science is on *the interaction between human and natural systems*" (emphasis in original).

The schematic visualizations we provide in Figure 2 and 3 are simplified representations of complex systems in which certification standards are

embedded. With this, we do not mean to suggest that the predominant linear impact logic be completely replaced by a systems-based logic but rather propose that using different impact logics in a complementary manner can help develop a more comprehensive understanding of the impacts of certification schemes and indeed of the impacts of any other social initiatives with potentially complex and multifaceted ramifications.

Implications for Further Research

As part of our cross-disciplinary review of the scholarship on the impact of certification standards, we have problematized the linear view that is currently predominant in impact assessment studies and questioned whether sufficient progress can be made in this field if such a linear approach continues to be widely applied. In accordance with the nature of a problematizing review (Alvesson & Sandberg, 2020), we have problematized the assumptions behind this prevailing linear view and based on this critique have offered an alternative theoretical perspective to orient and guide further research.

Our review has a number of implications for research on the impact of certification standards in the business and society domain. While acknowledging that linear studies can provide certain insights into the effects of certification for local stakeholders in a particular context, we have shown that such studies fall short of capturing the full impact of adopting a certification standard. A first implication of this finding for future impact studies, therefore, is that research needs to be more equally balanced between linear, configurational, and complex systems-based approaches to impact assessment. Combining such approaches in research has clear benefits in that it enables researchers to zoom in on specific effects while zooming out to allow for the detection of interdependencies and outcomes within and across socio-ecological systems. Applying such a combination of approaches, of course, requires that researchers recognize the value of these different approaches and of using them alongside one another. In the absence of such recognition, researchers continuing to apply a linear approach (Barnett et al., 2020) risk repeatedly coming up against the inevitable limits of such thinking and thus of underrepresenting the true extent of the impacts of a certain certification standard. The exercise of balancing a system-wide perspective with a delineated research scope, however, remains a complex task. Below we provide several suggestions on how scholars may implement this alternative research avenue through their theorizing and choices concerning their research designs.

Future research could usefully apply the concepts of complex systems thinking more directly in their studies by, for example, examining the key transition or tipping points that trigger impacts across socio-ecological

systems and sustain these impacts over longer periods of time. Future research could also intentionally work with a broader scope by assessing impact on (the interactions between) multiple stakeholder groups. Rather than zooming in on individual farmers, producers, or consumers, for example, studies may also include other, fewer-researched, actor groups that are part of that system, such as regulatory intermediaries (e.g., auditors and certification bodies) or different supply chain tiers (e.g., manufacturers, distributors, and retailers). In this way, researchers may already be able to capture cascading processes and potential interdependencies occurring between and across different actors.

In addition to the implications for impact research in terms of where, when, and for whom it is measured, a second implication of the systems-based approach we propose is that it may guide individual researchers to conceptualize and examine differently how impact occurs; that is, in terms of the causal pathways through which the impacts of certification may be realized. Instead of applying a net effect logic that conceives of certification standards acting as a direct force or as an impetus for the studied impact, researchers can instead use the proposed framework and associated concepts as a heuristic when designing and conducting studies on impact. This may entail, among other things, adopting a more probabilistic mind-set in figuring out which other conditions need to be in place for impact to be realized and sustained. Adopting such a mind-set may further require that researchers expand their zone of vision beyond the local effects of a standard for particular stakeholders to take account of other interdependencies and interactions—both in terms of direct and indirect observable outcomes, but also broader socio-cultural, political, economic, and structural forms of impact—across contexts and over time (Galaz et al., 2018) to trace impacts along more extended causal pathways.

The upshot of including a systems perspective to complement existing studies, is that researchers of impact will also have to accommodate this deeper, more complex, conceptualization in their methodologies and research designs. Due to the prevalent linear logic in extant work on impact, studies tend to be limited in scope and often overlook the importance of including a broader time horizon, as we have observed in our review. The issue of scope may be addressed by incorporating more cross-sectional data or conducting more complex analyses (such as network analysis, loop analysis or other scoping tools) involving multiple stakeholder groups, as we have suggested above. Paying more attention to the temporal dimensions of how impact may be generated, in turn, is also something that research may factor into their research designs so as to allow for a more nuanced picture of when and how impact may yield longer-lasting outcomes as opposed to impact being equated to measurements at single points in time. Process and longitudinal research

designs and methods may be used for this purpose, with data and analyses linked to an overall systems perspective.

In addition to these broader implications for the business and society field, we also wish to highlight several important research directions for studies on certification standards. Our review has shown how past and current research predominantly focuses on MSIs like Fairtrade that operate in a limited number of sectors such as forestry or coffee. We have also identified a dearth of research on SSIs as a corollary of this predominant focus, in spite of the fact that SSIs are becoming more increasingly prominent in practice. Future research could thus usefully explore different types of certification standards that are currently emerging and that have so far been under-researched. It may be fruitful, for instance, to compare impact assessments of MSIs and SSIs more explicitly and directly, and also to contrast how nonprofit standard setters and for-profit standard setters influence how certification schemes affect outputs and outcomes. Although some scholars have already begun to explore such dynamics, including the role of professionalism in the dynamics of standard-setting (Henriksen & Seabrooke, 2016; Henriksen, 2015) and issues of competition and legitimization between MSIs and more business-driven programs (Fransen, 2012; Fransen et al., 2019), we argue that there is a pressing need for further research into the effects of the emergence of SSIs and other emergent standards on the sector as a whole.

To sum up, this review has offered insights into the multifaceted nature of the impact of certification standards, identifying and problematizing existing approaches to assessing such impacts and developing an alternative framework to guide further research. As we have argued, to better understand the behavior of certification standard systems, and thus their impact, scholars may need to extend their ideas regarding the definition, scope, and scale of such impact. With impact-related questions increasingly taking center-stage in research and society at large, we hope that the alternative vantage point we offer here may contribute to ongoing research and debate over the next decade in *Business & Society* and beyond.

Acknowledgements

The authors thank *Business & Society* co-editor Kathleen Rehbein and three anonymous reviewers for their constructive feedback, insightful comments, and helpful guidance.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD


Sylke F. Jellema https://orcid.org/0000-0002-2052-6099

Supplemental Material

Supplemental material for this article is available online.

List of Works Reviewed*

*This list includes only those studies that have been referred to in this article. See the online appendix for the full list of works reviewed

Akoyi, K. T., Mitiku, F., & Maertens, M. (2020). Private sustainability standards and child schooling in the African coffee sector. *Journal of Cleaner Production*, *264*, Article 121713. https://doi.org/10.1016/j.jclepro.2020.121713

Barrios, A., de Valck, K., Shultz, C. J., Sibai, O., Husemann, K. C., Maxwell-Smith, M., & Luedicke, M. K. (2016). Marketing as a means to transformative social conflict resolution: Lessons from transitioning war economies and the Colombian coffee marketing system. *Journal of Public Policy & Marketing*, *35*, 185–197.

Becchetti, L., Conzo, P., & Gianfreda, G. (2012). Market access, organic farming and productivity: The effects of Fair Trade affiliation on Thai farmer producer groups. *Australian Journal of Agricultural and Resource Economics*, *56*, 117–140.

Beuchelt, T. D., & Zeller, M. (2011). Profits and poverty: Certification's troubled link for Nicaragua's organic and Fairtrade coffee producers. *Ecological Economics*, *70*, 1316–1324.

Blackman, A., Goff, L., & Planter, M. R. (2018). Does eco-certification stem tropical deforestation? Forest Stewardship Council certification in Mexico. *Journal of Environmental Economics and Management*, *89*, 306–333.

Blackman, A., & Naranjo, M. A. (2012). Does eco-certification have environmental benefits? Organic coffee in Costa Rica. *Ecological Economics*, *83*, 58–66.

Bouslah, K., M'Zali, B., Turcotte, M., & Kooli, M. (2010). The impact of forest certification on firm financial performance in Canada and the US. *Journal of Business Ethics*, *96*, 551–572.

Campbell, C. L., Heinrich, D., & Schoenmüller, V. (2015). Consumers' reaction to fair trade motivated price increases. *Journal of Retailing and Consumer Services*, *24*, 79–84.

Chiputwa, B., & Qaim, M. (2016). Sustainability standards, gender, and nutrition among smallholder farmers in Uganda. *The Journal of Development Studies*, *52*, 1241–1257.

Cramer, C., Johnston, D., Mueller, B., Oya, C., & Sender, J. (2017). Fairtrade and labour markets in Ethiopia and Uganda. *The Journal of Development Studies*, *53*, 841–856.

Dare, L. (2018). From global forests to local politics: Unwrapping the boundaries within forest certification. *Australian Journal of Political Science*, *53*, 529–547.

Doremus, J. (2019). Unintended impacts from forest certification: Evidence from indigenous Aka households in Congo. *Ecological Economics*, *166*(1), Article 106378. https://doi.org/10.1016/j.ecolecon.2019.106378

Dwivedi, P., Tumpach, C., Cook, C., & Izlar, B. (2018). Effects of the Sustainable Forestry Initiative fiber sourcing standard on the average implementation rate of forestry best management practices in Georgia, United States. *Forest Policy and Economics*, *97*, 51–58.

Feng, Y., Lai, K., & Zhu, Q. (2020). Legitimacy in operations: How sustainability certification announcements by Chinese listed enterprises influence their market value? *International Journal of Production Economics*, *224*, Article 107563.

Fonner, R., & Sylvia, G. (2015). Willingness to pay for multiple seafood labels in a niche market. *Marine Resource Economics*, *30*, 51–70.

Garrett, R. D., Carlson, K. M., Rueda, X., & Noojipady, P. (2016). Assessing the potential additionality of certification by the Round Table on Responsible Soybeans and the Roundtable on Sustainable Palm Oil. *Environmental Research Letters*, *11*(4), Article 045003. https://doi.org/10.1088/1748-9326/11/4/045003

Goossens, Y., Berrens, P., Charleer, L., Coremans, P., Houbrechts, M., Vervaet, C., de Tavernier, J., & Geeraerd, A. (2017). Qualitative assessment of eco-labels on fresh produce in Flanders (Belgium) highlights a potential intention–performance gap for the supply chain. *Journal of Cleaner Production*, *140*, 986–995.

Gutierrez, A., & Thornton, T. F. (2014). Can consumers understand sustainability through seafood eco-labels? A US and UK case study. *Sustainability*, *6*, 8195–8217.

Herédia-Colaço, V., do Vale, R. C., & Villas-Boas, S. B. (2019). Does fair trade breed contempt? A cross-country examination on the moderating role of brand familiarity and consumer expertise on product evaluation. *Journal of Business Ethics*, *156*, 737–758.

Heyes, A., Kapur, S., Kennedy, P. W., Martin, S., & Maxwell, J. W. (2020). But what does it mean? Competition between products carrying alternative green labels when consumers are active acquirers of information. *Journal of the Association of Environmental and Resource Economists*, *7*, 243–277.

Hilson, G. (2014). "Constructing" ethical mineral supply chains in Sub-Saharan Africa: The case of Malawian fair trade rubies. *Development and Change*, *45*, 53–78.

Ho, T. Q., Hoang, V., Wilson, C., & Nguyen, T. (2018). Eco-efficiency analysis of sustainability-certified coffee production in Vietnam. *Journal of Cleaner Production*, *183*, 251–260.

Joo, S., Min, H., Kwon, I. G., & Kwon, H. (2010). Comparative efficiencies of specialty coffee retailers from the perspectives of socially responsible global sourcing. *The International Journal of Logistics Management*, *21*, 490–509.

Kitayama, K., Fujiki, S., Aoyagi, R., Imai, N., Sugau, J., Titin, J., Nilus, R., Lagan, P., Sawada, Y., Ong, R., Kugan, F., & Mannan, S. (2018). Biodiversity observation

for land and ecosystem health (BOLEH): A robust method to evaluate the management impacts on the bundle of carbon and biodiversity ecosystem services in tropical production forests. *Sustainability*, *10*(11), Article 4224.

Konopka, R., Wright, M. J., Avis, M., & Feetham, P. M. (2019). If you think about it more, do you want it more? The case of Fairtrade. *European Journal of Marketing*, *53*, 2556–2581.

Krumbiegel, K., Maertens, M., & Wollni, M. (2018). The role of Fairtrade certification for wages and job satisfaction of plantation workers. *World Development*, *102*, 195–212.

Latynskiy, E., & Berger, T. (2017). Assessing the income effects of group certification for smallholder coffee farmers: Agent-based simulation in Uganda. *Journal of Agricultural Economics*, *68*, 727–748.

Liu, R., Gao, Z., Yan, G., & Ma, H. (2018). Why should we protect the interests of "Green food" certified product growers? Evidence from kiwifruit production in China. *Sustainability*, *10*, 1–14.

Makita, R. (2018). Application of fair trade certification for wild plants: Lessons from a FairWild project in India. *International Journal of Sustainable Development & World Ecology*, *25*(7), 619–629.

McCarthy, J. F. (2012). Certifying in contested spaces: Private regulation in Indonesian forestry and palm oil. *Third World Quarterly*, *33*, 1871–1888.

Meemken, E., Sellare, J., Kouame, C. N., & Qaim, M. (2019). Effects of Fairtrade on the livelihoods of poor rural workers. *Nature Sustainability*, *2*, 635–642.

Meemken, E., Spielman, D. J., & Qaim, M. (2017). Trading off nutrition and education? A panel data analysis of the dissimilar welfare effects of organic and fairtrade standards. *Food Policy*, *71*, 74–85.

Michal, J., Březina, D., Šafařík, D., Kupčák, V., Sujová, A., & Fialová, J. (2019). Analysis of socioeconomic impacts of the FSC and PEFC certification systems on business entities and consumers. *Sustainability*, *11*(15), Article 4122. https://doi.org/10.3390/su11154122

Mitiku, F., De Mey, Y., Nyssen, J., & Maertens, M. (2017). Do private sustainability standards contribute to income growth and poverty alleviation? A comparison of different coffee certification schemes in Ethiopia. *Sustainability*, *9*(2), Article 246. https://doi.org/10.3390/su9020246

Mook, A., & Overdevest, C. (2018). Does Fairtrade certification meet producers' expectations related to participating in mainstream markets? An analysis of advertised benefits and perceived impact. *Sustainable Development*, *26*, 269–280.

Naylor, L. (2018). Fair trade coffee exchanges and community economies. *Environment and Planning A: Economy and Space*, *50*, 1027–1046.

Orzes, G., Jia, F., Sartor, M., & Nassimbeni, G. (2017). Performance implications of SA8000 certification. *International Journal of Operations & Production Management*, *37*, 1625–1653.

Paluš, H., Parobek, J., Šulek, R., Lichý, J., & Šálka, J. (2018). Understanding Sustainable Forest Management certification in Slovakia: Forest owners' perception of expectations, benefits and problems. *Sustainability*, *10*, 1–17.

Parvathi, P., & Waibel, H. (2016). Organic agriculture and fair trade: A happy marriage? A case study of certified smallholder black pepper farmers in India. *World Development*, *77*, 206–220.

Pichler, M. (2013). "People, planet & profit": Consumer-oriented hegemony and power relations in palm oil and agrofuel certification. *The Journal of Environment & Development*, *22*, 370–390.

Ponte, S. (2012). The Marine Stewardship Council (MSC) and the making of a market for "sustainable fish." *Journal of Agrarian Change*, *12*, 300–315.

Raynolds, L. T. (2014). Fairtrade, certification, and labor: Global and local tensions in improving conditions for agricultural workers. *Agriculture and Human Values*, *31*, 499–511.

Shorette, K. (2014). Nongovernmental regulation and construction of value in global markets: The rise of fair trade, 1961–2006. *Sociological Perspectives*, *57*, 526–547.

Song, Y., Qin, Z., & Yuan, Q. (2019). The impact of eco-label on the young Chinese generation: The mediation role of environmental awareness and product attributes in green purchase. *Sustainability*, *11*(4), Article 973. https://doi.org/10.3390/su11040973

Swartz, W., Schiller, L., Sumaila, U. R., & Ota, Y. (2017). Searching for market-based sustainability pathways: Challenges and opportunities for seafood certification programs in Japan. *Marine Policy*, *76*, 185–191.

Tey, Y. S., Brindal, M., Darham, S., Sidique, S. F. A., & Djama, M. (2020). Early mover advantage in Roundtable on Sustainable Palm Oil certification: A panel evidence of plantation companies. *Journal of Cleaner Production*, *252*, Article 119775. https://doi.org/10.1016/j.jclepro.2019.119775

Tolentino-Zondervan, F., Berentsen, P., Bush, S., Idemne, J., Babaran, R., & Lansink, A. O. (2016). Comparison of private incentive mechanisms for improving sustainability of Filipino tuna fisheries. *World Development*, *83*, 264–279.

Vagneron, I., & Roquigny, S. (2011). Value distribution in conventional, organic and fair trade banana chains in the Dominican Republic. *Canadian Journal of Development Studies*, *32*, 324–338.

Vanderhaegen, K., Akoyi, K. T., Dekoninck, W., Jocqué, R., Muys, B., Verbist, B., & Maertens, M. (2018). Do private coffee standards "walk the talk" in improving socio-economic and environmental sustainability? *Global Environmental Change*, *51*, 1–9.

Vos, J., & Boelens, R. (2014). Sustainability standards and the water question. *Development and Change*, *45*, 205–230.

Additional References

Abbott, A. (1988). Transcending general linear reality. *Sociological Theory*, *6*(2), 169–186.

Abbott, A. (1992). From causes to events: Notes on narrative positivism. *Sociological Methods and Research*, *20*(4), 428–455.

Alvesson, M., & Sandberg, J. (2020). The problematizing review: A counterpoint to Elsbach and Van Knippenberg's argument for integrative reviews. *Journal of Management Studies*, *57*(6), 1290–1304.

Andorfer, V. A., & Liebe, U. (2012). Research on fair trade consumption—A review. *Journal of Business Ethics*, *106*(4), 415–435.

Arton, A., Leiman, A., Petrokofsky, G., Toonen, H., Neat, F., & Longo, C. S. (2020). What do we know about the impacts of the Marine Stewardship Council seafood ecolabelling program? A systematic map protocol. *Environmental Evidence*, *7*(1), 1–8.

Auld, G., Gulbrandsen, L. H., & McDermott, C. L. (2008). Certification schemes and the impacts on forests and forestry. *Annual Review of Environment and Resources*, *33*, 187–211.

Barnett, M. L., Henriques, I., & Husted, B. W. (2020). Beyond good intentions: Designing CSR initiatives for greater social impact. *Journal of Management*, *46*(6), 937–964.

Blackman, A., & Rivera, J. (2011). Producer-level benefits of sustainability certification. *Conservation Biology*, *25*(6), 1176–1185.

Clark, C., Rosenzweig, W., Long, D., & Olsen, S. (2004). *Double bottom line project report: Assessing social impact in double bottom line ventures* (Working Paper 13). Center for Responsible Business, University of California, Berkeley.

Dammert, A. C., & Mohan, S. (2015). A survey of the economics of fair trade. *Journal of Economic Surveys*, *29*(5), 855–868.

de Bakker, F. G. A., Rasche, A., & Ponte, S. (2019). Multi-stakeholder initiatives on sustainability: A cross-disciplinary review and research agenda for business ethics. *Business Ethics Quarterly*, *29*(3), 343–383.

Dragusanu, R., Giovannucci, D., & Nunn, N. (2014). The economics of fair trade. *Journal of Economic Perspectives*, *28*(3), 217–236.

Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, *54*(2), 393–420.

Fransen, L. (2012). Multi-stakeholder governance and voluntary programme interactions: Legitimation politics in the institutional design of corporate social responsibility. *Socio-Economic Review*, *10*(1), 163–192.

Fransen, L., Kolk, A., & Rivera-Santos, M. (2019). The multiplicity of international corporate social responsibility standards. *Multinational Business Review*, *27*(4), 397–426.

Furnari, S., Crilly, D., Misangyi, V. F., Greckhamer, T., Fiss, P. C., & Aguilera, R. (2020). Capturing causal complexity: Heuristics for configurational theorizing. *Academy of Management Review*, *46*(4), 778–799.

Galaz, V., Crona, B., Dauriach, A., Jouffray, J. B., Österblom, H., & Fichtner, J. (2018). Tax havens and global environmental degradation. *Nature Ecology & Evolution*, *2*(9), 1352–1357.

Gilbert, D. U., Rasche, A., & Waddock, S. (2011). Accountability in a global economy: The emergence of international accountability standards. *Business Ethics Quarterly*, *21*(1), 23–44.

Grabs, J., Auld, G., & Cashore, B. (2021). Private regulation, public policy, and the perils of adverse ontological selection. *Regulation & Governance*, *15*, 1183–1208.

Greenpeace International. (2021). *Destruction: Certified*. https://www.greenpeace.org/static/planet4-international-stateless/2021/04/b1e486be-greenpeace-international-report-destruction-certified_finaloptimised.pdf

Grewatsch, S., Kennedy, S., & Bansal, P. (2021). Tackling wicked problems in strategic management with systems thinking. *Strategic Organization*. https://doi.org/10.1177/14761270211038635

Grieco, C., Michelini, L., & Iasevoli, G. (2015). Measuring value creation in social enterprises: A cluster analysis of social impact assessment models. *Nonprofit and Voluntary Sector Quarterly*, *44*(6), 1173–1193.

Harrington, R., Anton, C., Dawson, T. P., de Bello, F., Feld, C. K., Haslett, J. R., Kluvánkova-Oravská, T., Kontogianni, A., Lavorel, S., Luck, G. W., Rounsevell, M. D. A., Samways, M. J., Settele, J., Skourtos, M., Spangenberg, J. H., Vandewalle, M., Zobel, M., & Harrison, P. A. (2010). Ecosystem services and biodiversity conservation: Concepts and a glossary. *Biodiversity and Conservation*, *19*(10), 2773–2790.

Hedström, P., & Swedberg, R. (1998). Social mechanisms: An introductory essay. In P. Hedström & R. Swedberg (Eds.), *Social mechanisms: An analytical approach to social theory* (pp. 1–31). Cambridge University Press.

Henriksen, L. F. (2015). The global network of biofuel sustainability standard-setters. *Environmental Politics*, *24*(1), 115–137.

Henriksen, L. F., & Seabrooke, L. (2016). Transnational organizing: Issue professionals in environmental sustainability networks. *Organization*, *23*(5), 722–741.

Impact Management Project. (2021). *Impact management norms*. https://impactmanagementproject.com/impact-management/impact-management-norms/#anchor2

International Association for Impact Assessment. (2015). *Social impact assessment: Guidance for assessing and managing the social impacts of projects*. https://www.iaia.org/uploads/pdf/SIA_Guidance_Document_IAIA.pdf

Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2010). *Handbook on impact evaluation: Quantitative methods and practices*. World Bank.

Maciejewski, K., & Baggio, J. (2021). Network analysis. In R. Biggs, A. de Vos, R. Preiser, H. Clements, K. Maciejewski, & M. Schlüter (Eds.), *The Routledge handbook of research methods for social-ecological systems* (pp. 321–331). Routledge.

Mills-Scofield, D. (2012, November 26). It's not just semantics: Managing outcomes vs. outputs. *Harvard Business Review*, *26*. https://hbr.org/2012/11/its-not-just-semantics-managing-outcomes.

MSI Integrity. (2020, July). *Not fit-for-purpose: The grand experiment of multi-stakeholder initiatives in corporate accountability, human rights and global governance*. https://www.msi-integrity.org/not-fit-for-purpose/

Nesadurai, H. E. (2013). Food security, the palm oil–land conflict nexus, and sustainability: A governance role for a private multi-stakeholder regime like the RSPO? *The Pacific Review*, *26*(5), 505–529.

Oya, C., Schaefer, F., & Skalidou, D. (2018). The effectiveness of agricultural certification in developing countries: A systematic review. *World Development*, *112*, 282–312.

Polasky, S., Crépin, A.-S., Biggs, R., Carpenter, S. R., Folke, C., Peterson, G., Scheffer, M., Barrett, S., Daily, G., Ehrlich, P., Howard, R. B., Hugher, T., Levin, S. A., Shogren, J. F., Troell, M., Walker, B., & Xepapadeas, A. (2020). Corridors of clarity: Four principles to overcome uncertainty paralysis in the Anthropocene. *BioScience*, *70*(12), 1139–1144.

Prieto-Sandoval, V., Alfaro, J. A., Mejía-Villa, A., & Ormazabal, M. (2016). ECO-labels as a multidimensional research topic: Trends and opportunities. *Journal of Cleaner Production*, *135*, 806–818.

Ragin, C. C. (2009). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago Press.

Rasche, A. (2010). Collaborative governance 2.0. *Corporate Governance: International Journal of Business in Society*, *10*(4), 500–511.

Rasche, A., Waddock, S., & McIntosh, M. (2013). The United Nations Global Compact: Retrospect and prospect. *Business & Society*, *52*(1), 6–30.

Reinecke, J., Manning, S., & Von Hagen, O. (2012). The emergence of a standards market: Multiplicity of sustainability standards in the global coffee industry. *Organization Studies*, *33*(5–6), 791–814.

Rocha, J. C., Peterson, G., Bodin, Ö., & Levin, S. (2018). Cascading regime shifts within and across scales. *Science*, *362*(6421), 1379–1383.

Ruggeri, G., Orsi, L., & Corsi, S. (2019). A bibliometric analysis of the scientific literature on fairtrade labelling. *International Journal of Consumer Studies*, *43*(2), 134–152.

Sartor, M., Orzes, G., Di Mauro, C., Ebrahimpour, M., & Nassimbeni, G. (2016). The SA8000 social certification standard: Literature review and theory-based research agenda. *International Journal of Production Economics*, *175*, 164–181.

Seager, T. P. (2008). The sustainability spectrum and the sciences of sustainability. *Business Strategy and the Environment*, *17*(7), 444–453.

Simon, H. A. (1952). On the definition of the causal relation. *Journal of Philosophy*, *49*(16), 517–528.

Sitas, N., Ryan, P., & Schultz, L. (2021). Systems Scoping. In R. Biggs, A. de Vos, R. Preiser, H. Clements, K. Maciejewski, & M. Schlüter (Eds), *The Routledge handbook of research methods for social-ecological systems* (pp. 83–94). Routledge.

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., de Vries, W., de Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., & Sörlin, S. (2015). Sustainability. planetary boundaries: Guiding human development on a changing planet. *Science*, *347*(6223), 736–746.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*(1), 49–100.

Tröster, R., & Hiete, M. (2018). Success of voluntary sustainability certification schemes: A comprehensive review. *Journal of Cleaner Production*, *196*, 1034–1043.

Waddock, S. (2008). Building a new institutional infrastructure for corporate responsibility. *Academy of Management Perspectives*, *22*(3), 87–108.

Author Biographies

Sylke F. Jellema is a PhD candidate at the department of Business-Society Management at the Rotterdam School of Management, Erasmus University. Her main research interests include organizational purpose and identity issues in impact-driven organizations, as well as the certification industry in particular.

Mirjam D. Werner is associate professor at the department of Business-Society Management at the Rotterdam School of Management, Erasmus University. Her research focuses mainly on social movements, citizen activism, social enterprises and framing processes.

Andreas Rasche is a professor of Business in Society at the Copenhagen Business School's (CBS) Center for Sustainability. He also acts as the associate dean for the CBS Full-Time MBA Program. His research examines the intersection of global governance, corporate sustainability, and the political role of firms. More information is available at http://www.arasche.com.

Joep Cornelissen is a professor in Corporate Communication at Rotterdam School of Management, Erasmus University and a Chair in Strategy and Organization (part-time) at the University of Liverpool. He researches processes of innovation, communication, and change and has a general interest in questions of scientific reasoning and theory development in management and organization theory. He is currently the Editor-in-Chief of *Organization Theory*, an online, open access journal publishing theoretical work on organizations, organizing, and the organized.