Technology and Innovation Management at Hamburg University of Technology

Working Paper

STANDARD IMPLEMENTATION TRAJECTORIES FOR SUSTAINABLE PRODUCT DESIGN: A CONFIGURATIONAL APPROACH

Armand Smits, Viktoria Drabe, Cornelius Herstatt

September 2016
Working Paper No. 95

Hamburg University of Technology (TUHH)

Institute for Technology and Innovation Management

Am Schwarzenberg-Campus 4
D-21073 Hamburg

Tel.: +49 40 42878 3775
Fax: +49 40 42878 2867

Email: a.smits@fm.ru.nl
viktoria.drabe@tuhh.de
Web : www.tuhh.de/tim

Abstract

While sustainability issues increasingly gain importance for new product design, they also further complicate the NPD process. In many cases it is hard to exactly measure the socio-environmental impact of new products, and sustainability targets may conflict with other ones. Innovators can aim to manage these challenges by turning to voluntary sustainability standards (VSS), like the practices and certificates that come with the EU Ecolabel, Greenguard or Cradle to Cradle standards. VSS are predefined rules, procedures, and methods for common and voluntary use and focus on social and environmental performance. It is proposed that the local implementation of these general standards from outside the organization will likely lead to a variety of firm-specific implementation trajectories, ultimately leading to different levels of VSS implementation extensiveness across firms. This variety that is not sufficiently addressed in extant research, is researched in the current study. Using organizational learning as theoretical lens this study investigates configuration(s) of factors, including the embeddedness of the relationship between the focal firm and standard specific organizations that drive VSS implementation extensiveness. In doing so, it uses the configurational research approach fuzzy set Qualitative Comparative Analysis (fsQCA). Empirically the study draws on qualitative and quantitative data from an international collection of firms that implemented the Cradle to Cradle standard. The study shows that VSS are multifaceted and that configurations that consistently drive VSS certification extensiveness differ from the ones that drive VSS practice implementation extensiveness. Additionally, it is found that configurations that consistently lead to the absence of high implementation extensiveness do not simply mirror the ones for high implementation extensiveness but have unique properties. Finally the study illustrates that similar levels of implementation extensiveness can result from multiple distinct configurations. The study mainly contributes to extant research on sustainable product design and how to integrate general principles of sustainable design into the NPD process.

Keywords: Sustainable product design, voluntary sustainability standards, fsQCA, implementation, Cradle to Cradle

1 Introduction

Spurred by drivers such as regulatory compliance, competitive advantage, stakeholder pressure, and environmental awareness, product innovators increasingly take environmental and social considerations into account (e.g. Markham and Lee, 2013; Varadarajan, 2015; Schiederig, Tietze, and Herstatt, 2012; Luchs, Swan, and Creusen, 2016). Adding this perspective to more traditional areas of attention, such as product quality and customer requirements, further complicates the new product development (NPD) process (Berchicci and Bodewes, 2005). For instance, the socio-environmental impact of new products is hard to fully understand and precisely measure (Wijen, 2014), and sustainability targets may conflict with other ones (Chen, 2001).

With the aim to deal with these challenges innovators may adopt voluntary sustainability standards (VSS), which are predefined rules, procedures, and methods for common and voluntary use and a focus on social and environmental performance (Delmas and Young, 2009; Brunsson, Rasche, and Seidl, 2012; Rubik, Scheer, and Iraldo, 2008). VSS are developed and maintained by a variety of standard setters such as consultants, non-governmental organizations, policy makers, and industry associations. Examples are the procedures that come with the EU Ecolabel,

Greenguard, Cradle to Cradle, or SMART standards. VSS offer product designers a sense of direction through design guidelines and benchmarks (Goggin, 1994; Rubik, Scheer, and Iraldo, 2008; Eppinger, 2011), and can be used as coordination mechanism when innovating with external stakeholders (Wagner, 2008; Kolk, 2005). Furthermore, complying with standards often is rewarded with so-called "positive sanctions" like certifications and eco-labels that signal that new products are designed in a more social-environmentally friendly way than uncertified offerings of competitors (King and Toffel, 2009) and therefore support the strategic positioning of new products (Rubik, Scheer, and Iraldo, 2008; Bratt et al., 2011). This last feature further differentiates VSS from more traditional voluntary NPD standards such as Stage-Gate (Cooper, 2008; O'Connor, 1994) and Lean Product Development (Karlsson and Åhlström, 1996) were certification is absent or less prominent.

While the implementation of VSS provides firms with various benefits, there are also significant difficulties (e.g. Horne, 2009, Rasche, 2015; Könnölä and Unruh, 2007). For instance, the opaque and evolving nature of the sustainability concept and the low barriers of entry of the VSS market have resulted in a significant increase in VSS over the last decades (Marx and Wouters, 2015). This proliferation of standards that often partly compete with each other, confuses product designers and customers and has increased skepticism (Delmas, Nairn-Birch, and Balzarova, 2013). Moreover, monitoring and auditing mechanisms that come with standard certification can turn out incomplete, unreliable, or inconsistent (Fransen and Kolk, 2007; O'Rourke, 2006). This may lead to consumers distrusting selected VSS and the associated companies and new products.

Taken together, the ambiguous nature of VSS, combined with the great variety of motivations of organizations to pursue sustainable product design, raises questions about how individual firms engage with VSS for this purpose. Prior studies have shown that firm level implementation of general practices, such as VSS, often lead to locally translated versions of the practices with different levels of implementation extensiveness, or degrees of implementation, across firms (Ansari, Fiss, and Zajac, 2010). While research on VSS has started to recognize variations in implementation extensiveness (e.g. Boiral, 2007; Yin and Schmeidler, 2009; Qi et al., 2012; Chappin et al., 2015; Bowler, Castka, and Balzarova, 2016), its drivers are understudied (Heras-Saizarbitoria and Boiral, 2013). Current studies have mainly focused on motivations for VSS adoption and have found that efficiency related motivations lead to more extensive implementation than legitimacy related ones or responding to social pressures. Motivations for adoption, though, have mostly been studied in isolation, and have yet not sufficiently been combined with other factors into more holistic configurations that can reveal and explain more nuanced firm-level implementation trajectories (Ivanova, Gray, and Sinha, 2014; Heras-Saizarbitoria and Boiral, 2013). Moreover, studies on VSS implementation extensiveness have mainly focused on the ISO14001 standard, which is only loosely coupled with NPD and product design.

This research is motivated by the abovementioned gap in the literature. More specifically, we address the following research question: What configuration(s) of factors drive VSS implementation extensiveness in the context of sustainable NPD? The framework guiding this theory elaboration research is rooted in organizational learning theory because studies have referred to standard implementation as a learning process (e.g. Naveh, Meilich, and Marcus, 2006; Aravind, 2012). Empirically, we focus on the Cradle to Cradle standard, which promotes closed-loop product design and therefore is closely associated with NPD (Bakker et al., 2010; Eppinger, 2011; Luchs, Swan, and Creusen,

2016). Our research follows earlier recommendations to conceptualize drivers of standard implementation as systemically interdependent and as configurations as opposed to isolated variables (e.g. Klein and Sorra, 1996; Ivanova, Gray, and Sinha, 2014). In line with this aim we apply fuzzy set Qualitative Comparative Analysis (fsQCA) as research approach because it is specifically suited to deal with complex interdependencies of variables underlying organizational outcomes. Although quite new to the field, fsQCA is increasingly used in management research, including studies on innovation management (see Kan et al., 2015 for an overview).

Our study mainly extends literature on sustainable design and NPD by further explaining implementation processes of VSS to manage sustainable design and innovation (see also Luchs, Swan, and Creusen, 2016). More specifically we show that VSS are multifaceted and that configurations of factors that drive VSS certification extensiveness differ from the ones that drive VSS practice implementation extensiveness. Moreover, we identify that configurations leading to the absence of high implementation extensiveness do not simply mirror the ones of high implementation extensiveness but have unique properties. Finally we show that similar levels of implementation extensiveness can result from multiple distinct configurations. We provide standard setters with relevant new knowledge to increase implementation extensiveness and thereby increase the societal impact of VSS. NPD managers and product designers can use our findings to better craft VSS implementation trajectories and integrate principles of sustainable design into the NPD process.

2 Theoretical background

Sustainability issues are increasingly gaining importance for product innovators (Markham and Lee, 2013). To design new products with increased socio-environmental benefits firms can use externally developed voluntary sustainability standards (VSS) (Rubik, Scheer, and Iraldo, 2008; Dangelico, Pontrandolfo, and Pujari, 2013, Goggin, 1994). These standards are voluntary in the sense that their adoption and use is not stipulated by any law or regulation. Instead, their regulative capacity is rooted in their perceived legitimacy and usefulness, and the potential pressure executed by external stakeholders (Helms and Webb, 2014).

VSS can regulate practices or outcomes. Some standards focus on practices without explicitly specifying outcomes, such as the ISO14001 standard. Others, such as FSC certified paper, may prescribe certain outcomes, new product features in our case, without detailing the practices to get there. However, practices are increasingly linked to outcomes and output measures can serve as guidelines and benchmarks for practices, which blurs the distinction (Rasche, 2015). VSS may vary in their governance mechanisms. Their rules can have different levels of detail, and monitoring can range from no monitoring at all to third party monitoring together with awarding and sanctioning in the case of compliance or non-compliance (Rasche, De Bakker, and Moon, 2013). This research focuses on standards with relatively elaborate and stringent governance mechanisms such as third party monitoring and a certificate or label to reward firms that comply with the standard. It is assumed that specifically these standards will have a significant impact on NPD processes because firms only have limited latitude to adapt these standards to their existing practices, and the associated labels impact new product positioning. Moreover, these types of standards are becoming increasingly important for governing sustainability (Delmas, Nairn-Birch, and Balzarova, 2013; Marx and Wouters, 2015).

The implementation of VSS as learning process

Nord and Tucker (1987: p.9) define the implementation of standards and other

administrative innovations[1] as "the 'pay off' stage of the [administrative] innovating process; the innovation is put in place, and the process of embedding it in the organization becomes a central activity." In process models of innovation diffusion, implementation is often positioned in-between adoption and routinization. Although implementation is often treated as binary (Staw and Epstein, 2000; Heras-Saizarbitoria and Boiral, 2013), the organization has or has not implemented the standard, increasingly studies emphasize heterogeneity in implementation extensiveness (e.g. Ansari, Fiss, and Zajac, 2010; Qi et al., 2012; Boiral, 2007; Chappin et al., 2015; Bowler, Castka, and Balzarova, 2016). Here, implementation extensiveness refers to the scale or degree of implementation, being low or high. In the context of this research an organization may, for instance, use a standard in a single NPD project (low level of implementation extensiveness) while another organization uses the standard for multiple projects in the firm's innovation portfolio and significantly integrates it in firm-level policies (high level of implementation extensiveness).

Prior studies have explicitly conceptualized the implementation of standards and other administrative innovations as a learning process (Naveh, Meilich, and Marcus, 2006; Aravind, 2012; Gaba and Dokko, 2015). Implementation is associated with a complex process in which intentions and motivations are based on assumptions that may not fully materialize due to inherent internal and contextual uncertainties. Organizational experiences result from earlier actions and inform further activities. Learning takes place if initial actions are modified based on reflection (Edmondson, Bohmer, and Pisano, 2001). Organizational learning as theoretical lens neatly suits the relatively uncertain NPD process because firms may experiment with VSS on a project by project basis with room for reflection in between.

Using organizational learning as theoretical lens, this research first develops an initial theoretical model to guide empirical analyses. The model relates to initial actions (motivations for adoption) and later reflection ((dis)satisfaction with implementation) and takes into account two additional factors connected to organizational learning processes (Figure 1). The selection of factors to include in the model was also based on recommended factor/case ratios for fsQCA analysis (Fiss, 2011; Marx and Dusa, 2011).

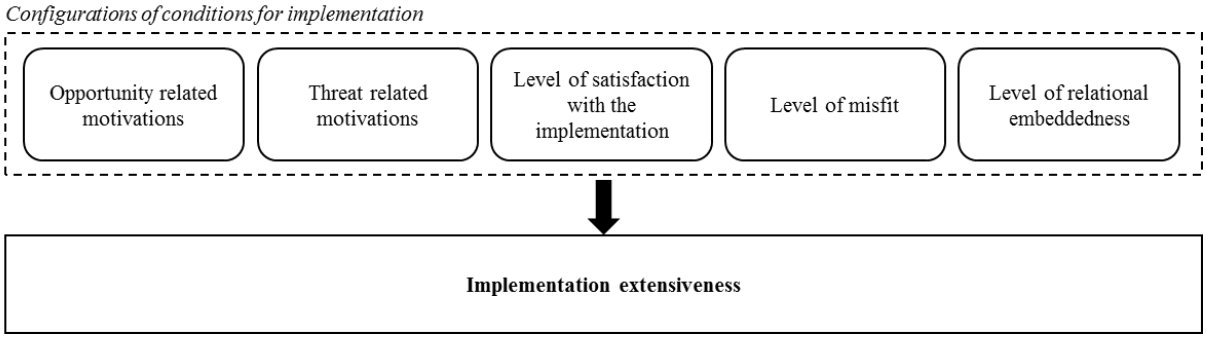

Figure 1: Theoretical model

[1] Administrative innovations refer to innovations in the managerial procedures, administrative processes, and rules of the organization (Daft, 1978) and include the implementation of standards (Naveh, Meilich, and Marcus, 2006). Administrative innovations are often contrasted with technical or technological innovations.

Motivations for adoption

Motivations for adoption have been regarded as important determinants of the nature of implementation trajectories. Based on neo-institutional arguments (Dimaggio and Powell, 1983) several studies that focus on the implementation of ISO14001 distinguish between efficiency related motivations and legitimacy related motivations (e.g. Boiral, 2007; Yin and Schmeidler, 2009; Qi et al., 2012). Efficiency related motivations are about the improvement of the performance of internal practices. Legitimacy related motivations are reactions to social pressures exerted by powerful agents from the institutional environment to adopt a standard. It appears that ISO14001is more superficially implemented when legitimacy related motivations are stronger than efficiency related ones (Yin and Schmeidler, 2009; Qi et al., 2012). At the extreme, firms decouple implementation from changes in practices and only adopt for the benefits of association rather than for performance guidance (Boiral, 2007; Delmas and Montes-Sancho, 2010). Other research has shown that differentiation in motivations for adoption and the associated implementation is more accurately captured by distinguishing opportunity related motivations from threat related ones (e.g. Kennedy and Fiss, 2009; Crilly, Zollo, and Hansen, 2012). When firms see a standard as an opportunity and are motivated by achieving gains they implement more extensively than organizations that do not see benefits. Conversely, firms that perceive the standard as threat and are motivated by avoiding losses will only implement the standard to the extent that the perceived threat is neutralized, a finding that is consistent with the threat-rigidity hypothesis from psychology (Staw, Sandelands, and Dutton, 1981). Hence for this study it is anticipated that, for a certain implementation timeframe, 'motivations for adoption', and specifically the extent to which these motivations are opportunity or threat related can be an important firm-level condition for implementation extensiveness of VSS in NPD.

Level of satisfaction with the implementation

At the heart of learning processes lies the re-evaluation of assumptions based on new and emerging information (Edmondson, Bohmer, and Pisano, 2001). Actual implementation of VSS, and standards in general, can lead to dissatisfactory results, which, subsequently, may trigger change. Unexpected better ways to execute actions may emerge and potential miscalculations made at the start of the process are identified and corrected. Change in perspective can be rooted in developments inside or outside of the organization. For instance, standards and standardization are typically associated with stability and sameness. A closer look, however, reveals that VSS are in fact quite dynamic, and that their governance, content, and legitimacy can change over time (Helms and Webb, 2012, Rasche, 2015).

Following Naveh, Meilich, and Marcus (2006) and Aravind (2012) 'adaptation in use' and 'change catalysis' are identified as two distinct implementation learning modes. Adaptation in use means refinement or further exploitation (March, 1991) of the implementation activities in order to better capture implementation benefits and increase the fit between the VSS and the organization to improve satisfaction. Change catalysis relates to distant search during implementation (March, 1991). Relative to adaptation in use, change catalysis is a deeper form of learning that challenges basic assumptions, norms, and objectives. Initial implementation experiences can act as organizational resources (Lewis and Seibold, 1993: 333) that serve as springboard for new plans, ideas, and activities. While organizations are likely to maintain the intended implementation trajectory after adaptation in use, though in refined form, the effects of change catalysis can have more drastic consequences. Specifically based on dissatisfaction with the

implementation, organizational actors may question earlier motivations and eventually 'jump ship' (Greve, 1995) in the midst of implementation. They may halt the implementation trajectory, replace a standard with another one (Abrahamson and Fairchild, 1999), or abandon (Terlaak and Gong, 2008) VSS for NPD altogether. For instance, Ecover, a Belgian-based producer of cleaning products, decided to abandon their use of the EU Ecolabel after noticing sloppy auditing, which was claimed to hurt the company's superior environmental reputation (Delmas, Nairn-Birch, and Balzarova, 2013). Therefore it is anticipated that, for a certain implementation timeframe, the level of experience related satisfaction with VSS and their implementation can be an important firm-level condition for implementation extensiveness of VSS in NPD.

Level of misfit

Standards with relatively stringent governance mechanisms do not lend themselves easily to multiple interpretations (Ansari, Fiss, and Zajac, 2010). Consequently, firms often have to adapt their products, resources, and practices to the standard and have little room to adjust the content of the standard to local circumstances. In this case it is expected that the level of misfit is important in implementation related learning processes. Level of misfit refers to the extent to which the characteristics and requirements of the standard are incompatible with the resources, practices, external ties, and products already in use by adopters (Ansari, Fiss, and Zajac, 2010; Simpson, Power, and Klassen, 2012). In cases of high levels of misfit, standard implementation will most likely require higher levels of product and process adaptations and therefore come with more radical forms of NPD. In contrast, implementation in cases of low misfit will likely result in more incremental NPD.

It is anticipated that the implementation challenges that come with the level of misfit can influence implementation extensiveness beyond the motivations for adoption and satisfaction with the implementation process. In a certain implementation timeframe, similar motivations and implementation efforts could lead to less extensive implementation for companies that experience a high misfit in comparison to the ones that experience a lower misfit because the gap to be closed for the former companies is much larger. A more indirect way by which higher levels of misfit could potentially influence implementation extensiveness is through escalation of commitment (Staw, 1976; Schmidt and Calantone, 1998): Dissatisfaction with initial implementation results does not lead to putting implementation on hold or VSS abandonment but to more extensive implementation in hope of a turnaround. Research has shown that escalation of commitment will be stronger with higher previous resource expenditures (Sleesman et al., 2012), which seem particularly required in implementation trajectories with higher levels of misfit. Thus, it is anticipated that, for a certain implementation timeframe, the level of misfit can be an important firm-level condition for implementation extensiveness of VSS in NPD.

Level of relational embeddedness

VSS are seldom created in isolation (Tamm Hallström and Boström, 2010). To secure legitimacy many standardization organizations have been designed as multi-stakeholder associations, often including firms that will implement the standards later on (Blind and Mangelsdorf, 2016; Helms, Oliver, and Webb, 2012). Standard related organizations frequently have a stake in diffusing and promoting standards as well, and therefore also interact with firms during further stages of the standard's life cycle (Bowler, Castka, and Balzarova, 2015). Hence, there are multiple opportunities for relationships to develop between implementing firms and standard specific organizations. These relationships can have different levels of 'embeddedness': the degree of reciprocity and closeness between

partners in the relationship (Rindfleisch and Moorman, 2001: p. 3).

The level of relational embeddedness could influence implementation extensiveness in multiple ways. Standards are dynamic and collectively developed and revised over time (Brunsson, Rasche, and Seidl, 2012). Through close relationships with standard related organizations, firms can consciously or unconsciously enforce own content in standard development (Bowler, Castka, and Balzarova, 2015). In this manner close collaborations may reduce standard–organization misfits over time and thereby facilitate more extensive implementation. Close and reciprocal relationships can also enable knowledge spillovers between different standard stakeholders (Blind and Mangelsdorf, 2016) that accelerate implementation related learning. Furthermore, based on the alliance literature (e.g. Adobor, 2006; Bruner and Spekman, 1998), it can be inferred that close relationships with standard specific organizations may motivate managers to hang on to the VSS and further implementation of the standard due to investments made in building the relationship or (public) identification with the standard and the standard related organizations. This may occur despite relatively unsatisfactory learning outcomes that otherwise would have led to standard replacement or abandonment and therefore superficial implementation. Therefore it is expected that, for a certain implementation timeframe, the level of relational embeddedness among the implementer and standard specific organizations can be an important firm-level condition for implementation extensiveness of VSS in NPD.

3 Methods

Empirical setting

To thoroughly examine and explain different levels of implementation extensiveness of VSS, which are administrative innovations for the implementing organizations, this research follows the recommendations of Klein and Sorra (1996) to focus on a single administrative innovation. We opted for the Cradle to Cradle (C2C) VSS (see also Braungart, McDonough, and Bollinger, 2006 for more background information on C2C) for three main reasons: (1) As product design standard, C2C is directly linked to NPD; (2) The C2C standard has a global focus and is not restricted by country or industry borders, which contributes to a broader applicability of the research findings; (3) By gaining its current structure in 2005, C2C strikes the attractive balance between sufficient history to study implementation trajectories on the one hand and recency on the other hand.

C2C, a phrase originally coined by Walter Stahel in the 1980s, emphasizes the importance of closed materials loops to protect and enrich ecosystems and focuses on the recycling of materials and products without quality deterioration, a process named ‘upcycling’. C2C argues that the traditional eco-efficiency perspective, which assumes a ‘cradle-to-grave’ flow of materials and focuses on the reduction of waste or negative externalities in different phases of the product lifecycle, should be replaced by an eco-effectiveness perspective. This latter perspective is about creating closed loops in which materials maintain their quality and can be continuously recycled or reused, eliminating the concept of ‘waste’ altogether. As such it has significant impact on product design (Bakker et al., 2010; Eppinger, 2011). C2C was mainly developed by German chemist Michael Braungart and American architect William McDonough. As the standard’s ‘gurus’ Braungart and McDonough promote C2C through consultancy assignments, seminars, articles, and books. From 2005 onwards, the standard has been complemented by a proprietary product certification system with different criteria, like material health and material reutilization, and five progressive

achievement levels: basic, bronze, silver, gold, platinum.

Research approach

We aim to elaborate theory on sustainable design and NPD by exploring how VSS are implemented in the context of NPD and what configuration(s) of factors drive VSS implementation extensiveness. This purpose is well served by applying fuzzy set Qualitative Comparative Analysis (fsQCA). This research approach uses Boolean algebra to systematically derive configurations of factors, or 'conditions' in fsQCA language, that lead to a certain outcome (Ragin, 2008). fsQCA is well suited for theory elaboration (e.g. Crilly, Zollo, and Hansen, 2012; Munoz and Dimov, 2015) and suffers less from computational and interpretation limitations that are typically associated with modeling higher-order interactions using traditional regression techniques (Fiss, 2007). fsQCA unravels three elements of causal complexity: conjunction, equifinality, and asymmetry (Ragin, 2008; Schneider and Wagemann, 2012). Conjunction refers to single conditions not impacting the outcome in isolation from each other. Equifinality allows for multiple sets of conditions to be linked to a single outcome. Asymmetry means that the set of conditions that leads to an outcome can be qualitatively different than the set that leads to the absence of the outcome and therefore each situation requires separate theoretical and empirical consideration. While an in-depth explanation of the method is beyond the scope of this article, we provide further background information in Appendix 1.

Data collection

Data collection progressed in three main stages. First, certified firms were identified through consulting the website of the Cradle to Cradle Certified Product Registry. This registry is maintained by the standard setter and publicly available on the internet (www.c2ccertified.org). It continuously documents the products that hold C2C certificates and their associated companies. Utilizing an internet archive tool (www.waybackmachine.nl) we were able to develop a longitudinal overview of the certification behavior of all firms that had one or more certifications for the period 2005-2015. In parallel, interviews with 11 firms were conducted, transcribed, and analyzed in order to assess firm motivations for implementation, implementation barriers and enablers, and how implementation had influenced their internal practices.

Second, based on information from the previous step and an analysis of the literature on VSS and related concepts an online questionnaire was developed that focused on implementation extensiveness. A pre-test with five academics and four practitioners, both having experience with C2C as well as being new to the topic, helped to refine some of the survey questions but also revealed that generally the questionnaire was clear. Then, a list of contact details was developed. A good source turned out to be press releases. In many cases certification came with press releases that included names and contact details of the persons closely related to implementation. In the remaining cases we asked for respondents who were considered as most knowledgeable on the C2C implementation trajectory (Campbell, 1955). A German questionnaire was sent to companies from Germany, Austria, and Switzerland. An English version was sent to the remaining companies, mostly from the United States and the Netherlands. Surveys were sent out in July 2015 and followed-up with up to two reminder e-mails. The population consisted of 205 firms and ultimately 76 firms participated in the study. In keeping with the process view on implementation, we decided to focus our analysis on the firms that gained their first certificate before 2013. This allowed for a minimum implementation period of three years after gaining the first certificate. There is no

official benchmark on the length of VSS implementation trajectories but our longitudinal data showed that most of the dynamics in certification happened prior to the three year cut-off point and stabilized after this point. Due to the use of 2013 as cut-off point and missing values we could use 41 of the 76 cases for the further analysis. Table 1 presents the descriptive results about the set of respondents.

In a third step we gathered additional qualitative data on the 41 cases included in the analysis through publicly available archival data such as websites, media appearances, published case studies, and press releases. This helped us to further interpret preliminary results. At the moment we are planning follow up in-depth interviews with respondents from several of the 41 cases for the same purpose.

Table 1: Summary statistics of the 41 cases included in the fsQCA analyses

Sample size		41			
Countries	USA	15	**Industry***	Textile fabric	11
	Netherlands	11		Building materials	7
	Germany	5		Furniture	5
	Denmark	3		Paper & Print	5
	Switzerland	2		FMCG	3
	Austria	1		Floor covering	3
	Belgium	1		Interior design	3
	Spain	1		Packaging	3
	Taiwan	1		Chemicals	2
	United Kingdom	1		Office supplies	2
				Personal care	1
Functional background of respondent	Sustainability	12		Other	9
	R&D	8			
	Executive Board	7			
	Marketing	5			
	Production	3			
	Sales	2			
	Product Management	2			
	Other	2			

*Multiple responses possible

fsQCA preparations

A starting point in fsQCA is to define set-membership scores and create meaningful case groupings (Ragin, 2008). In fuzzy sets, set membership varies on a continuum from 0 to 1, where 0 represents non-membership and 1 full-membership. Values between 0 and 1 represent partial membership, for example 0.66 represents more in than out, while 0.33 represents more out than in. Through a process called calibration the degree of membership of individual cases is determined. The basis for the calibration is a combination of theoretical knowledge and empirical evidence (Ragin, 2000: p.150). Ideally calibration makes use of criteria rooted in established sources external to the data (Schneider and Wagemann, 2012). However, as in our case, this knowledge is often not readily available. Hence, following earlier fsQCA research (e.g. Munoz and Dimov, 2015; Chappin et al., 2015) calibration was conducted based on the data gathered. We applied the direct method of calibration (see Ragin, 2008: p.89-94). This method uses a logit function based on three qualitative anchors, full membership (1), full

non-membership (0), and a cross-over point of maximum ambiguity of membership (0.5), to rescale survey data into fuzzy sets. For both the outcome and all the individual conditions we choose the 80th percentile of the scores as breakpoint for full membership, the 20th percentile for full non-membership, and the median as cross-over point. Furthermore, following current convention, a 0.001 constant was added to all scores to avoid theoretical difficulties of analyzing sets with membership scores of exactly 0.5 (Ragin, 2008; Fiss, 2011; Greckhamer, 2016). In the following paragraphs the outcome variables and conditions are discussed. The list of items can be found in Appendix 2.

Outcome variable The outcome variable is 'implementation extensiveness'. In keeping with the notion of VSS regulating both means and ends, we included both certification extensiveness and practice implementation extensiveness. For 'certification extensiveness' we relied on two survey questions. The first question focused on the percentage of firm products with a certificate on a scale ranging from 1 (below 10%) to 5 (over 70%). The second question focused on the highest level of certification achieved on a scale ranging from 1 (basic) to 5 (platinum). The score for certification extensiveness was calculated by taking the mean of these two questions. For the cases that had abandoned certification at the time of survey administration we entered a 0.00 for both questions. The certification extensiveness scale ranged from 0.00 to 4.50 with a mean of 2.41. For 'practice implementation extensiveness' we developed a scale based on Chappin et al. (2015) and Kennedy and Fiss (2009). We used 5-point Likert scales and the overall score was calculated by taking the mean of the 7 items (Cronbach's α = .878). The scale ranged from 2.00 to 5.00 with a mean of 3.62. Six items ranged from 1 = Strongly disagree to 5 = Strongly agree. The 7th item was a more general item that ranged from 1 = Not at all implemented to 5 = To a great extent implemented.

Conditions The variables describing motivations for adoption were developed based on Bansal and Roth (2000), Crilly, Zollo, and Hansen (2012), and Kennedy and Fiss (2009) and insights from the early interviews. The current study focuses on 'opportunity related motivations' and 'threat related motivations'. Both variables were measured on a 5-point Likert scale ranging from 1 = Unimportant to 5 = Very important. Opportunity related motivations consisted of a four item-scale. The scale ranged from 1.00 to 5.00 with a mean value of 3.92 (Cronbach's α = .755). Threat related motivations consisted of a three item-scale. The scale ranged from 1.00 to 5.00, with a mean value of 2.57 (Cronbach's α = .766).

'Level of satisfaction with the implementation' was measured with a scale that was based on Chun and Davies (2006). 5-Point Likert scales were used to measure three items. For two items the scale ranged from 1 = Strongly disagree to 5 = Strongly agree. For the third item the scale ranged from 1 = Very low to 5 = Very high. The scores on the scale ranged from 2.00 to 5.00 with a mean of 3.90 (Cronbach's α= .838).

'Level of misfit' was measured by a new variable. Based on the semi-structured interviews we learned that some firms had to significantly adjust their products, NPD process, and/or value chain to align with the C2C standard. For instance, one of our respondents argued that in the context of re-designing new products it is "*not always easy to find something that substitutes the existing substances one-to-one. Here, you really have to research extensively and conduct tests to obtain the required quality*." We used these insights to develop a three item scale. 5-Point Likert scales were used ranging from 1 = Strongly disagree to 5 = Strongly agree. The scale ranged from 1.00 to 4.67 with a mean of 3.12 (Cronbach's α = .787).

The 'level of relational embeddedness' was measured based on the scale developed by Rindfleisch and Moorman (2001). Mainly through the interviews we learned that C2C was particularly associated with two organizations that had been present since the certification system was initiated in 2005, and had been involved in standard development and consultancy: MBDC (McDonough Braungart Design Chemistry, founded by McDonough and Braungart) and EPEA (Environmental Protection Encouragement Agency, founded by Braungart). Our measure particularly focused on these two standard specific organizations. Four items were measured on a five point Likert scale ranging from 1 = Strongly disagree to 5 = Strongly agree. The scale ranged from 1.75 to 5.00 with a mean of 3.36 (Cronbach's $\alpha = .815$).

Analytical procedures

Initial analyses revealed that the correlation between the two implementation extensiveness measures, certification extensiveness and practice implementation extensiveness, was significant at the $p > 0.01$ level but poor (0.466). Therefore we decided to treat both measures separately in further analyses. Furthermore, following prior research (e.g. Fiss, 2011, Crilly, Zollo, and Hansen, 2012; Greckhamer, 2016), and in line with the asymmetric understanding of causality in configurations, we investigate both the configurations leading to a high outcome (i.e. high implementation extensiveness) as well as configurations leading to the absence of a high outcome.

A further step in fsQCA is to identify necessary or sufficient subset relations (Ragin, 2006). Conditions are necessary if they must be present for an outcome to occur and they are sufficient if they can produce the outcome by themselves. The analysis of the conditions and their negations showed an absence of necessary conditions when applying the recommended consistency benchmark of ≥0.90 (Ragin, 2008; Schneider and Wagemann, 2012). Then, sufficiency analyses were conducted using Ragin's (2008) truth table algorithm to identify configurations of conditions that consistently lead to the outcome. In so doing, we followed extant research and used a consistency benchmark of ≥0.80 (Ragin, 2008, Fiss, 2011) for each configuration. Following Greckhamer (2016) we also applied a proportional reduction in inconsistency score (PRI) benchmark of ◎0.65 to avoid the inclusion of cases that passed the consistency benchmark score for both the presence and the absence of an outcome. All analyses were carried out using fsQCA2.5 (Ragin, Davey, and Drass, 2009). Running the truth table algorithm produces several simplified or logically reduced solutions with different treatments of counterfactuals: the logically possible combinations of conditions that do not exist in the dataset. Intermediate solutions lie in the middle of the complexity/parsimony continuum and only integrate 'easy' counterfactuals. Parsimonious solutions integrate both 'easy' and 'difficult' counterfactuals and therefore represent the solutions in most reduced form (Ragin, 2008). Following extant research both solutions are reported in integrated form (e.g. Fiss, 2011, Crilly, Zollo, and Hansen, 2012; Greckhamer, 2016). Core conditions are both present in the intermediate and parsimonious solutions and have the strongest evidence linking them to the outcome. Peripheral conditions are complementary and are absent from the most reduced, parsimonious solution.

Finally, to better deal with temporal order we followed De Meur, Rihoux, and Yamasaki's (2009) suggestion to return to the cases in a more qualitative manner after the fsQCA analyses. In so doing we relied on our in-depth interviews and combined this information with our archival data and the longitudinal certification overview.

4 Results

Tables 2 and 3 present the fsQCA results. Following extant research we use Ragin and Fiss' (2008) notation. For all solutions, or

configurations, full circles represent the presence of a condition and a crossed-out circles represent its absence. Blank spaces indicate that this condition may or may not be present. Large circles indicate core conditions while small circles indicate peripheral conditions. Solutions are sorted by unique coverage and the ones that share core conditions, so called natural permutations (Fiss, 2011), are grouped. This research aims to further explore VSS implementation extensiveness and its driving factors in the context of sustainable NPD. First the configurations linked to high certification extensiveness and its absence are presented and compared, followed by the configurations linked to high practice extensiveness and its absence.

Certification extensiveness

Three configurations were consistently linked to achieving high levels of certification extensiveness (Table 2).

Table 2: Configurations for certification extensiveness

	Presence of high certification extensiveness			Absence of high certification extensiveness		
	S1a	S1b	S2	S3	S4	S5
Motivations opportunity	●	⊗	⊗	⊗	●	●
Motivations threat	●	●	⊗		●	⊗
Satisfaction	●	●	●	⊗		●
Misfit	●	⊗	●	⊗	⊗	⊗
Relational embeddedness	●	●	●	⊗	⊗	●
Consistency	0.80	0.95	0.93	0.89	0.86	0.92
Raw coverage	0.35	0.11	0.13	0.38	0.20	0.09
Unique coverage	0.25	0.03	0.06	0.29	0.12	0.03
Overall solution consistency	**0.82**			**0.88**		
Overall solution coverage	**0.44**			**0.54**		

Core conditions are represented by ● (presence) and ⊗ (absence); Peripheral conditions are represented by ● (presence) and ⊗ (absence).

The main drivers in dominant path S1a are the presence of threat related motivations for adoption, in combination with satisfaction with the implementation and a close relationship with standard specific organizations. The presence of opportunity related motivations and an initial misfit between the firm and the standard are complementary conditions. S1b differs on these complementary conditions, which are absent in this path. In contrast to the literature described in the theoretical background, it seems that for this path to high certification, opportunity related motivations are relatively unimportant. S2 has rather different characteristics when compared to S1a and S1b. Here, opportunity nor threat related motivations trigger implementation. The qualitative interviews suggest that reasons for

adoption in this case could be better captured by an articulation of the firm's philosophy through the C2C standard instead of linking the standard to concrete opportunities or threats, as was explained by one of the respondents: *"The C2C message, that you can do and consume what you like because the products are not harmful, (…) underlines our value proposition in the market."* In contrast to S1a and S1b, path S2 also has the presence of a misfit as core condition. For all three paths the presence of significant relational embeddedness is a core condition, together with the presence of satisfaction as core or peripheral condition. Three different configurations were consistently connected to the absence of high certification extensiveness. Here a relatively dominant path is S3 with an absence of all conditions at its core except for threat related motivations, which can be either present or absent. This is the only consistent path for the absence of high certification extensiveness with an absence of satisfaction as core condition. Follow up analyses revealed that for several of the firms linked to this path this absence of satisfaction even led to certification abandonment despite the fact that earlier they extended their level of certification. S4 shows a presence of both motivational factors combined with the absence of an initial misfit and significant relational embeddedness as core conditions, while satisfaction can be present or absent. This path corroborates best with the conventional idea that firms that are mainly motivated by threats will only show superficial implementation. S5 is the only path in which the presence of satisfaction combined with a presence of significant relational embeddedness are core conditions. Overall the absence of misfit is a core condition in all three paths leading to the absence of high certification extensiveness. This could be because firms in this group made a conscious choice to only certify a very small part of their product portfolio, or abandoned certification after having certified only a small parts of their portfolio and experienced unsatisfactory results. As one informant responded: *"We decided to put C2C on hold for now, and wait for the development of the standard without having to further invest money, time, and efforts from our side."*

Asymmetry can be observed when comparing the two groups of contrasting levels of certification extensiveness: paths to the presence of high levels of certification extensiveness are not mirror images from the ones leading to the absence of high levels of certification extensiveness. Relatedly, partly overlapping paths lead to sharply contrasting outcomes. This becomes particularly clear when comparing S1a and S5. Both paths have two similar core conditions (the presence of satisfaction and relational embeddedness) and both show the presence of opportunity related motivations. The difference is made through the contrasting remaining conditions. The analysis of qualitative data reveals that although the firms of both paths have several similar conditions, they had different certification strategies. S1a firms were mainly early certifiers and used extensive certification to develop a broad portfolio of C2C products. S5 firms, in contrast, used certification mainly to differentiate a small percentage of their products.

Practice implementation extensiveness

We found eight different solutions when analyzing practice implementation extensiveness, six configurations leading to its presence, two driving its absence (Table 3).

Table 3: Configurations for practice implementation extensiveness

	Presence of high practice implementation extensiveness						Absence of high practice implementation extensiveness	
	S6a	S6b	S7	S8a	S8b	S8c	S9	S10
Motivations opportunity	●	●		●	⊗		⊗	⊗
Motivations threat	●		⊗	●	●	●		⊗
Satisfaction			●	●	●	●	⊗	⊗
Misfit	●	●	●		⊗	●	⊗	⊗
Relational embeddedness		⊗	●	⊗	●	⊗	⊗	
Consistency	0.83	0.99	0.96	0.95	0.95	0.95	0.97	0.96
Raw coverage	0.43	0.28	0.26	0.21	0.10	0.18	0.48	0.42
Unique coverage	0.20	0.06	0.13	0.04	0.02	0.01	0.13	0.07
Overall solution consistency	**0.87**						**0.98**	
Overall solution coverage	**0.72**						**0.55**	

Core conditions are represented by ● (presence) and ⊗ (absence); Peripheral conditions are represented by ● (presence) and ⊗ (absence).

The first two paths (S6a and S6b) are similar for two out of five conditions. Leading to high practice implementation extensiveness are the presence of opportunity related motivations and the presence of a misfit. The satisfaction with the implementation can be present or absent. In S6a, the presence of threat related motivations presents a complementary condition whilst relational embeddedness does not play a critical role. This differs in solution S6b, where the absence of relational embeddedness drives practice implementation and threat related motivations do not play a role for the achievement of the outcome. Solution S7 has only one core condition in common with S6a and S6b, the presence of a misfit. The second core condition for S7 is the presence of significant relational embeddedness. The complementary conditions are the absence of threat related motivations and the presence of satisfaction with the implementation. Motivations based on expected opportunities can be present or absent.

The final three paths to a high practice implementation extensiveness (S8a, S8b, S8c) again differ from the previous configurations with regard to core conditions. In all three sub-solutions (S8a, S8b, S8c) threat related motivations as well as the satisfaction with the implementation are the only core conditions. Relational embeddedness is absent in S8a and S8c, however its presence is a complementary condition in S8b. Furthermore, it is remarkable in the light of extant theory that S8b has a presence of threat related motivations and an absence of opportunity related motivations and is still connected with high levels of implementation extensiveness.

Also for practice implementation extensiveness, the configurations of conditions leading to the presence of a high level of practice implementation cannot be mirrored to derive the configurations leading to the absence of high practice implementation extensiveness. As

shown in Table 3, the configurations that consistently lead to the absence of a high practice implementation extensiveness (S9 and S10) both have an absence of conditions at their core. The absence of motivations related to opportunities and satisfaction are core conditions in both configurations. In dominant path S9, the absence of a misfit and of relational embeddedness are additional core conditions. Only threat related motivations can either be present or absent. In the last solution, S10, the absence of threat related motivations is core, the absence of misfit is a complementary condition, and relational embeddedness does not play a significant role. Overall, there are no conditions, neither core or complementary, that need to be present for the absence of a high level of practice implementation extensiveness.

5 Results

This research aimed to investigate how VSS are implemented in the context of sustainable NPD and what configuration of factors drive VSS implementation extensiveness. Based on fsQCA analyses of firms that engaged with the C2C standard in their design processes we found that VSS are multifaceted: Configurations of factors that drive VSS certification extensiveness differ from the ones that drive VSS practice implementation extensiveness. Moreover we identified that configurations consistently leading to an absence of implementation extensiveness do not simply mirror the ones for high implementation extensiveness but have unique properties. We also showed that similar levels of implementation extensiveness can result from multiple distinct configurations. Our research has several implications.

A learning approach to VSS implementation

Taking organizational learning as a theoretical lens and considering configurations of implementation conditions further enriches our understanding of VSS implementation variation and its drivers. Previous research has mainly focused on motivations for adoption as driver of standard implementation extensiveness (e.g. Kennedy and Fiss, 2009; Crilly, Zollo, and Hansen, 2011). In our study the results of these prior studies were mainly corroborated for practice implementation extensiveness and not so much for certification extensiveness. To some extent this is understandable because previous research has mainly focused on practice implementation extensiveness and has not explicitly considered the level of certification extensiveness (e.g. Boiral, 2007; Yin and Schmeidler, 2009; Qi et al., 2012; Chappin et al., 2015; Bowler, Castka and Balzarova, 2016).

Beyond corroboration, we extend research on VSS implementation. In our research we found consistent trajectories leading to high implementation extensiveness that started with threat related motivations (e.g. S8b), which goes against current understandings. Based on our results an explanation is that for the companies with these trajectories the perspective on VSS changed over time. In these trajectories the presence of threat related motivations came together with the presence of satisfaction with implementation as core condition. It could be that, over time, this satisfaction provided incentives to increase implementation extensiveness. This further underlines the relevance of a more configurational and processual approach to VSS implementation than is usually done.

We further contribute by drawing attention to the presence of potential 'escalation of commitment' (Staw, 1976; Schmidt and Calantone, 1998) and its determinants in VSS practice implementation trajectories, as was already suggested in the theoretical framework. In dominant paths S6a and S6b we found that the presence of opportunity related motivations together with the efforts to bridge a significant misfit led to high practice implementation with a possible absence of satisfaction. Hence, in these cases learning trajectories could be disturbed due to high levels of prior investments: Some of

the firms maintained high levels of practice implementation, without being satisfied with C2C.

Finally, we uncovered alternative learning processes for certification extensiveness when compared to practice implementation extensiveness, which suggests that future research should consider both types of implementations as equally important but separate trajectories. For certification extensiveness, a stepwise implementation process was documented. Certification was explicitly tied to individual NPD projects and, in our case, a system with progressive achievement levels. This invoked step-by-step implementation and intermediate evaluation. In this context, the presence of satisfaction with the implementation trajectory seems to be a more powerful driver for implementation extensiveness than in the case of practice implementation where such a stepwise implementation trajectory was far less prominent. This stepwise implementation process could also explain the importance of significant levels of relational embeddedness that was found in all configurations leading to high certification extensiveness. For the majority of firms, high levels of certification extensiveness seem not to happen overnight, and firms seem to work together with standard specific organizations for a longer period of time for advice and certification. Such a trajectory would therefore benefit from partnerships between implementing firms and standard specific organizations with significant levels of relational embeddedness.

VSS implementation and product design strategies

Beyond learning processes, the variety of implementation paths that was found in this research could be further explained by connecting these paths to firms applying VSS for different product design strategies, like technological innovation vs. stylistic innovation (Ravasi and Stigliani, 2012).

In technological innovation, design is seen as combining technological product parameters to determine the functionality of a product. For stylistic innovation the emphasis is on a combination of signs, like language and symbols, to give meaning to a product. For paths with an initial technical misfit between the standard and the firm's existing products and processes as core condition, we infer that the standard used for designing technological innovations, in this case the (re)design of products in a more sustainable way. We infer from the absence of this condition in some paths that for other firms the emphasis was more on stylistic innovation through symbolically underlining already present sustainability characteristics with certifications. As in path S1a, both of these strategies could also occur together.

Another distinction in design strategies is a portfolio design strategy vs. a product design strategy (Karjalainen and Snelders, 2010). In a portfolio design strategy firms build a coherent portfolio of products that share explicit design features. In a product design strategy each product or product line has its own stylistic design and explicit design features. In contrast to the technological and stylistic innovation strategies described above, there is little room for the co-occurrence of both the portfolio and the product design strategies in a single firm. The current study found that VSS are used for both the portfolio and the product design strategies. This could further explain different levels of implementation extensiveness. In several cases of high implementation extensiveness, VSS became a distinctive design feature for the entire product portfolio and, beyond this, even became part of the very identity of the organization (Ravasi and Schultz, 2006). In contrast, an explicit product design strategy was found in several cases of low implementation extensiveness, like S5. Here firms used the standard to differentiate a single product or product line and not the overall

portfolio. The presence of portfolio design strategies challenges the completeness of the VSS evaluation systems currently is use. These systems mainly focus on the impact of VSS on individual products (e.g. Teisl, Roe, and Hicks, 2002) and not so much on the impact on firms' identities. This implies that more complete systems are needed that include a greater variety of evaluation indicators.

Boundary condition, limitations, and further research

As every study also this one has several boundary conditions and limitations that deserve attention and have to be addressed in future research. Our findings relate to a standard with specific governance mechanisms (Rasche, De Bakker, and Moon, 2013) and a broad applicability. Additional studies may investigate and compare with the implementation of VSS with different characteristics, such as industry specific VSS.

This study was restricted by the number of conditions that can be included in set-theoretic studies with a medium n (Fiss, 2011; Marx and Dusa, 2011). Other studies could investigate related conditions that seem relevant in the context of implementation extensiveness, like firm size or available resources. Using organizational learning as theoretical lens we regarded the firm mostly as a whole. In so doing we largely ignored intra-firm dynamics that may affect implementation extensiveness as well, such as the variety of reactions that implementation decisions may evoke for the different functionalities that contribute to NPD and the associated leadership aspects. Additional studies could complement our organizational level analysis by more detailed within firm studies. Finally, this study did not associate VSS implementation with performance. Further studies may investigate and compare the performance consequences of the different trajectories that were distilled.

6 Conclusion

For VSS in product design, firms do not just implement or not. Opening the black box of implementation resulted in several implementation trajectories that consistently lead to the presence or absence of high levels of implementation extensiveness. These trajectories are driven by distinct configurations of factors including motivations to adopt VSS, different learning processes with associated learning requirements and potential learning disturbances, and different design strategies. We hope that these insights will lead to more conscious VSS design and implementation trajectories.

7 References

Abrahamson, E., and Fairchild, G. (1999). Management fashion: Lifecycles, triggers, and collective learning processes. Administrative Science Quarterly, 44(4), 708-740.

Adobor, H. (2006). The role of personal relationships in inter-firm alliances: Benefits, dysfunctions, and some suggestions. Business Horizons, 49(6), 473-486.

Ansari, S. M., Fiss, P. C., and Zajac, E. J. (2010). Made to fit: How practices vary as they diffuse. Academy of Management Review, 35(1), 67-92.

Aravind, D. (2012). Learning and innovation in the context of process-focused management practices: The case of an environmental management system. Journal of Engineering and Technology Management, 29(3), 415-433.

Bakker, C. A., Wever, R., Teoh, C., and De Clercq, S. (2010). Designing cradle-to-cradle products: a reality check. International Journal of Sustainable Engineering, 3(1), 2-8.

Bansal, P., and Roth, K. (2000). Why Companies Go Green: A Model of Ecological Responsiveness. Academy of Management Journal 43 (4), 717–736.

Berchicci, L., and Bodewes, W. (2005). Bridging environmental issues with new product development. Business Strategy and the Environment, 14(5), 272-285.

Blind, K., and Mangelsdorf, A. (2016). Motives to standardize: Empirical evidence from Germany. Technovation 48, 13-24

Boiral, O. (2007). Corporate greening through ISO 14001: a rational myth? Organization Science, 18(1), 127-146.

Bowler, K., Castka, P., and Balzarova M., (2016). Understanding Firms' Approaches to Voluntary Certification: Evidence from Multiple Case Studies in FSC Certification. Journal of Business Ethics, 1-16.

Bratt, C., Hallstedt, S., Robèrt, K. H., Broman, G., and Oldmark, J. (2011). Assessment of eco-labelling criteria development from a strategic sustainability perspective. Journal of Cleaner Production, 19(14), 1631-1638.

Braungart, M., McDonough, W., and Bollinger, A. (2006). Cradle-to-cradle design: creating healthy emissions – a strategy for eco-effective product and system design. Journal of Cleaner Production 15 (13-14), 1–12.

Brunsson, N., Rasche, A., and Seidl, D. (2012). The dynamics of standardization: Three perspectives on standards in organization studies. Organization Studies, 33(5-6), 613-632.

Bruner, R., and Spekman, R. (1998). The dark side of alliances: Lessons from Volvo–Renault. European Management Journal, 16(2), 136-150.

Campbell, D. T. (1955). The Informant in Quantitative Research. American Journal of Sociology, 60 (4), 339-342.

Chappin, M. M., Cambre, B., Vermeulen, P. A., and Lozano, R. (2015). Internalizing sustainable practices: a configurational approach on sustainable forest management of the Dutch wood trade and timber industry. Journal of Cleaner Production, 107, 760-774.

Chen, C. (2001). Design for the environment: A quality-based model for green product development. Management Science, 47(2), 250-263.

Chun, R., and Davies, G. (2006). The Influence of Corporate Character on Customers and Employees: Exploring Similarities and Differences. Journal of the Academy of Marketing Science 34 (2), 138–146.

Cooper, R. G. (2008). Perspective: The Stage-Gate® idea-to-launch process—Update, what's new, and NexGen systems. Journal of Product Innovation Management, 25(3), 213-232.

Crilly, D., Zollo, M., and Hansen, M. T. (2012). Faking it or muddling through? Understanding decoupling in response to stakeholder pressures. Academy of Management Journal, 55(6), 1429-1448.

Daft, R.L. (1978). A dual-core model of organizational innovation. Academy of Management Journal 21, 193–210.

Dangelico, R. M., Pontrandolfo, P., and Pujari, D. (2013). Developing sustainable new products in the textile and upholstered furniture industries: role of external integrative capabilities. Journal of Product Innovation Management, 30(4), 642-658.

Delmas, M. A., and Montes-Sancho, M. J. (2010). Voluntary agreements to improve environmental quality: Symbolic and substantive cooperation. Strategic Management Journal 31(6): 575-601.

Delmas, M. A., Nairn-Birch, N., and Balzarova, M. (2013). Choosing the right eco-label foryour product. MIT Sloan Management Review, 54(4), 10-12.

Delmas, M., and Young, O. (2009). Introduction: New perspectives on governance for sustainable development. In M. Delmas and

O. Young (Eds.), Governance for the environment: New perspectives: 3-11. Cambridge: Cambridge University Press.

De Meur, G., Rihoux, B. and Yamasaki, S. (2009). Addressing the critiques of QCA, in: Rihoux, B., and Ragin, C.R. (Eds.), Configurational Comparative Methods, Sage, Los Angeles, CA, 147-165.

DiMaggio, P., and Powell, W. W. (1983). The iron cage revisited: Collective rationality and institutional isomorphism in organizational fields. American Sociological Review, 48(2), 147-160.

Edmondson, A. C., Bohmer, R. M. and Pisano, G. P. (2001). Disrupted Routines: Team Learning and New Technology Implementation in Hospitals. Administrative Science Quarterly 46, 685–716.

Eppinger, S. (2011). The fundamental challenge of product design. Journal of Product Innovation Management 28 (3), 399–400.

Fiss, P. C. (2007). A set-theoretic approach to organizational configurations. Academy of Management Review, 32(4), 1180-1198.

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. W., and Kolk, A. (2007). Global rule-setting for business: A critical analysis of multi-stakeholder standards. Organization, 14(5), 667-684.

Gaba, V., and Dokko, G. (2015). Learning to let go: Social influence, learning, and the abandonment of corporate venture capital practices. Strategic Management Journal. Forthcoming: DOI: 10.1002/smj.2404

Goggin, P. A. (1994). An appraisal of Ecolabelling from a design perspective. Design Studies, 15(4), 459-477.

Greckhamer, T. (2016). CEO compensation in relation to worker compensation across countries: the configurational impact of country level institutions. Strategic Management Journal. Forthcoming: DOI: 10.1002/smj.2370

Greve, H. R. (1995). Jumping ship: The diffusion of strategy abandonment. Administrative Science Quarterly, 40(3), 444-473.

Helms, W. S., Oliver, C., and Webb, K. (2012). Antecedents of settlement on a new institutional practice: Negotiation of the ISO 26000 standard on social responsibility. Academy of Management Journal, 55(5), 1120-1145.

Helms, W., and Webb, K. (2014). Perceived voluntary code legitimacy: Towards a theoretical framework and research agenda. Journal of Management & Organization, 20, 287-312.

Heras-Saizarbitoria, I., and Boiral, O. (2013). ISO 9001 and ISO 14001: Towards a research agenda on management system standards. International Journal of Management Reviews, 15(1), 47-65.

Horne, R.E. (2009). Limits to labels: The role of eco-labels in the assessment of product sustainability and routes to sustainable consumption. International Journal of Consumer Studies (33), 2, 175-182.

Ivanova, A., Gray, J., and Sinha, K. (2014). Towards a unifying theory of management standard implementation: The case of ISO 9001/ISO 14001. International Journal of Operations & Production Management, 34(10), 1269-1306.

Kan, A. K. S., Adegbite, E., El Omari, S., and Abdellatif, M. (2015). On the use of qualitative comparative analysis in management. Journal of Business Research 69(4), 1458-1463.

Karlsson, C., and Åhlström, P. (1996). The difficult path to lean product development. Journal of Product Innovation Management, 13(4), 283-295.

Karjalainen, T. M., and Snelders, D. (2010). Designing visual recognition for the brand. Journal of Product Innovation Management, 27(1), 6-22.

Kennedy, M. T., and Fiss, P. C. (2009). Institutionalization, framing, and diffusion: The logic of TQM adoption and implementation decisions among US hospitals. Academy of Management Journal, 52(5), 897-918.

King, A., and Toffel, M. (2009). Self-regulatory institutions for solving environmental problems: Perspectives and contributions from the management literature. In M. Delmas and O. Young (Eds.), Governance for the environment: New perspectives: 98-115. Cambridge: Cambridge University Press.

Klein, K. J., and Sorra, J. S. (1996). The challenge of innovation implementation. Academy of Management Review, 21(4), 1055-1080.

Kolk, A. (2005). Corporate Social Responsibility in the Coffee Sector: The Dynamics of MNC Responses and Code Development. European Management Journal, 23(2), 228-236.

Könnölä, T., and Unruh, G. C. (2007). Really changing the course: the limitations of environmental management systems for innovation. Business Strategy and the Environment, 16(8), 525-537.

Lewis, L. K., and Seibold, D. R. (1993). Innovation modification during intraorganizational adoption. Academy of Management Review, 18(2), 322-354.

Luchs, M. G., Swan, K. S., and Creusen, M. E. (2016). Perspective: A Review of Marketing Research on Product Design with Directions for Future Research. Journal of Product Innovation Management, 33(3), 320-341

March, J. G. (1991). Exploration and exploitation in organizational learning Organization Science 2(1), 71-87.

Markham, S. K., and Lee, H. (2013). Product development and management association's 2012 comparative performance assessment study. Journal of Product Innovation Management, 30(3), 408-429.

Marx, A., and Dusa, A. (2011). Crisp-set qualitative comparative analysis (csQCA), contradictions and consistency benchmarks for model specification. Methodological Innovations Online, 6(2), 103-148.

Marx, A. and Wouters, J. (2015). Competition and cooperation in the market of voluntary sustainability standards. In P. Delimatsis (Ed.), International standardization – Law, Economics and Politics. Cambridge: Cambridge University Press.

Muñoz, P., & Dimov, D. (2015). The call of the whole in understanding the development of sustainable ventures. Journal of Business Venturing, 30(4), 632-654.

Naveh, E., Meilich, O., and Marcus, A. (2006). The effects of administrative innovation implementation on performance: an organizational learning approach. Strategic Organization, 4(3), 275-302.

Nord, W. R., and Tucker, S. (1987). Implementing routine and radical innovations. Free Press.

O'Connor, P. (1994). Implementing a stage-gate process: a multi-company perspective. Journal of Product Innovation Management, 11(3), 183-200.

O'Rourke, D. (2006). Multi-stakeholder regulation: privatizing or socializing global labor standards? World development, 34(5), 899-918.

Qi, G., Zeng, S., Li, X., and Tam, C. (2012). Role of internalization process in defining the relationship between ISO 14001 certification and corporate environmental performance. Corporate Social Responsibility and Environmental Management, 19(3), 129-140.

Ragin, C. (1987). The comparative method: Moving beyond qualitative and quantitative strategies. University of California Press: Berkeley, CA.

Ragin, C. (2000). Fuzzy-Set Social Science. London: The University of Chicago Press.

Ragin, C. (2006). Set relations in social research: Evaluating their consistency and coverage. Political Analysis 14, 291–331.

Ragin, C. (2008). Redesigning Social Inquiry: Fuzzy Sets and Beyond. London: The University of Chicago Press.

Ragin, C., and Fiss, P. (2008). Net effects analysis versus configurational analysis: An empirical demonstration. In Ragin, C. (Ed.), Redesigning Social Inquiry: Fuzzy Sets and Beyond. London: The University of Chicago Press.

Ragin, C., Davey, S., and Drass, K. (2009). FS/QCA 2.5. Department of Sociology, University of Arizona: Tucson, AZ.

Rasche, A. (2015). Voluntary Standards as Enablers and Impediments to Sustainable Consumption. In Handbook of Research on Sustainable Consumption. Edward Elgar Publishing, Incorporated.

Rasche, A., De Bakker, F. G., and Moon, J. (2013). Complete and partial organizing for corporate social responsibility. Journal of Business Ethics, 115(4), 651-663.

Ravasi, D., and Schultz, M. (2006). Responding to organizational identity threats: Exploring the role of organizational culture. Academy of Management Journal, 49(3), 433-458.

Ravasi, D., and Stigliani, I. (2012). Product design: a review and research agenda for management studies. International Journal of Management Reviews, 14(4), 464-488.

Rindfleisch, A and Moorman, C. (2001). The Acquisition and Utilization of Information in New Product Alliances: A Strength-of-Ties Perspective. Journal of Marketing (65), 2, 1-18.

Rubik, F., Scheer, D., and Iraldo, F. (2008). Eco-labelling and product development: potentials and experiences. International Journal of Product Development, 6(3-4), 393-419.

Schiederig, T., Tietze, F., and Herstatt, C. (2012). Green innovation in technology and innovation management–an exploratory literature review. R&D Management, 42(2), 180-192.

Schneider, C. and Wagemann, C. (2012). Set-Theoretic Methods for the Social Sciences. A Guide to Qualitative Comparative Analysis. Cambridge: Cambridge University Press.

Simpson, D., Power, D., and Klassen, R. (2012). When one size does not fit all: A problem of fit rather than failure for voluntary management standards. Journal of Business Ethics, 110(1), 85-95.

Sleesman, D. J., Conlon, D. E., McNamara, G., and Miles, J. E. (2012). Cleaning up the big muddy: A meta-analytic review of the determinants of escalation of commitment. Academy of Management Journal, 55(3), 541-562.

Schmidt, J. B., and Calantone, R. J. (1998). Are really new product development projects harder to shut down? Journal of Product Innovation Management, 15(2), 111-123.

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. Organizational behavior and human performance, 16(1), 27-44.

Staw, B. M., and Epstein, L.D.. (2000). What bandwagons bring: Effects of popular management techniques on corporate performance, reputation, and CEO pay. Administrative Science Quarterly 45(3), 523-556.

Staw, B. M., Sandelands, L. E., and Dutton, J. E. (1981). Threat rigidity effects in organizational behavior: A multilevel analysis. Administrative Science Quarterly, 501-524.

Tamm Hallström, K., and Boström, M. (2010). Transnational multi-stakeholder standardization: Organizing fragile non-state authority. Cheltenham, UK: Edward Elgar.

Teisl, M. F., Roe, B., and Hicks, R.L. (2002). Can Eco-Labels Tune a Market? Evidence from Dolphin-Safe Labeling. Journal of Environmental Economics & Management, 43(3), 339-359.

Terlaak, A., and Gong, Y. (2008). Vicarious learning and inferential accuracy in adoption processes. Academy of Management Review, 33(4), 846-868.

Varadarajan, R. (2015). Innovating for sustainability: a framework for sustainable innovations and a model of sustainable innovations orientation. Journal of the Academy of Marketing Science, 1-23.

Wagner, M. (2008). Empirical influence of environmental management on innovation: evidence from Europe. Ecological Economics, 66(2), 392-402.

Wijen, F. (2014). Means versus ends in opaque institutional fields: Trading off compliance and achievement in sustainability standard adoption. Academy of Management Review (39), 302-323.

Yin, H., and Schmeidler, P. J. (2009). Why do standardized ISO 14001 environmental management systems lead to heterogeneous environmental outcomes? Business Strategy and the Environment, 18(7), 469-486.

Zadeh, L. (1972). A fuzzy-set-theoretic interpretation of linguistic hedges. Journal of Cybernetics, 2: 4-34.

APPENDIX

Appendix 1: Background information on fuzzy set Qualitative Comparative Analysis (fsQCA)

This appendix briefly describes central aspects of fuzzy set Qualitative Comparative Analysis (fsQCA). For an in-depth discussion the reader is referred to, for instance, Ragin (1987; 2000; 2008), Schneider and Wagemann (2012), and Fiss (2007). fsQCA is specifically designed for investigating configurations, or recipes, of conditions associated with an outcome of interest.

QCA's set theoretic approach contrasts with correlation based approaches (as used in general regression methodologies) and relies on Boolean algebra to study relationships among sets of cases (Fiss, 2007). Each case is assessed for its membership in each of the sets studied. For crsip-set QCA (csQCA) membership is evaluated in a dichotomous ('crisp') way: cases are either in (1) or out (0). In fsQCA there are graduations of membership from zero to one. Fuzzy sets are consistent to which humans would assess empirical cases (Zadeh, 1972) like, for instance, the sets of big houses, old men, or large cities, and, as in this study, implementation extensiveness. Creation, or calibration, refers to defining sets and deciding on criteria for set membership (Schneider and Wagemann, 2012). Sets are developed for the outcome of interest and for individual conditions that are may lead to the outcome.

To reduce complexity in analyses, individual cases are pooled together in 'configurations' of conditions leading to an outcome. Technically this is done via a truth table that lists all theoretically possible combinations of conditions (Schneider and Wagemann, 2012). (2^k where K = number of conditions. For instance if the analysis includes 5 conditions and one outcome, the truth table consists of 32 rows (2^5)). Each case can only be present in one row of the truth table. Each row can contain one or more cases or none. QCA software further reduces complexity through *Boolean minimization* by which configurations of conditions are simplified (Ragin, 1987: p. 93). For instance, if two configuration differ in only one causal condition but produce the same outcome, than the causal condition that distinguishes the two configurations can be considered irrelevant and can be removed to create a simpler configuration.

QCA allows researchers to identify *sufficient* and *necessary* conditions. Suppose that two configurations of three conditions (A, B, C, and A, B, D) lead to outcome Y. In this case each configuration is sufficient to lead to outcome Y. However, since these are two different configurations, each of them is not necessary to produce Y. Assessing the individual conditions, A and B are both necessary conditions because they are present in both configuration leading to Y. However, A and B are not sufficient conditions, because they need to be combined with other conditions (C or D) to produce Y. The QCA research approach requires separate analyses for evidence of necessary and sufficient relationships, the former preceding the latter (Ragin 2008; Schneider and Wagemann 2012).

The measures *consistency* and *coverage* enhance the interpretation of QCA results. Consistency measures the "degree to which cases sharing a given combination of conditions […] agree in displaying the outcome in question" (Ragin, 2008, p. 44). Consistency equals 1 when all cases sharing a configuration also share the outcome. However in fsQCA, cases frequently have partial membership in multiple configurations, which lowers consistency scores. High consistency scores support the validity of the analyzed causal path. Consistency levels are functionally analogous to statistical significance

levels used in regression based approaches (Greckhamer, 2016).

Sufficient levels of consistency is a precondition for inferring a relationship between an outcome and a configuration of conditions and interpreting its coverage (Greckhamer, 2016) Coverage measures how much a configuration 'accounts for' instances of an outcome and thus determines its empirical relevance (Ragin 2008). Three measures are usually applied (Ragin, 2008; Schneider and Wagemann, 2012). (1) Raw coverage, which is the proportion of cases of a specific outcome covered by a configuration. (2) Unique coverage, which is the proportion of cases uniquely covered by a configuration that is not covered by any other configuration. Both raw and unique coverage are configuration specific. (3) Solution coverage, which is the combined coverage of all configurations consistently linked to an outcome. Although set theoretic approaches aim to identify subset relationship instead of variance, solution coverage is functionally analogous to a regression equation's coefficient of determination (Greckhamer, 2016).

Appendix 2: Items of condition and outcome variables

Condition variables	Min value	Max value
Motivations opportunity	1.00	5.00
The expectation to increase sales		
The expectation of C2C to be a source of new opportunities		
The expectation to improve the quality of the company's product(s)		
The expectation to be perceived as a market leader		
Motivations threat	1.00	5.00
The potential loss of market share if the company does not implement C2C standards		
The competition from other C2C certified companies		
The demand of the company's customers for C2C		
Satisfaction	2.00	5.00
We would recommend the implementation of C2C to our business partners		
We are pleased to be associated with C2C		
Our company's overall satisfaction with the C2C implementation		
Misfit	1.00	4.67
In the course of implementing C2C standards, the company made fundamental changes to the existing product(s) and processes		
The innovation process had to be significantly adapted to the C2C standards		
The implementation of C2C standards had a significant impact on the entire value chain of the company		
Relational embeddedness	1.75	5.00
We feel indebted to EPEA and/or MBDC for what they have done for us.		
The company's employees share close social relations with the employees from EPEA and/or MBDC.		
Our relationship with EPEA and/or MBDC can be defined as "mutually gratifying."		
We expect that we will be working with EPEA and/or MBDC far into the future.		
Outcome variables		
Certification extensiveness	0.00	4.50
Share in the total number of products of your company's C2C certified products		
The highest C2C certification level that one (or more) of your products achieved		
Practice implementation extensiveness	2.00	5.00
The company has integrated the C2C standards in procedures and work instructions		
The company has identified specific persons and positions responsible for C2C implementation		
The company has adapted the C2C implementation procedures to its various business departments, business units or plants/warehouses		
The company has integrated the C2C standards in its computerized and other administrative systems		
The company keeps records of the training provided to staff in relation to the implementation of C2C standards		
The company has obliged its supply base to supply according to C2C standards		
The extent to which at this point in time C2C philosophy, standards, and methods have been implemented throughout your company		