THE STRATEGIC ROLE OF EXTERNAL SUSTAINABILITY ASSESSMENT IN THE FOOD VALUE CHAIN

Philipp Müller

University of Jyväsklyä
School of Business and Economics

2014

JYVÄSKYLÄN YLIOPISTO

ABSTRACT

Philipp Müller
The Strategic Role of External Sustainability Assessment in the Food Value Chain
Jyväskylä: Jyväskyklä University School of Business and Economics, 2014, 115 p.

Purpose: This research seeks to shed more light on how external sustainability assessment in the food value chain could achieve a more significant strategic impact, as most organizations fail to implement sustainability measures beyond marketing efforts. A model is proposed that aims to clarify the understanding of the process, including the influence of the actors and the way they are affected.
Design/methodology/approach: The topic is initially approached rather broadly from the perspective of performance measurement and management due to a lack of relevant literature, creating the link to sustainability and its assessment in a second step. Nine essential themes are developed from the theory and their applicability in the relevant field tested through two case studies.
Findings: The identified themes are all relevant in the context of sustainability assessment in the food value chain and furthermore show a strong interconnectivity. This interconnectivity is particularly complicated by the externalization of the assessment and can lead to conflicts, which then limits the success and the strategic opportunities of the assessments.
Research Limitations: Selecting a case study approach, this research cannot claim general applicability, particularly since the topic is relatively unexplored. Furthermore the primary data had a strong focus on experiences in western countries.
Practical Implications: Due to the complexity of sustainability assessment, involved actors need to cooperate closely and share their expertise. The developed model can help to create awareness regarding this aspect, and sensitize and prepare organizations for potential areas of conflict, facilitating a more efficient collaboration.
Originality/value: This research makes an ambitious attempt to view sustainability assessment as a modern form of performance measurement and management. In doing so it demonstrates that there are solutions for organizations that wish to manage sustainability on a strategic level.

Keywords: sustainability, sustainability assessment, performance measurement, performance management, food value chain, external assessment

Author's address Philipp Müller
Corporate Environmental Management
School of Business and Economics
University of Jyväskylä
mueller.philipp@live.com

Supervisor Tiina Onkila, Ph.D.
Corporate Environmental Management
School of Business and Economics
University of Jyväskylä

Reviewer Hanna-Leena Pesonen, Professor
Dean, Head of Corporate Environmental Management
School of Business and Economics
University of Jyväskylä

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

1 INTRODUCTION

The significance of sustainability continuously rises and whereas there is still an ongoing debate on what sustainability actually means, organizations try more and more to integrate it into their strategic management. As many of them struggle in their attempt to do so, this thesis will have a closer look at one particular phenomenon and start in this chapter by providing some background information and the rationale for studying the topic, before introducing the actual research question and the delimitations of this paper.

1.1 Background Information

The events of the recent past have risen the public awareness regarding human activity on earth: everything is connected. Significant warnings about the health of the planet appear more and more frequently and are no longer just a topic for scientists, e.g. the US department of defense recently declared that climate change was a threat to the national security. Warnings about negative impact of climate change on the economy have also become more concrete in the last years, e.g. through the work of Stern (2007), the Stern Review, that tries to monetize the effects of climate change. At the same time, each economic crisis facilitated a clear perception on the impact on the society, which could be best witnessed in Greece or Portugal. Thus the concept of sustainability, particularly the triple bottom approach with the three dimensions ecology, economy, and society, has gained recognition, especially since the World Commission on Environment and Development's (WCED) addressed the topic and defined sustainable development as: "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (1987: 54). The idea of not depleting the resources faster than they can be renewed, can be traced back to literature on forestry of the 18th century (von Carlowitz 1713); ironically, its author was worried about sufficient wood supply in order to extract and process ore.

One sector that is strongly affected by sustainability concerns is the food industry. How deeply this connection goes, is best illustrated by Lagi et al. (2011),

who analyzed the relation between the development of food prices and the political situation in the middle east and found that social unrest, such as the Arab spring, is most likely to break out in times of food price spikes. A major role falls and will fall to agriculture, as projections indicate a strong global population growth, from currently around 7.2 billion people to a possible headcount of 8.1 billion in 2025 and 9.6 billion in 2050 (United Nations Department of Economic and Social Affairs/Population Division 2013: 1), an increase of 12.5 and 33.3 percent respectively. At the same time, many inhabitants of less developed countries are expected to increase their standard of living which is generally accompanied by a more resource demanding nutrition. In a nutshell, there will be more people to feed with a higher consumption rate per capita, on less resources, including the availability of arable land. Simultaneously, the possible consequences of overconsumption such as longer drought periods due to climate change, are likely to lead to poorer farming conditions. However, the agricultural sector is already one of the key contributors to global greenhouse gas emissions, the land sectors (including forestry and other land uses) accounting for almost 30 percent of human-induced emissions (FAO 2013: 219-220), which will further increase if the growing demand is satisfied with current practices.

Meanwhile, a trend to more sustainable production standards is perceivable. Whereas some actors from within the value chain push the topic, e.g. due to concerns about resource security, there is also a strong push coming from external stakeholders, particularly NGOs and consumers. Thus, sustainable development is now a basic requirement for the food industry, which labels more products and processes sustainable than any other industry (Frank et al. 2014).

1.2 Rationale for Studying the Topic

Sustainability has thus become increasingly relevant for the business sector. Nevertheless, the topic is still often not integrated into the core business of organizations and too often limited to Corporate Social Responsibility (CSR) activities for marketing purposes (Porter & Kramer 2011). At the same time, the literature points out how decisive the integration is in order to achieve sustainability (e.g. Hart 1995, 1997; Roome 1998; Figge et al. 2002) and practitioners across the world do not believe that their actions are sufficient to meet the global sustainability changes as their sustainability initiatives have reached a plateau (UN & Accenture 2013). The most important reasons for this stagnation are problems to quantify and capture the business value of sustainability.

The resulting demand for solutions has led to a growing number of sustainability performance assessment tools in the food industry, mostly focused on agricultural activities. These tools range from simple self-assessments for farmers to complex, holistic LCA-methods across the value chain. Interestingly, the experience of providers of such tools has shown that their clients again tend to see the opportunities mainly from a CSR-perspective instead of a strategic management point of view and often struggle to implement recommendations

from the assessments. The potential of these tools to complement or replace traditional performance measurement approaches has also been rarely touched upon in the literature.

This particular situation has led to the decision to have a closer look at the topic and examine the link between performance measurement and management and sustainability performance assessment in the food sector.

1.3 Research Question and Objectives

As the previous parts have shown, organizations struggle to scale their sustainability efforts as they fail to measure and concretize them, and solutions by external providers often do not succeed to generate the desired impact as well. Also the literature so far largely neglected the strategic management potential of these methods.

Therefore, this research aims to examine the strategic role of external sustainability performance assessments in the food value chain, more precisely how the assessments might achieve a more significant strategic impact.

Two objectives can be derived from this research question. First of all, the process of an external sustainability performance assessment in the context of the food value chain needed to be clearly understood and in order to do so, its essential elements had to be identified. Second, it had to be analyzed how the actors are involved and affected during the process to ultimately understand the role of the sustainability assessment and its strategic impact.

The existing literature did not provide sufficient information regarding the question, in particular not in the context of the food value chain, therefore a broader and more general review was conducted across involved disciplines in Chapter 2. Subsequently, experts along the food value chain from two different, popular tools who had different roles in the process of sustainability performance assessments were interviewed after having developed a set of essential themes from the literature. This enabled the researcher to test the theoretical findings in a new context, thus largely complementing the first research objective. The primary findings as presented in Chapter 4 were then furthermore used to develop a model in Chapter 5 that depicts not only the process of sustainability assessment, but the role, as well as the connections and interactions among the initially identified themes, thus meeting the second objective.

1.4 Delimitations

Sustainability and particular sustainability performance assessment is still quite new to the corporate world, nevertheless, many different theories and disciplines are involved, and consequently the topic could be approached from various angles and points of perspective. For the purpose of this thesis it seemed most appropriate to start with an introduction into performance measurement and management. As the assessment tools of the case studies are already established, the focus will not so much be on system design issues, but rather to analyze how

these systems affect the performance and which challenges occur. Afterwards, the topic of sustainability will be presented, and also here, the question will not be, which behavior is most sustainable – a question that is fiercely debated, particular in agriculture – or whether or not the tools actually manage to assess sustainability, but which role sustainability plays in the corporate context, how sustainability assessment impacts on the organizational performance, and which difficulties appear.

Regarding the scope of the study certain limitations have to be considered as well. Due to the sample size of nine interviews across two assessment tools, the results will not be generalizable. Both tools were developed for the food industry and thus the findings stem from experiences in that specific sector. More precisely, most of the experience had been gathered in agriculture, yet the scope often also included more steps of the value chain. Therefore, this research might focus on food, but largely on experiences from the agricultural sector and subsequent steps. Furthermore, many of the clients, including those that were interviewed, were no farmers, besides the fact that the assessment results largely address farm activities. Most of them were umbrella associations, consultants, or from further down the value chain (Tier 2), e.g. processors. Even when the clients were on the producer level (Tier 1), they were usually large producers with several farms and their own management department that dealt with the assessment, and not the farmers themselves. The scope is illustrated in FIGURE 1.

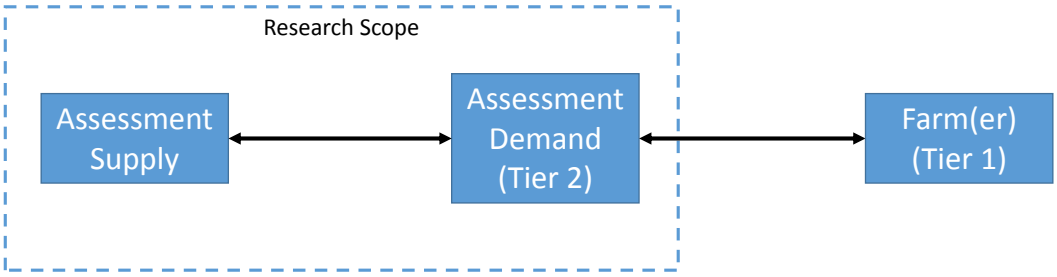


FIGURE 1 Research Scope

This part served to point out the main reasons and objectives for examining this particular phenomenon, and introduce the topic. Subsequently, the relevant framework and the central concepts of this thesis are presented in more detail.

2 LITERATURE REVIEW

The chosen topic involves several areas that overlap in certain aspects. Nonetheless, for the subsequent literature review, they shall be addressed individually at first, before connecting the dots. The review will start by having a closer looking at performance measurement, as the collection of data generally forms the basis of all decision-making. The logical next step, to draw a conclusion from the available data shall then be regarded subsequently including the practical framework of management systems. Though these systems might work well as long as the focus stays within a firm's boundaries, it will become apparent that covering parts of or the whole supply chain poses a significantly more complex challenge. Today's global economy nevertheless requires such a holistic approach, especially when it comes to sustainability, which will be addressed isolated at first, before creating the link to performance measurement and management.

Before starting the literature review, it is important to remember the task of this research, analyzing external sustainability performance assessments and their strategic role for organizations in the food value chain. Therefore, rather than designing new tools, the underlying processes of the existing methods were examined to highlight the areas that appeared to have a strong influence on the clients' performance. The purpose of this literature review is to get a better understanding of the involved concepts in preparation for the subsequent primary data collection and further analysis. Therefore, the vast body of literature dealing with the proper design of performance measures and measurement systems was considered under the light of how certain system characteristics influence the performance, thus allowing the author to draw conclusions on the applicability of the existing tools. As the literature on the relevant topics in the context of the food value chain was rather limited, the author decided to conduct the review from a broader perspective, as the applicability of the theoretical findings to the food value chain could be tested during the interviews with experts from the field.

2.1 Performance Measurement

Performance measurement and its application in the business world increasingly received attention, particularly within the last 20 years, as the rising numbers of publications and citations dealing with performance measurement (Neely 2005, Taticchi et al. 2010) and a continuous evolvement of software solutions well illustrate. Some type of measurement has most likely been part of business since the first (professional) trades took place and thus, its crucial role for business performance is part of the literature since decades as well, originating from organizational control theory (Demartini 2014: 10). In the 19th century, Lord Kelvin (1824-1907) pointed out that "When you can measure what you are speaking about, and express it in numbers, you know something about it". Nonetheless, Bourne et al. (2005) still observe a constant debate whether performance measurement actually positively influences the performance of a business.

2.1.1 Definition of Performance Measurement

A possible explanation for this apparent contradiction can be found in the development of business performance measurement that involved practitioners and academics alike, with very diverse backgrounds such as accounting, management, marketing, and human resources (Neely 1999, Marr & Schiuma 2003), without being a research discipline in its own right. This has led to the lack of a cohesive body of knowledge (Marr & Schiuma 2003).

It therefore makes sense to have a closer look at the definition of performance measurement. Neely et al. (1995: 1229) suggest to define:

- Performance measurement "as the process of quantifying the efficiency and effectiveness of action" with a
- Performance measure being "a metric used to quantify the efficiency and/or effectiveness of an action".

It is worthwhile to highlight this dual function of performance because it can lead to internal as well as external factors influencing the decision process (Slack 1991) given a specific context, e.g. customer satisfaction: in order to outperform their competitors, firms need to be more effective in terms of meeting their customers' requirements (external factor) and more efficient in their resource utilization (internal factor) (Kotler 1984).

One word that is rather rarely defined in the literature is performance itself, despite its frequent utilization. An exception regarding the definition is Lebas (1995: 23), who clearly distinguishes between performance and results, as he views performance as the future potential of a firm:

> "Performance, especially in the case of management, is not so much about past achievements, as generally accepted, but about the future, about the capability of the unit being evaluated. It is so because, in our mind, the purpose of management is about creating and shaping the future of the organization, as well as that of society."

However, he also admits its complexity (Lebas & Euske 2002). This makes a more precise definition difficult and after all, performance is highly contextual depending on who applies it and what it is to be used for (Lebas 1995). Admittedly, this makes it hard in practice and might explain why other authors chose a more concrete approach to identify performance, such as examining the profit and loss account (e.g. Bourne et al. 2005).

2.1.2 Definition of Performance Measurement System

Interestingly, the real discussion is not so much about the definition of individual metrics, but rather their function and interaction within a system, that Neely et al. (1995) define as a "set of metrics used to quantify both the efficiency and effectiveness of actions", a definition Neely will later update to "*past* actions" (Neely et al. 2002), a refinement owed to one of the main points of critique on Performance Measurement Systems (PMSs): the fixed alignment towards past events, making them rather unsuitable for strategic support, an issue that will still be addressed more in detail later on.

Even though PMSs lack a clear definition and as a consequence, research outcomes are hard to compare and generalize (Franco-Santos et al. 2007), it is possible to structure existing definitions according to three professional perspectives:

- *Operations*: The focus lies on measuring the performance of actions as defined by Neely et al. (1995, 2002). Additionally, a PMS can also be perceived as a tool to communicate outcomes of actions to employees (Bititci et al., 1997).
- *Strategic Control*: Here two roles can be fulfilled by a PMS, first of all, it illustrates how certain measures interact with each other while providing the required data to review the status of the current strategy (Ittner et al. 2003). Secondly, it also allows to draw conclusions on the process of strategy implementation based on which performance metrics were developed (Gates, 1999).
- *Management Accounting*: Three different roles are possible, firstly as a financial management tool for planning and budgeting. Second, PMSs inform about the overall business performance. Last, they can be utilized as tools of motivation and control (Otley 2006).

These perspectives not only highlight the various fields that influence PMSs, but also show how versatile such systems can be according to their design.

2.1.3 Roles of Performance Measurement Systems

Nonetheless, one aspect that needs to be stressed is that the different roles and functions need to be recognized, as measurement systems designed for a specific purpose will most likely fail when they are used to fulfil another role (Otley 2006). In some cases, marketing is listed as a separate perspective as well (see Clark 2006 for a historic overview of marketing performance measurement) and

although it is certainly crucial to consider metrics such as customer satisfaction, it is rather a part of the previously noted perspectives and supports them (e.g. to shape or assess the external perception of strategies), than a perspective on its own. Still, the key take away should not be the discussion about possible classification, which is a difficult task due to overlaps, but the awareness that a PMS should be designed according to its purpose. For this reason, possible roles that Franco-Santos et al. (2007) identified in the literature shall be listed subsequently, independent from the perspective. It is quite intriguing to see that there is a very close overlap with Henri (2006) suggesting *Monitoring*, *Attention focusing*, *Strategic decision making*, and *Legitimizing*, and thus rather naming the goal of each role:

- *Measure performance*: Besides measuring, the role is to evaluate the performance and observe the progress. It serves to answer the question 'how am I doing' (Simon et al. 1954);
- *Learning and improvement*: Both articles list learning through feedback, however Henri (2006) embeds it in *Strategic decision making*;
- *Strategy management*: Answering 'What problems should we look into' and 'Of the several alternatives, which is rationally the best' (Simon et al. 1954), the authors incorporated both *Attention focusing* and *Strategic decision making*, furthermore including the functions planning, strategy formulation, and alignment;
- *Communication*: Often used to justify and validate company actions from the past, present, and future to internal and external stakeholders, *Legitimizing* is precisely describing the category as well; and
- *Influence behavior*: It comprises the roles of directing behavior through rewards or compensation, as well as managing and controlling relationships.

The overlap between the two works shows that certain characteristics might be assigned to and combined in different roles, yet the major purposes a PMS can fulfill become clear.

2.1.4 Influence of PMS on Organizational Performance

Whereas the importance of the awareness about the different roles a PMS can play has now been stressed, the question whether such systems generally have a positive influence on business performance still remains unanswered. As previously noted, there is no clear answer to this questions and results from literature are inconclusive and in some cases even contradictory (Bourne et al. 2013). But as Bourne et al. (2005) point out, the more alluring question for practitioners is under which circumstances performance measurement has a positive impact on organizational performance. In an extensive literature review they adopted the framework suggested by Pettigrew et al. (1989): *Context* (internal and external), *Content* (what is measured and how the measures are structured), and *Process*. The factors are presented in TABLE 1. One finding stuck

out particularly: when multiple business units within one organization were compared, the main difference between high- and average-performers seemed to appear in the handling of the system; those managers that were more interactive in their approach considering local prevalent circumstances rather than global company targets (that might not be optimal for their unit), achieved better performances (Bourne et al. 2005). This demonstrates how contextual performance measurement is and the significance of adaptation, even though it might mean giving less priority to targets coming from the top.

TABLE 1 Overview of Factors and their Influence on Performance (adapted from Bourne et al. 2005)

Factors	Influence	Source
External Context		
Industry competitiveness	More intense market competition leads to more effective use of information	Lokman & Clarke 1999
Economy (market uncertainties, supplier characteristics, econ. situation)	MS's effectiveness depends on the speed of change and measurability of performance	Smith & Goddard 2002, Waggoner et al. 1999
Political environment	Economic constraints and regulatory regimes influenced the use of measurement systems	Hussain & Hoque 2002
Internal Context		
System maturity	More mature systems are more effective	Evans 2001, Martins 2002
Organizational structure	Aligning structure and measurement is important	Hendricks et al. 1996, Bourne et al. 2002
Organizational size	Measurement is easier in larger companies and more complicated for smaller ones	Hoque & James 2000, Hudson et al. 2001a, b
Organizational culture	MS profits from aligning users' cultural preference and the cultural elements embedded in the system	De Waal 2002, Gates 1999, Johnston et al. 2002, Lingle & Schiemann 1996, Lockamy & Cox 1995, Maisel 2001, Malina & Selto 2001, Bititci et al. 2004
Management style	Appropriate style matters (depending on circumstances, use, and level of implementation of MS)	Gelderman 1998, Libby & Luft 1993, Hunton et al. 2000, Simon 1987, Bititci et al. 2004
Competitive strategy	Benefits through alignment of measures according to strategy	Kaplan & Norton 1996, 2001, Lockamy 1998, Mendoza & Saulpic 2002, McAdam & Bailie 2002, Neely 1998

(continues)

TABLE 1 (continues)

Resources & Capability	Resources and capabilities are required by companies for implementing and refreshing the MSs	Kennerley & Neely 2002
Information systems & Infrastructure	Importance of high data integrity and a low burden of data capture	Bititci et al. 2002, Eccles 1991, Lingle & Schiemann 1996, Manoochehri 1999
Other management practices and systems	MS and other systems should be aligned (e.g. budgeting, compensation)	De Toni & Tonchia 2001, Eccles 1991, Eccles & Pyburn 1992, Kaplan & Norton 1996, 2001, Moon & Fitzgerald 1996, Otley 1999
Content		
Definition of performance measures	MS are more effective when they are appropriately designed	Neely et al. 1997
Dimensions measured	Used dimensions direct the management focus and can lead to more effective MS	Kaplan 1994, Lingle & Schiemann 1996
Structure & presentation	Structure that reflects the strategy and interrelation-ship makes MS more effective	Lipe & Salterio 2000, 2002
Process		
Alignment with strategic objectives	Alignment can improve motivation	Atkinson 1998, Otley 1999
Data capture	More individual data collection leads to higher performance	Lynch & Cross 1991, McGee 1992, Simons 1991, Neely 1998
Data analysis	More individual approa-ches by managers improve performance	Lynch & Cross 1991, Neely 1998
Interpretation & Evaluation	Comparing results beyond global company targets supports higher perfor-mance	Simons 1991, Neely 1998, Ittner et al. 2003, Kerssens-Van Drongelen & Fisscher 2003
Decision making & Taking action	Continuous acting on trends and issues beyond company targets improves performance	Ittner et al. 2003, Neely 1998, Flamholtz 1983, Flamholtz et al. 1985, Simons 1991
Communication & Information provision	Constant spreading and discussion of performance leads to higher performance	Bititci et al. 1997, Forza & Salvador 2000, Kerssens-Van Drongelen & Fisscher 2003, Lebas 1995, Lynch & Cross 1991, Simons 1991, McGee 1992, Neely 1998, Otley 1999

MS = Measurement System

As the previous parts suggest, performance measurement can positively impact on performance, however this depends on many internal and external factors. At this stage it makes sense to point out again that the literature analyzed so far assumes performance measurement and according systems to be designed, set up, and conducted internally, possibly with external support. This does not precisely reflect the circumstances of the subsequently investigated cases: the assessment of firms' sustainability performance were conducted with established tools by an external provider. It is not unlikely that this aspect has an influence on the impact, as Bourne et al. (2005) found that even within one organization, those business units whose managers took a more individual and contextual approach during the measurement processes performed better than business units whose managers relied on the standard company systems.

2.1.5 Shortcomings

The previous part gives an insight into many important factors related to the design, implementation, and utilization of a PMS, and therefore also provides an impression of what can go wrong in practice. Another point of criticism against the practical application of PMS that continuously reappears throughout the literature can be summed up as neglecting non-financial dimensions of performance (Drucker 1954; Johnson & Kaplan 1987; Fitzgerald 1988; Goold & Quinn 1990; Ghalayini & Noble 1996). The main reasons why the sole use of traditional financial measures is considered incomplete are that they:

- encourage shot-term planning (Banks & Wheelwright 1979; Hayes & Abernathy 1980);
- do not provide a strategic focus nor data on quality and flexibility (Skinner 1974);
- can easily lead units to only optimize local results, e.g. keeping machines and people busy by manufacturing inventory (Goldratt & Cox 1986; Hall 1983);
- stimulate management to rather decrease deviations from the standard than trying to achieve continuous improvement (Schmenner 1988; Turney & Andersen 1989); and
- lack data on customer wishes and the performance of competitors (Camp 1989; Kaplan & Norton 1992).

Even though this call for additional measurement beyond the financial dimension can be found throughout the literature on performance measurement, it does not seem to be easy to solve. Two aspects document this: Firstly, Choong (2013) finds that the criteria discussed in the PMS literature are still largely of financial nature. Secondly, there is no clear agreement on what the other dimensions should be (Franco-Santos et al. 2007). Also the often suggested balance between the measures - eponymous for one of the most popular frameworks for performance measurement and management: the Balanced

Scorecard, developed by Kaplan and Norton - cannot really be proven conclusively (Kennerley & Bourne 2003).

One last potential shortcoming that still needs to be addressed is the question whether measurement is really relevant for business or in other words, whether the saying 'what gets measured gets managed' holds any truth. Though he did not reject it, Eccles was already more carefully in his formulation when he exclaimed that "what gets measured gets attention, particularly when rewards are tied to measures" (1991: 131). An empirical study on the topic also found no significant relationship between indicating and acting, but proposes the introduction of mobilizing (moving a firm from passive- to activeness) to produce a more fitting model: "What gets mobilized gets managed, especially if it gets measured" (Catasús et al. 2006: 516). Thus the situation can be summed up as: successful management by solely measuring is rather unlikely; it requires active care, which can be initiated, directed, and aligned through measures, and supported by compensation.

2.2 Performance Management

Performance management generally enables organizations to get from the point where they know what to do, to how to do it, overcoming the 'knowing-doing' gap (Cohen 1998), "effectively translating information coming from the measurement […] into effective tasks" (Taticchi et al. 2009: 48), thus it becomes relevant for this research. Without acting on the information of the assessment, it cannot directly impact on the organizational performance.

2.2.1 Definition of Performance Management System

The borders between performance measurement and management seem to be quite fluid, particularly so for measurement and management systems. An illustrative example of the blurriness is the previously mentioned concept of the Balanced Scorecard which appears in literature on performance management as well as measurement.

In order to clarify the differences, a short reminder of the previously used definition of a PMS: a set of metrics used to quantify both the efficiency and effectiveness of past actions. In contrast, a performance management system can be described as a set of

> "the evolving formal and informal mechanisms, processes, systems, and networks used by organizations for conveying the key objectives and goals elicited by management, for assisting the strategic process and ongoing management through analysis, planning, measurement, control, rewarding, and broadly managing performance, and for supporting and facilitating organizational learning and change" (Ferreira & Otley 2009: 264)

Thus, the definition lists three functions that a management system should serve: controlling, supporting, and enabling, bearing a striking resemblance to the Deming-Cycle consisting of Plan, Do, Check, and Act, that is generally used as the basis for standardized Management Systems provided by the ISO or the EU (EMAS), which will still be addressed later on. For now, it is sufficient to

highlight two further elements of Ferreira & Otley's definition: first, it includes formal as well as informal mechanisms, and second, it refers to the effectiveness in strategy achievement (Demartini 2014: 9). These elements also richly illustrate the evolutionary path of such systems, starting as a controlling tool towards more efficient management (Taticchi et al. 2010a) that allows continuous improvement of the performance (Neely et al. 1995), to developing, implementing, and diffusing strategies (Kaplan & Norton 1996), to aligning operations with strategic objectives, and finally, to enabling organizational learning (Kueng et al. 2001). In a way, control is still the primary function of the systems (Amaratunga & Baldry 2002), but the perception and comprehension of control has changed: whereas the initial meaning had a rather negative touch due to the constraining character, nowadays controlling is perceived as constructive and guiding (Demartini 2014: 10).

At first glance, there is a close resemblance between performance management systems and PMSs indeed, especially if the roles of a PMS as given by Henri (2006) are recalled: Monitoring, Attention focusing, Strategic decision making, and Legitimizing. However, keeping in mind that the results of performance measurement generally show what happened in the past and not why it happened, and the definition of performance being rather about potential future than past achievements, plus the elements of the previous paragraph (especially the enabling function), should make it apparent that Ferreira & Otley intended to define a more holistic approach to managing and controlling organizational performance. After all, an increasingly complex business world requires organizations to better understand cause-effect relationships in order effectively provide support to decision-making processes (Taticchi & Balachandran 2008). The idea of measurement and PMS assisting in the management process and therefore being a part of the whole system instead of a holistic approach in itself is nicely captured by Amaratunga & Baldry (2002: 218): "Measurement is not an end in itself but a tool for more effective management." Nevertheless, in practice, systems can be designed and adopted according to an organization's individual needs which makes it probably impossible to draw a clear line. A compromise that can be found in more recent literature would be to combine the two aspects under a new name: performance measurement and management systems (PMMSs) (Taticchi et al. 2010a), which will be adopted subsequently.

As a precise definition is not of primary concern for this research, the more relevant question, how management systems influence the organizational performance, will be addressed subsequently.

2.2.2 Influence of Performance Management Systems on Organizational Performance

Earlier it was emphasized that in the case of PMSs, it is not so much about if, but rather under which circumstances the organizational performance is positively influenced. De Waal & Counet (2009) claim that utilizing performance management systems is one of the few techniques available to management that

indeed helps organizations to achieve better results. However, this should not be taken as generally valid, even though there are a number of case studies supporting this claim (Ahn 2001, de Waal & Coevert 2007).

First of all, as every organization is different, it faces specific challenges that their systems need to incorporate, which makes the systems rather individual as well. Secondly, successful management systems often come at high costs: literature reports the rate of performance management system implementations that failed to be around 70 percent (McCunn 1998, Neely & Bourne 2000). Though experts assume the rate to be rather around 56 percent nowadays, possibly due to a larger body of knowledge and experience (De Waal & Counet 2009), this still means that more than every second implementation attempt did not succeed. So even if everything goes well, the implementation was successful, the system well designed, and it can be assumed to positively impact on the organizational performance, the performance increase still needs to make up for more than one failed attempt.

In order to allow for a more holistic description of performance management systems, Ferreira & Otley (2009) developed a framework (see ANNEX 1) that can serve as a template to capture the key characteristics of systems. There are two aspects that are quite intriguing to see about the framework: First of all, it is an extended version that is largely based on a previous paper from Otley (1999). It is worth to look at how it developed within those ten years. Whereas the original only consisted of five questions (see ANNEX 2), the new version includes seven additional questions (see ANNEX 3). Though the original questions can still be recognized in the new framework, confirming their relevance, the focus has slightly shifted or rather extended. Now the framework takes a broader perspective, e.g. strategic objectives have been complemented by taking the overall vision and mission of the organization into account as well as the structure and culture. Also the communicational aspect, how information is spread and received across the organization is more emphasized. Furthermore, the type of evaluation, subjective, objective, or mixed, formal or informal, and financial or non-financial is specifically addressed in the updated framework as well as a reward system. Finally, the new questions also target a more dynamic nature of a management system and the way the systems are used, as well as the interaction and interrelations between the elements of the system. The newly added questions thus largely react to criticism towards the original framework (Demartini 2014: 72).

The second remarkable aspect within the later framework is the strong similarity to the previously discussed influences of PMS (see TABLE 1) in many areas, in particular in relation to organizational culture and structure. Other categories are similarly covered as well, such as evaluation, alignment with strategic objectives, well designed and chosen key performance measurement, illustrating interrelationships, and communication (Information flows and infrastructure). The clear communication of performance expectations is crucial to allow employees to deduce the cause-effect attributions (Bowen & Ostroff 2004). There are two conclusions that can be drawn out of this information: first,

these aspects really seem to be of significance for organizational systems in general. Second, it once more demonstrates the overlap between performance measurement systems on the one hand and management systems on the other hand.

2.2.3 Standardized Management Systems

Standardized Management Systems, such as the exemplary QMSs (Quality Management Systems) and EMSs (Environmental Management Systems) presented later on, are relevant for this research question for two reasons. First of all, unlike the previous systems that are usually developed within the organizational boundaries, they are provided by an external third party, meaning that they are far less custom-tailored to any specific organization, even though they are usually kept in a general manner and thus still leave room for customization. Secondly, since they are standardized (to a certain extent) they are also better to compare and thus more suited to analyze specific effects, such as their influence on organizational performance.

According to the International Organization for Standardization (ISO), "a standard is a document that provides requirements, specifications, guidelines or characteristics that can be used consistently to ensure that materials, products, processes and services are fit for their purpose" (ISO 2014a). The ISO also states that standards are strategic tools for businesses that allow them to save costs by minimizing errors and waste while increasing productivity, so in other words they enhance the overall performance of an organization. The ISO provides some of the best-know management systems standards, such as the ISO 9001 that sets out the criteria for quality management systems or the ISO 14001 for environmental management systems. As previously noted, the structure is usually based on the Deming Cycle (Plan-Do-Check-Act) with the overall objective on continuous improvement and thus not different from the previous performance management system definition. Though standards are often criticized for being too expensive, it should be pointed out that it is not the standards that are costly – the EMAS (Eco-Management and Audit Scheme), an environmental management system standard very similar to the ISO 14001, is freely provided by the European Commission – but the certification process, which is not required (ISO 2014b) in order to run the management system. Due to its standardized character, there is more literature with explicit results about the impact of management systems on organizational performance, which will be regarded subsequently.

When it comes to QMS and thus an ISO 9000 certification, the case is slightly easier, as there is a direct link to organizational performance. EMSs on the other hand, are a two-step process as their primary purpose is the improvement of an organization's environmental performance. Whether this has an impact on the overall corporate performance will be addressed later on. The measures to indicate the overall performance used in the literature were often of financial nature, such as earnings before taxes or operational costs growth rate (Martínez-Costa & Martínez-Lorente 2007), despite the general opinion that financial

measures are likely to fail in capturing the overall picture. Due to the more direct link, a possible QMS impact will be examined before having a look at EMSs. Even though those are not the only available standards, EMSs and QMSs are presented as they come with the largest body of literature.

2.2.3.1 QMS

When examining how organizations can profit from implementing a QMS (generally the ISO 9000 family), one can distinguish between internal and external benefits (see TABLE 2 for an overview). Though the benefits are attributed to the certified QMS, the certification is mainly required to achieve the external benefits. At least in theory, the internal benefits could be achieved without certifying the QMS.

TABLE 2 ISO 9001 Certification Benefits (Sampaio et al. 2009)

Internal benefits	**External benefits**
Productivity improvements	Access to new markets
Product defect rate decreases	Corporate image improvement
Quality awareness improvements	Market share improvement
Delivery times improvements	Customer relationship improvements
Internal organization improvements	Customer satisfaction
Nonconformities decreases	Customer communication improvements
Customers complaints decreases	ISO 9000 certification as a marketing tool
Internal communication improvements	
Product quality improvement	
Competitive advantage improvement	
Personnel motivation	
Definition of the personnel responsibilities and obligations	

Both sides can serve as a motivation to seek quality certification, and though both reasons can be present, one of the two factors usually predominates (Sampaio et al. 2010). Interestingly, the effect of the QMS seems to depend on the original motivation for the certification: whereas companies that sought certification due to external reasons were more likely to achieve external improvements, those firms that were rather driven by internal motivation obtained higher profits from the implementation and were more likely to move towards total quality management (Llopis & Tarí 2003), maximizing their benefits from the implementation (Sampaio et al. 2012). It is noteworthy that much of the literature focused on the relevance of the certification of and not the QMS itself.

Although motivation seems to be highly relevant, all in all, there is no agreement yet on the ultimate influence of the ISO 9000 on performance, as the results of a growing body of literature are still contradictive regarding the necessity of certification (Sampaio et al. 2012); Martínez-Costa & Martínez-Lorente (2007) even found a possible negative impact of the certification as potential market benefits do not offset the costs of implementing and maintaining a certified ISO 9000. This strong variance of outcomes could be explained by the

multitude of variables that influence the performance (Heras et al. 2002) or as an indicator that the effects are still only poorly understood (Pavlov & Bourne 2011). For this research, however, it is sufficient to know that the driving motivation is of great relevance.

2.2.3.2 EMS

After all the disagreement on the effect of (certified) QMS, it comes as a surprise to find more consent among the EMSs literature regarding the influence on performance. De Vries et al. (2012) discovered that a large majority of papers (30 out of 34) found a positive relationship between adopting ISO 14001 and business and/or environmental performance. Analyzing whether there is a difference between lacking a formal EMS, having a formal EMS, and having a formal certified EMS, Melnyk et al. (2003) concluded that while the presence of a formal EMS already brings benefits, the certification further improves the performance, which also helps to explain varying research outcomes. They suggest that three reasons could lead to this result:

- the certification process helps to involve more people from within the company, raising awareness for the environmental activities of the organization;
- the evaluation through a third party could encourage additional improvements; and
- the certification requires to actually examine the underlying processes instead of just focusing on the outcomes.

Although Melnyk et al. (2003) examined only cases using ISO 14001, Iraldo et al. (2009) reached similar conclusions for EMASs, finding EMAS registered companies (the equivalent of certified ISO 14001 bodies) to handle their EMS in a more comprehensive and effective manner. However, they argue that adopters of both EMAS and ISO 14001 scarcely perceive any competitive advantages as the standards' designs fail to provide such, e.g. the EMAS logo is not permitted to be used on products. Nevertheless, as the standards are updated on a regular basis – a new version of the ISO 14001 is expected by the end of 2015 (ISO 2014c) – the aspect could be considered in future designs.

To round up this chapter on EMS, it is worth looking at the factors that seem to positively influence the EMS's impact on business and/or environmental performance, TABLE 3 provides an overview. What sticks out once more, is the reoccurrence of *Internal motivation* that was already highlighted for QMSs. Noteworthy are furthermore the factors *Maturity of the system*, as well as *Company size*, since both also appear in TABLE 1 as factors that influence the impact of PMSs.

TABLE 3 Factors Positively Influencing the Impact of ISO 14001 (adapted from de Vries et al. 2012)

Factor	Source
Maturity of the system	Melnyk et al. 2002, 2003
Top and middle management commitment	Darnall et al. 2000, Tan 2005, Yin & Ma 2009, Zutshi & Sohal 2004
Internal motivation	Boiral 2007, Tien et al. 2005
Company size	Szymanski & Tiwari 2004
Well-defined responsibilities	Vastag & Melnyk 2002
Employee training and involvement	Babakri et al. 2003, Boiral 2007, Darnall et al. 2000, Morrow & Rondinelli 2002, Rondinelli & Vastag 2000, Tien et al. 2005, Turk 2009, Yin & Ma 2009, Zutshi & Sohal 2004
Employee awareness	Boiral 2007, Darnall et al. 2000, Melnyk et al. 2003, Morrow & Rondinelli 2002, Newbold 2006, Rondinelli & Vastag 2000, Turk 2009, Vastag & Melnyk 2002
Stakeholder involvement	Delmas 2001, Mohammed 2000, Tien et al. 2005

2.2.4 Shortcomings

Like for PMSs, there are also multiple things that can go wrong for a management system, as the high failure rate illustrates. Interestingly, the research of de Waal & Counet (2009) indicates that the main problems are no longer really in the implementation-, but the utilization-phase: among the top five issues (according to practitioners), there is only one aspect purely related to the implementation phase – management putting low priority on the system implementation – ranked third place, and not listed at all by academics. More pressing are a lack of performance management culture in the organization as well as management commitment, followed by perceiving insufficient benefits from the system and a too low priority of the system. The list is complemented by academics who view insufficient information- and communication technology as most crucial, followed by the organization being in an unstable phase (for an overview of the complete top ten problems see ANNEX 4). In all those cases, the problems came from within the organization, however the flaw can also lie in the systems, that often require too much time and financial investments, have too many and too complicated measures, misleading signaling, and are too mechanistic as well as too monotonous and therefore discourage entrepreneurial intuition (Martinez et al. 2010).

A critical point that finds more and more attention is the role of performance measurement and management in inter-organizational activities. One reason that is likely to bear responsibility is that even though more holistic approaches have begun to appear, the interrelationships among systems are still neglected (Demartini 2014: 3), often only focusing on logistics control systems

(Folan & Browne 2005). The criticism in relation to suitability for supply chain points out that there is (Shepherd & Gunter, 2005):

- no proper connection with strategy;
- a preference of cost over non-cost indicators;
- an unbalanced perspective, e.g. insufficient consideration of customers; and
- no real system thinking.

Due to this criticism of PMMSs for supply chains, it is worth having a closer look at inter-organizational management.

2.2.5 Inter-organizational Management

Besides the problems between PMMSs and management across the organizational boundaries, the area of inter-organizational management, there are two more reasons to have a closer look at it. Firstly, the research question is embedded in the context of the food value chain which makes a better understanding of the concept necessary. Second, the subsequent chapter on sustainability will show, how significant and efficient actions require coordinated approaches along the whole chain of production.

When talking about inter-organizational management, it generally refers to managing the supply chain of an organization. A supply chain can be defined as a set of at least three entities (including individuals, e.g. final consumer) that are directly involved in any kind of flow – up- or downstream, material (goods) and/or non-material (services, finances, information) – from a source to a customer (Mentzer et al. 2001). Supply Chain Management (SCM) is often a source of confusion among practitioners as well as academics, as some perceive it in operational terms involving material flows, others as a management philosophy, and some as a management process (Tyndall et al. 1998). For this paper, SCM will be defined as:

"The management of upstream and downstream relationships with suppliers and customers in order to deliver superior customer value at less cost to the supply chain as a whole." (Christopher 2011: 3)

This definition already includes the term value and indeed, in the last years it could be observed that the term value chain was increasingly introduced and used in an almost identical context. This was experienced by the author throughout the case studies, therefore the terms value chain and value chain management will be used subsequently, although the definition stems from SCM. This particularity can be explained by having a closer look at the original concept of the value chain introduced by Porter: it was developed as a strategic tool to determine the competitive advantage based on the activities and the way they are organized *within* a firm (1985: 33). As FIGURE 2 illustrates, there are two types of value chain activities, primary and support activities, that a firm has to perform in a unique way or more efficiently than their rivals in order to gain competitive advantage.

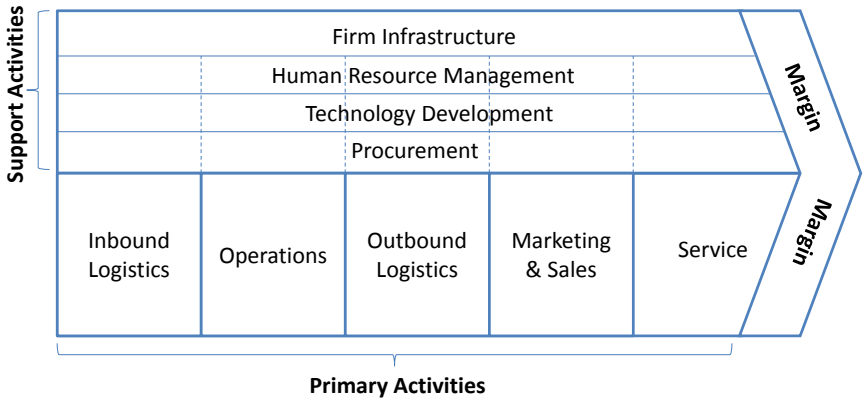


FIGURE 2 Porter's Original Value Chain (Porter 1985: 37)

Organizations need to assess each activity and should there be other actors capable of performing the task more efficiently, the activity should be outsourced, forming a partnership. Thus, many former in-house activities are nowadays outsourced and extend beyond the organizational boundaries; as a consequence value chains are getting increasingly global and hence more complex (Christopher 2011: 11), requiring firms to measure and manage performance across their value chain. For performance measurement, the attention generally lies pretty much entirely on the logistics control system and consequently fails to create a more holistic picture, e.g. the efficiency of interaction among firms or benchmarking different supply chains (Folan & Browne 2005).

So despite the economic logic of outsourcing along the value chain, a question that needs to be asked is whether it is possible to achieve a better performance as the chains get increasingly complex. One of the main reasons for the complexity is that participants of the chain might not share the same goals, what is good for one actor might harm the objectives of another one, therefore hindering the sharing of information between the chain partners (Aramyan et al. 2007), crucial for proper managment of the chain. So in a nutshell, better value chain performance does not necessarily equal better individual organizational performance. In addition, agri-food value chains are even more complex due to certain challenges that other type of chains do not face, e.g. (Aramyan et al. 2007): perishability of products, long production throughput time, seasonality and dependence on natural conditions, special transportation and storage requirements, and high external pressure.

A popular framework that was developed by the Supply-Chain Council (2014), is the SCOR model, which advises performance measurement throughout the six key processes of each chain: plan, source, make, deliver, return, and enable. Taking an additional step, Aramyan et al. (2007) created an integrated framework over all these processes suggesting efficiency, flexibility, responsiveness, and food quality as the key categories to measure. It is quite interesting that only the last category is agri-food chain specific and that all in all, four categories seem to be sufficient. Whereas sharing information,

communicating clearly, acknowledging mutual benefits, and cooperating closely, increase the chances of the chain to do well (Bowersox & Closs 1996), organizational performance improvements rather depend on a successful integration, particularly internally (Flynn et al. 2010), meaning the alignment of value chain and organizational performance measurement and management. Once properly integrated, value chain management can help to improve the quality of data and reporting, customer satisfaction, and internal operational as well as financial performance (Ou et al. 2010). Again, an overlap with previous findings becomes apparent.

2.2.6 Organizational Development

Two aspects of performance measurement and management were only indicated so far, but not addressed explicitly. Due to their importance – also for sustainability initiatives – the role of change and learning shall be touched upon subsequently. Despite the vast literature on the topics, a brief introduction and analysis of their relation to performance will be sufficient for this research, considering its resource limitations.

2.2.6.1 Change

The objective to improve a company's performance will logically involve some sort of change. Brignall (1992) finds performance measurement to be one key agent to that change. He is supported by Amarantunga & Baldry (2002), who consider the lack of suitable performance measurement as a potential barrier, not only to change but improvement consequently. Biticti et al. (2000) point out the dualism of the relationship between PM: whereas PM can initiate change and ensure achievements to be maintained, it also needs to be sensitive to changes in the internal and external environment and has to be adapted accordingly. This ongoing development of the measurement system is generally necessary for it to remain efficient and deliver relevant results that can be actually used; because irrelevant results are unlikely to be used, and results that are not used will be dismissed (Neely 1998). But the ability to actually utilize the delivered results is crucial to drive change and thus move from performance measurement to performance management (Amarantunga & Baldry 2002). In a nutshell, measurement needs to be part of the change that it can initiate.

Looking at Kotter's eight steps to transforming an organization (1995), a popular framework for change – even though or maybe because it is rather practically than academically oriented – creates some quite interesting findings. He suggests the following steps:

1. *Establishing a sense of urgency*
2. *Forming a powerful guiding coalition*
3. *Creating a vision*
4. *Communicating the vision*
5. *Empowering others to act on the vision*

6. *Planning for and creating short-term wins*
7. *Consolidating improvements and producing still more change*
8. *Institutionalizing new approaches*

Indeed, there are several similarities to the previous findings on performance measurement and management. Several of the steps aim at creating internal motivation, amongst others *Establishing a sense of urgency* and *Planning for and creating short-term wins*. *Forming a powerful guiding coalition* once more highlights the importance of management commitment. Creation and communication of a vision, probably will not only raise employee awareness, but also facilitate a proper alignment with strategic objectives. Finally, *Empowering others to act on the vision* creates employee and other stakeholder involvement. A conclusion that could possibly be drawn from this overlap is that it shows, how deeply change is actually involved when it comes to performance measurement and management. Nevertheless, it should be mentioned that Kotter's theory lacks academic validation and is no guarantee for success (Appelbaum et al. 2012).

Kotter's steps highlight the importance of addressing the 'soft' issues: in many cases, barriers to change actually mainly stem from social and psychological circumstances that have not been adequately considered (Hoffman & Henn 2008). In terms of measurement systems this means that the reason for the implementation to fail might not be rooted in the system itself, but because the organization was not receptive. Lozano (2009) points out that the major obstacles to change are:

- *on the organizational side*: a lack of strategy and long-term planning, a lack of top management commitment, and bureaucracy; and/or
- *on the individual side*: a lack of communication, a lack of trust, a threat to one's job status.

Lozano (2013) categorizes the barriers into five groups: *Managerial* (in terms of leadership), *Organizational* (structure and alignment), *Supportive* (provided or denied support to employees), *Historical* (previous change attempts), and *External* (behavior of external stakeholders, e.g. pressure from competitors or regulators). He goes on to criticize that organizations either do not identify their barriers, or even if they do, they do not develop adequate strategies to overcome them. This goes hand in hand with March's observation: "Organisations are continually changing, routinely, easily and responsively, but change within them cannot be controlled arbitrarily. Organizations rarely do exactly what they are told to do" (1981: 563). So improving the circumstances that influence the success of an implementation becomes essential. Given the importance of these 'soft' issues, it becomes clear that organizations are individual and thus also do not all have the same receptiveness for change.

2.2.6.2 Learning

Previously, it was mentioned that PMMSs should enable an organization to control the effectiveness of the chosen strategy and indicate the reasons, should

it not work. This feedback mechanism thus allows companies to learn from the results of past activities and to modify their strategy accordingly. In short, PMMSs can equip organizations with the ability to learn and create a learning culture (Amarantunga & Baldry 2002), which can change behaviors and enhance an organization's competences (Fiol & Lyles 1985, Huber 1991, Slater & Narver 1995).

Organizational learning has also consistently been considered a foundation in achieving competitive advantage and key variable in performance improvements (Fiol & Lyles 1985, Brockmand & Morgan 2003, Santos-Vijande et al. 2012) as firms that possess the ability to learn are generally more flexible and therefore react faster to changes in their internal or external environment than their competitors (Day 1994, Slater & Narver 1995). There are also several case studies that found proof that an organization capable of learning usually tends to perform better (e.g. Baker & Sinkula 1999, Keskin 2006, Santos-Vijande et al. 2012). Furthermore, the actual learning process that can be distinguished in *Information acquisition*, *Information dissemination*, *Shared interpretation*, *Declarative memory*, and *Procedural memory*, seems to have a positive influence on firm performance both at the individual stage level (Tippins & Sohi 2003) as well as for the whole process considered (Darroch & McNaugton 2003). Zheng et al. (2010) also discovered that knowledge management can take the role of a mediator, a finding that will still be of interest later on. However, Jiménez-Jiménez & Sanz-Valle (2011) point out the inconclusiveness of most results, as the samples, as well as measurement of learning and performance, strongly vary. The authors furthermore found that organizational learning has a stronger impact on innovation than performance, suggesting that learning is rather enabling innovation, which then improves the organizational performance. Even though the exact way performance is influenced is debated, the key take away should be that PMMSs can facilitate organizational learning and ultimately has a positive impact on performance.

The previous chapter gave an insight into the concept of performance management within and across organizational boarders. The actual link to sustainability and the implications sustainability has on performance measurement and management will be addressed subsequently.

2.3 Sustainability

Sustainability and its meaning are fiercely debated, as well as its impact on the business world. Whereas some simply see it as a marketing tool, others perceive it as a chance to radically change their way of doing business. Though there might not be a one-size-fits-all-answer to what sustainability means to each and every business, the author will subsequently provide an overview of the most popular definitions and clarify, which one suits best to this specific research. Afterwards, it will be investigated how the ongoing sustainability trend influences and challenges performance measurement and management.

2.3.1 Definition

Going back in time, a starting point for the concept of sustainable business can be found in forestry. Already in the early 18th century, the term sustainability was used in the instructions of von Carlowitz (1713), advising to cut only as much timber as can be reforested. Almost 300 years later, the famous Brundtland-Report sentence might be viewed as initiator of the modern sustainability-rush: "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (WCED 1987: 24-25).

Possibly through the frequent utilization, the sentence is nowadays often cited – sometimes adapted to the firm's perspective referring to stakeholder needs (Dyllick & Hockerts 2002) – but its precise meaning rarely questioned. However, there are a couple of things that need to be said about it: First of all, the sentence is not so much a definition, but rather describes the consequence of acting sustainably. Second, the introduction of 'needs' has a rather blurry effect, as most humans would probably provide different answers to the question what they need, and that is already simplified, as there is actually no limitation to *human* needs; thus, the sentence becomes rather philosophical. In an attempt to make this rather vague concept more applicable to the business world, the triple bottom line (TBL) approach was introduced (Elkington 1994), presenting the three dimensions economy, ecology, and society that should be considered in order to operate sustainably. Unfortunately, facing 'social sustainability' instead of 'needs' did not necessarily make things easier, as authors find that with "respect to social sustainability, however, a lot of confusion has to be acknowledged" (Becker et al. 1997: 18) or consider it a "concept in chaos" (Vallance et al. 2011: 342). One explanation could be that regional and cultural differences make it almost impossible to find common ground.

A concrete definition of sustainability is given by Daly (1990a), stating three conditions that a sustainable society needs to meet:

- the rates at which renewable resources can regenerate should not be exceeded by the society's rate of use and consumption;
- the rates at which at which sustainable renewable substitutes are developed should not be exceeded by the society's rate of use and depletion of non-renewable resources; and
- the rates at which the planet can assimilate waste should not be exceeded by the society's rate of pollution production and emission.

However, it should be noted that Daly himself also defended the Brundtland-Report sentence and found benefit in the lack of a precise definition of sustainable development, as it "allowed a considerable consensus to evolve in support of the main idea that it is both morally and economically wrong to treat the world as a business in liquidation" (1990b: 32). In addition, the sentence also manages to highlight the concept of inter- and intra-generational equity by referring to

meeting the needs of present as well as future generations, which is in a way implicit in Daly's definition, but certainly not as clearly outlined.

2.3.2 Role of Sustainability in the Corporate Context

The concept of sustainability received increasing interest in the corporate context in the beginning of the last century. Clark addressed the social side of corporate sustainability – without using the term – as early as 1916 and notes that "if men are responsible for the known results of their actions, business responsibilities must include the known results of business dealings, whether these have been recognized by law or not" (1916: 223). Only in the 1950s, Bowen coined the term of corporate social responsibility (CSR) in the business world and defined it as "the obligation of businessmen to pursue those policies, to make those decisions, or to follow those lines of action which are desirable in terms of the objectives and values of our society" (1953: 6), which led Carroll (1999) to name him 'father of CSR'. The rapid economic growth and industrialization after the Second World War led to a shift that laid more focus on the environmental side and concerns about ecological conservation, which became popular after Meadows et al. publication in 1972 that wondered about the limits to growth in their report of the same title. Stakeholder theory also started to be more involved as the connection between society's interests and environmental pollution was made, e.g. Goyder (1961) proposed social audits so that stakeholders had a platform that allowed them to influence corporate behavior. The stakeholder pressure on companies, in particular through NGOs, increased in the 1990s as concerns over the effects of globalization intensified the existing ones over social and environmental damages (Kolk 2003). This moved organizations to defend themselves by disclosing their business practices beyond mandatory financial reports (KPMG/WIMM 1999, Krut & Moretz 2000, Line et al. 2002). Given the elementary role of food and its multi-faceted impacts on society (e.g. in terms of nutrition or environmental pollution), agriculture is under particular scrutiny (Aramyan et al. 2007) and developed defensive mechanisms early on (Park & Seaton 1996), which is best illustrated by the vast amount of certificates and standards that can be found in the sector nowadays. Reports mostly rely on the TBL structure, which is also propagated by organizations such as the Global Reporting Initiative (GRI) and the Dow Jones Sustainability Index (DJSI) in an effort to bring more uniformity and transparency into the growing numbers of publications. It could be argued that - similar to criticism on PMMSs – isolating each dimension will fail to capture the holistic picture, after all, sustainability is greater than the sum of its parts. The concept of TBL is admittedly the easiest to grasp and apply, however, one consequence that can be observed is that in reality, it is no so much three equal pillars, but the economic bottom line that predominates (Steger et al. 2007). This might be another reason why the TBL is favored, as this kind of sustainability can be achieved by adding two additional dimensions – social and environmental – on top. The strategic component of sustainability efforts was thus often limited to marketing, initially to prevent or repair reputational damage and later on to corporate branding. Yet, managers

showed increasing interest to integrate sustainable practices into their business beyond external communication which motivated Wilson (2003) to term corporate sustainability as a new management paradigm, as he considers it an alternative to the growth and profit-maximization model. This paradigm was supported by the business case for sustainability that assumed that efforts in each dimension are mutually reinforcing and therefore create a win-win situation for the private sector and society, best illustrated by the popular example of an increased eco-efficiency: using resources more efficiently leads to cost reductions for companies and less environmental impact (Dyllick & Hockerts 2002). This view has been challenged lately, e.g. Hahn et al. argue that

> "Given the complexity and multi-faceted nature of sustainable development, […] trade-offs and conflicts between economic, environmental and social aspects in corporate management and performance represent the rule rather than the exception" (2010: 218).

Therefore, the subsequent section will analyze the relation between sustainability and organizational performance in more detail.

2.3.3 Sustainability and Organizational Performance

When considering which circumstances actually create a competitive advantage and thus ultimately affect the performance, unfortunately, academics can only offer limited answers to the question. If external or so-called pull factors are to be considered, it can be said that – as so often – it depends, in this case on the stakeholders: whereas studies focusing on the development of stock market performance of firms showing environmental efforts generally found a positive relationship (Hart & Ahuja 1996, Klassen & McLaughlin 1996, Jacobs et al. 2010), those examining the customer-side could either discover no willingness to pay more for sustainable products (Anstine 2000) or even a negative relationship (Luchs et al. 2010). In other words, shareholders seemed to appreciate at least environmental efforts, but consumers valued more sustainable products less. In line with stakeholder theory (Freeman 1984, Donaldson & Preston 1995) both examples can be assigned to the *Market-drivers* category, *Environmental regulation* and *Societal values and norms* completing the set of external drivers. Values and norms in society in particular became an increasingly powerful element applied by NGOs, the media, or any group of stakeholders that shares the same norms and values (Wheeler et al. 2003).

On the other hand, there are also factors from within the organization that can impact on their performance. First of all, as already mentioned several times before, the strategy aspect: Sustainability principles need to be integrated into the overall firm-strategy and decision-making (Schaltegger & Burrit 2000, Labuschagne et al. 2005) and not simply be perceived as an additional aspect to the core strategy, a common behavior among firms (Etzion 2007). An internal force that was listed previously as well, is the organizational culture that seems to be especially influenced by: *Management commitment* (Lopéz-Gamero et al. 2009); *Communication*, meaning the timely and precise circulation of information (Sharma et al. 1999, Lenox & King 2004); an *Adequate time-horizon* (Schaltegger & Hasenmüller 2005), since concentrating on short-term profits is contradictory to

the concept of sustainability (Dyllick & Hockerts 2002); and finally, *Motivation*, which can be driven by legitimation, competitiveness, and responsibility (Bansal & Roth 2000). It is compelling to see that, except for the time-horizon, all influences were mentioned previously. In relation to motivation, it should be emphasized that competitiveness and legitimation are external drivers, whereas a feeling of responsibility is rather internal, and as was pointed out in the chapter on Standardized Management Systems, internal motivation appears to create stronger benefits. Since this was also valid for EMS, it might be concluded that this could apply to sustainability engagement as well. The final internal force that plays a role, are *Adequate resources*. Whereas physical property such as specific technologies and machinery are certainly of advantage, what is of high importance for successful sustainability initiatives are intangible assets such as knowledge and skills (Huang & Shih 2009, Melville 2010).

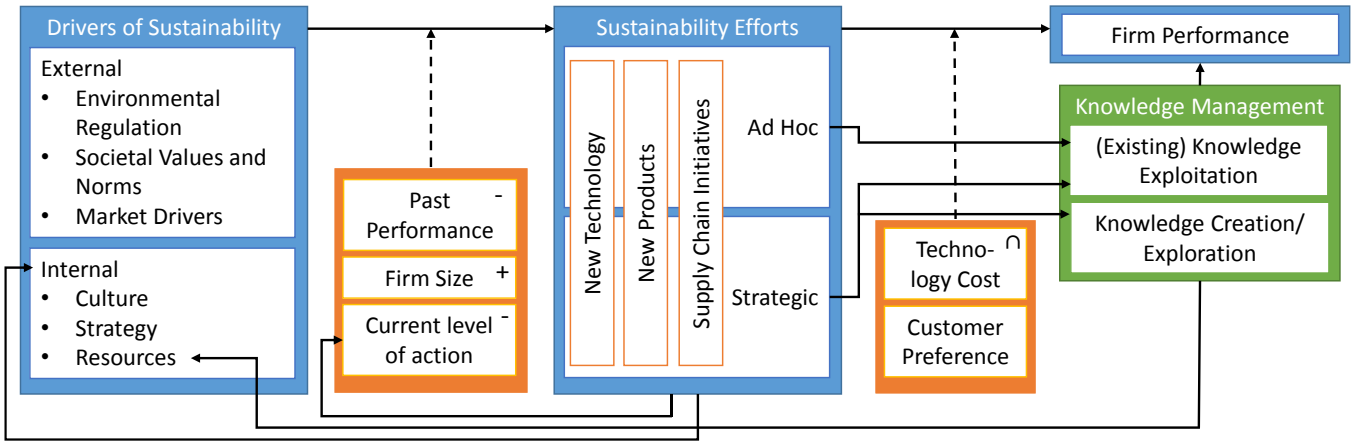


FIGURE 3 Conceptual Decision-Making Framework for Firms in the Context of Sustainability Drivers Firm Performance (adapted from Schrettle et al. 2014: 76)

Schrettle et al. 2014 processed this data into a conceptual framework, which is presented in FIGURE 3. It should be noted that the authors created the framework for manufacturing firms and considered sustainability only from the environmental perspective, as they presume economically sound behavior in all business activities and find the environmental dimension to have an impact on society. Though only conceptual and thus limited in its generalizability, it certainly sums up the previous finding in an illustrative way. There are some intriguing relations which shall be highlighted subsequently. The main message is clear, internal and external forces drive organizational sustainability efforts – either ad hoc, so rather designed for a specific task, or strategic and thus more general and long-term oriented – which impact on the firm performance. The influence can be positively or negatively moderated (orange boxes): interestingly, successful past performances and a high current level of sustainability action appear to slow down the sustainability ambition as firms feel less urgency, which relates to Kotter's (1995) change model. The cost for implementing new technology pays off performance-wise, but apparently only to a certain point,

hence the inverted u-shape. What has already been mentioned for organizational learning is knowledge management acting as a mediator (green box) in this framework: whereas ad hoc efforts can be conducted by using existing knowledge, more radical strategic efforts require to expand the knowledge, which then directly creates new resources. Initiating a feedback mechanism, sustainability efforts also positively influence internal drivers.

2.3.4 Challenges of Sustainability Assessment

Much has already been said about performance measurement and management, and the introduction of sustainability did not really change organizational performance, it simply expanded the horizon – by adding the environmental and social dimension in terms of the TBL – and thus, the previous lessons still apply. Therefore the focus will be laid on analyzing what new challenges arise or how existing ones are amplified. Initially a few clarifying words on measurement and assessment: though they are two concepts that often go hand in hand and the terms are often used interchangeably, the processes can differ. Poveda & Lipsett (2011) consider the identification of sustainability related variables followed by the according data collection and analysis as the measurement process, whereas assessments entail a judgmental element, such as weighting. As it is commonly applied in the subsequent cases, the term sustainability assessment will be used henceforth, which Hacking & Guthrie (2008) simply consider as a process that directs decision-making towards sustainability. Although quite generic, two aspects of this short definition should be pointed out. First of all, the process-character that highlights the continuity and second, the directing of decision-making, which clearly marks it as a strategic instrument that goes beyond marketing purposes. A similar development can be seen when looking at the background of sustainability assessment, which largely evolved from environmental impact assessment, but more recently it is rather connected to strategic environmental assessment (Sheate et al. 2001, 2003).

The major problem – for academics as well as practitioners – that already shone through previously, is the sheer complexity of sustainability, characterized by "pluralistic goals, ambiguity, uncertainty, emergence, and context dominance" (Searcy 2009a). As a consequence, many corporations already struggle to find a common understanding of what sustainability means for them (Searcy 2009b). The strong contextual dependence of sustainability issues also means that there is no one-size-fits-all-solution, meaning that each sustainability program requires a rather custom-tailored approach (Searcy 2009b), as geographical differences need to be considered (Schrettle et al. 2014). This ultimately leads to the question whether a regional approach should be chosen that can produce a more significant result for the individual case but lacks comparability, or a more universal approach that is less resource intensive as it does not require individual customization (Fraser et al. 2006, Agger 2010). Even if this matter can be decided, many organizations seem to lack the skills and knowledge to effectively move towards sustainability, which might be caused, amongst others, by missing appropriate educational and training programs,

failing to grasp the holistic impact of sustainability, and a lack of authority (Searcy 2012). What furthermore adds to the complexity is the collision of systems, indicated by pluralistic goals and what has been previously considered as the trade-offs and conflicts between the three dimensions. Additionally, firms do not only have to consider two new pillars, but they need to do this on a whole different time-scale, shifting from short-term profit maximization to long-term profitability. The dimension simultaneously expands horizontally (what to consider) and vertically (timeframe to consider). Especially regarding the social dimension, the question of what to consider gets further complicated by how to assess it: objectively or subjectively as perceived by the part of society in the assessment (Diener 2006, McCrea et al. 2006). Subsequently it also needs to be decided whether sustainability can or should be captured quantitatively or qualitatively. Another dilemma related to the temporal dimension is that assessments tend to measure past and present states, whereas sustainability is future-oriented (Magee et al. 2012), a critique that was already mentioned in the context of performance measurement. Another previously noted aspect that comes into play are inter-organizational relations that need to be taken into account as well. Although it is a well-known and researched concept, inter-organizational management is not oriented to sustainability (Skjott-Larsen & Schary 2007). Compared to the social side, much experience could already be gathered regarding the environmental dimension. Yet, it still poses many challenges for organizations and much of the potential of operational tools and instruments remains locked as firms fail to completely integrate the environmental supply chain management into corporate management dynamics (Iraldo et al. 2009). Furthermore, actual proof of sustainability initiatives being economically viable has been rather scarce to date (Tonelli et al. 2013: 42), even though it could be argued – at least in theory – that this is not a necessity when pursuing an improvement of the sustainability performance.

2.4 Conclusion of the Literature Review

The literature review shows that there are common points between performance measurement and management (PMM) and sustainability, but also new challenges as well as aspects that seem contradictory. Furthermore, the link between these two topics is still largely unexplored and does not allow any generalizations yet. The field of PMM in particular has been intensely researched, and though it is demonstrated that there is no one-size-fits-all solution for successful implementations, knowledge about factors influencing the process grows continually. Interestingly, the relevant aspects for both, performance measurement and management are quite similar, as the overlap between the works of Bourne et al. (2005) and Ferreira & Otley (2009) showed. The overarching structure provided by Bourne et al. (2005) – internal and external context, content, and process – will therefore be adopted and form a part of the subsequently developed themes.

The link to sustainability assessment still receives surprisingly little attention in terms of research, which might be due to the relative novelty of the topic compared to PMM, as well as its enormous complexity. Though no conscious effort has been made so far, to put the lessons learned from PMM into the context of sustainability and more precisely sustainability assessment, this literature review has focused on providing an overview and highlighting common themes among the research fields. Even though sustainability assessment is considered to stem largely from environmental impact assessment, it can be concluded that eventually, sustainability assessment could be PMM taken to the next level. To further stress the overlap, the work of Schrettle et al. (2014) will be synthesized with the findings of Bourne et al. (2005), resulting in common themes:

- *External Context*: Both works agree on the importance of context. Concerning external circumstances, both list regulatory as well as economic factors as influences. Understandably, social norms and values additionally play a major role for sustainability projects.
- *Internal Context*: Again, there are several similarities, in particular concerning characteristics of the organization, such as size and culture. Other aspects highlighted by both are the integration into and alignment according to existing strategies, the importance of providing adequate resources, plus past experiences.
- *Content*: There is also an overlap in terms of content, as the authors of both researches point out the importance of defining what to measure and design the system accordingly, depending on the scale of the effort, single aspects or dimensions in an ad-hoc approach, or a holistic strategic approach.

Many of the previously listed themes can also be found in other parts of the literature review. Furthermore, there are also some additional common themes that have not been listed so far:

- *Knowledge Development*: A part of resources, still it appears several times. The theme is called knowledge development instead of management to also include unorganized and unplanned learning processes, which overall can enhance an organization's competences (Fiol & Lyles 1985, Huber 1991, Slater & Narver 1995).
- *Motivation*: Though strongly related to context, motivation is added as a stand-alone theme, as it continuously appeared throughout the literature review as an important factor, e.g. Llopis & Tarí (2003) pointed out that the type of motivation influences the kind of impact on performance.
- *Implementation*: The review showed that efforts related to the implementation of performance measurement and management suffer a high failure rate. De Waal & Counet (2009) found furthermore that a successful implementation, does not lead to a successful utilization due to a number of factors, e.g. a lack of commitment.

Though much overlap exits, there is also some conflict between the theories, which becomes especially apparent in the following themes:

- *Process*: The main aspect stressed throughout the literature, especially for PMM, but also when it comes to sustainability initiatives, is the individual alignment of the process to an organization's need, not only in terms of data capturing but also interpretation and evaluation. However, this decreases comparability and increases the required resources (Fraser et al. 2006, Agger 2010).
- *Concepts*: Otley (2006) points out that it is elementary to be aware of the role of a system, yet the review exposes a drastic difference between PMM and sustainability, especially in agriculture. Whereas PMM is intended as a proactive, strategic tool, the recent boost in sustainability initiatives was originated as a defense mechanism, reactively protecting the organization against reputational damage (Park & Seaton 1996). Though a trend towards a more active approach can be perceived, it is still largely marketing-oriented. Furthermore, many organizations struggle to find a common understanding of what sustainability means to them (Searcy 2009b).
- *Externalization*: Sustainability, in particular in combination with PMM can certainly be challenging. Following Porter's value chain argumentation, it might therefore make sense to leave the sustainability assessment to an external provider who possesses the knowledge and experience and is therefore able to conduct it more efficiently. It was also pointed out that the external certification of an EMS can have positive influences, such as serving as additional encouragement (Melnyk et al. 2003). However, this directly contradicts with the demands of *Process*, as externals provide solutions that are at least to some degree standardized and might lack the necessary insight into the client organization to create a suiting assessment.

It becomes clear that there are certainly many parallels between PMM and sustainability in the literature, as well as potential for conflicts, not only within the themes, but also between each of them. Though Schrettle et al. (2014) make several suggestions on the relationship between individual factors on firm performance, these have not yet been tested in practice, particularly not in the context of sustainability assessment. Several other questions have also been raised during the course of this review that existing literature could not answer. To provide some more clarity and results regarding this particular phenomena, a case study approach, examining the experiences around two actual tools, will be conducted to provide an adequate support to the objectives of this research. The details concerning the case study approach are introduced in the next chapter including a presentation of the two tools involved.

3 METHODOLOGICAL CHOICES

The subsequent parts will serve to present the choices that were made in order to best answer the research question. First, the chosen research strategy is going to be justified, followed by the presentation of the cases used within this research. Afterwards, the research design will be explained to allow others to follow the development and achieve a clear understanding of the findings and the derived conclusions.

3.1 Case Study

This part will first explain and justify why a case study approach was chosen as research strategy and secondly why this was done for two specific tools available on the market in order to point out this study's reliability and validity. Afterwards, the individual tools will be briefly presented and the rationale behind the choice be pointed out.

3.1.1 Rationale for Using a Case Study

Though both performance measurement and management were already thoroughly investigated from theoretical as well as practical angles, the question to be analyzed in this research has so far barely been touched to the knowledge of the author. Sustainability assessment and according methods have started to appear in the literature, however, the findings were not connected to the potential impact it could have on an organization, particularly in the context of the food industry and from the perspective of PMM. Thus being a rather new or, more precisely, unexplored phenomenon, it appeared reasonable to select a research method that allows intensive research, the case study format being an appropriate research strategy for the task.

3.1.1.1 Features

When it comes to defining a case study, it is generally strongly linked to qualitative research, and might in some cases even be used as a synonym. Several understandings associated with a case study could be identified:

- only one case is selected, though analyzing multiple cases is accepted as well (Bryman 2001; Stake 2000);
- the study is a detailed and intensive examination (Bryman 2001; Piatt 1988);
- the research problem is studied in its context (Cresswell 1998; Holloway & Wheeler 1996; Robson 2002; Yin 1993,2014);
- multiple data collection methods are used (Creswell 1998; Hakim 2000; Holloway & Wheeler 1996; Robson 2002; Yin 1993,2014).

These points give a good idea of what constitutes a case study, but are not really a distinctive definition yet. Lewis (2003: 52) suggests to define a case study as featuring a "multiplicity of perspectives which are rooted in a specific context" with the perspectives stemming from multiple data collection methods or from a single data collection method from people with multiple perspectives. As such it is particularly useful for investigating in a real-life context (Yin 2014). In regards to the selection of the case, Mitchell (1983) proposes that researchers should pick the event or series of related events that they believe to exhibit the process of a general theoretical principle. A very rich understanding of one or several cases can be achieved through these features, one of the main advantages of case material (Amaratunga & Baldry 2001), which can also bring new discoveries (Shaughnessy & Zechmeister 1990). Therefore, it enables the researcher to study the process of sustainability assessment and put it in an organizational context. As the resources for this research were limited, the case study was also an appropriate choice, as the smaller number of sources posed a manageable data collection task.

3.1.1.2 Criticism

However, a case study approach also faces certain criticism and challenges, the most prominent ones being a lack of objectivity, as the researcher might have made too many subjective decisions, and insufficient data to allow the findings to be generalized.

Addressing the first point, the danger of the researcher's bias falsifying the results and neglecting to test findings is certainly real. Bromley (1986) finds that the subjective judgments of the researcher impact on the internal validity of the data. He is supported by Becker (1986) who suggests that the bias of researchers may lead them to draw conclusion that lack reliability. However, Berg (2001: 231) points out that the problem of bias is present in qualitative as well as quantitative research and that the decisive aspect is generating reproducibility: documenting the research process in a way that other researchers can recreate the same findings.

The second point of critique that is often brought forth, is the case study approach's limited capability to produce findings that can be generalized (Remenyi et al. 1998, Yin 2014): though case studies can lead to the proposition of new theories, their findings will not be universally applicable (Amaratunga & Baldry 2001). However, this should also not be the objective of any case study. Instead it should be applied in order to explore and understand unique situations (Berg 2001: 232).

Several steps can be taken that are explicitly thought to address and overcome these points of critique, which will be highlighted subsequently.

3.1.1.3 Overcoming the Criticism

In order to increase a research study's overall acceptability, Yin (2014) developed a series of design criteria that should be fulfilled; these also address the two earlier mentioned points of critique:

- *Construct validity*: The right operational measures need to be selected for the examined concepts.
- *Internal validity*: Causal relationships need to be explored and demonstrated.
- *External validity*: The degree to which the study results can be generalized needs to be established.
- *Reliability*: The repeatability of the study process needs to be demonstrated in order to allow other researchers to establish the same findings.

It should be pointed out that these criteria do not prescribe a specific level that should be achieved in any of the categories, but instead they highlight the necessity of proper documentation of the research process making the steps that led to the outcomes traceable. TABLE 4 provides an overview of the four tests and corresponding research strategies.

TABLE 4 Research Design Criteria for Judging the Quality of the Research Design (Yin 2014: 45)

Tests	Research strategy	Phase of research in which tactic occurs
Construct Validity	Include multiple sources of evidence	Data collection
	Establish chain of evidence	Data collection
	Have key informants review draft case study report	Composition
Internal Validity	Do pattern matching	Data analysis
	Do explanation building	Data analysis
	Address rival explanations	Data analysis
	Use logic models	Data analysis
External Validity	Use theory and replication logic	Research design
Reliability	Keep a case study protocol	Data collection
	Develop case study database	Data collection

The following measures were taken to fulfill Yin's criteria:

- *Reliability*: In order to facilitate reliability, a simple case study protocol was kept during the data collection phase (see ANNEX 5) to guide the researcher through the interviews. Furthermore, a case study database was developed containing transcripts of the interviews and field notes of the researcher taken during and after the interviews.
- *External Validity*: To increase external validity, the information was gathered not only for one specific tool, but data about experiences from a second tool was also gathered, independent from the first, thus following Yin's replication logic. As the context was expected to have a strong influence, findings that are similar across both tools can be perceived as more robust. At the same time, contradictory findings allow more precise conclusions and help to construct theories. Developing the theoretical background from the literature review before conducting the interviews also supported the construction of more robust cases.
- *Internal validity*: This research analyzed the primary data by doing a cross-case synthesis to find out how different themes are present in individual entities, within cases, as well as between the cases, and how they are interrelated. By using a software that is specifically designed for this kind of task, which is introduced later on, the robustness of the internal validity could further be strengthened.
- *Construct Validity*: Finally, the validity of the research construct was secured by discussing the main findings with experts from the field that have been strongly involved in the development and conducting of sustainability assessment methods for several years. Furthermore, triangulation could be achieved by deriving information from different sources with different perspectives - including those of the second tool – the so called data source triangulation (Denzin 1984, Yin 2014). One could moreover reach triangulation by having several investigators examine the same phenomenon or interpret the same results from different points of view, or by utilizing several methods (Denzin 1984), e.g. conduct a quantitative follow up survey after the qualitative interviews. Though these methods might have further increased the research's validity as the different types are assumed to compensate each other's weaknesses (Amaratunga & Baldry 2001), limited resources prevented such an extended approach.

3.1.2 Presentation and Selection of the Tools

Many different tools are nowadays available to conduct a sustainability assessment in the agricultural sector. They differ in numerous aspects, e.g. the dimensions captured, the method of measurement, the scope, the pricing, or whether they are conducted first hand or by an external provider. Subsequently, the two tools involved in this research are presented and it is explained how and why they came to be selected. For confidentiality reasons, this chapter

unfortunately cannot go into too much details on the tools but will provide a general overview and focus on aspects that are relevant for other parts of this research.

3.1.2.1 Central Case: Tool 1

3.1.2.1.1 Overview of Tool 1

Tool 1 was developed by a globally operating chemical company as a method to measure and assess sustainability in agriculture. Building on the experience of its precursors that only focused on capturing either the social dimension of sustainability or the relation between economic and ecological outcomes, it was originally initiated to compare different systems of agriculture against the background of overall sustainability performance. Although the goal was to provide scientifically based arguments for or against different ways of agricultural production to the public – illustrating the earlier mentioned role of context regarding sustainability efforts – it is now freely provided as a service to clients to enable them to start continuous improvement of their sustainability performance. Stakeholder dialogues were held throughout the development phases of the tool. Having a Life-cycle Assessment (LCA) logic, its system boundaries are flexible, which means the tool can be adapted to the individual requirements of the task. Most of the required data is provided by the client using spreadsheets that are designed for the purpose. The goal is for a study not to last more than two or three months in total. The report including the results, the interpretation of the results, and the recommendations is generally critically reviewed by an independent third party before the final report is delivered to and discussed with the client.

3.1.2.1.2 Rationale for Tool 1

There are several reasons that led to conducting this research with Tool 1 as the central case. The primary reason is of rather practical nature and related to the call for extensive knowledge: the author worked for the company that developed and offers sustainability assessments with Tool 1. During the several months of his employment in the sustainability department, the idea for the research task emerged and evolved, also through discussions with supervisors and co-workers, and ultimately led to the decision to continue the employment for the time of the research. Obviously, this affects the bias of the author, however, it should be highlighted that it was him who initiated the research idea out of personal interest. Having spent some time within an organization that provides sustainability assessment furthermore gave him a valuable practical impression of the operational aspects as well as access to years of practical experience in form of his colleagues. Some additional measures were taken to further reduce the bias: First of all, the research is published and therefore available to everyone, not keeping the findings within one organization. Second, the published version is anonymized and thus does not bring any reputational advantage. Last, the

research also includes interviews with experts from another assessment method. This provides an additional perspective, independent from the employing company.

Another circumstance made it intriguing to have a closer look at Tool 1: a business assurance company developed a sustainability standard that includes parts of Tool 1's methods in order to conduct an initial product-focused sustainability assessment. Unlike Tool 1, the implementation of the standard does not stop after providing the recommendations but is more extensive, as it also includes the actual management of sustainability in form of a system and the client receives a certificate. Therefore it could provide some insight regarding the design of Tool 1.

A third reason for studying Tool 1 was the fact that it is an established tool, building on years of experience. As mentioned, the organization developed and started using a first predecessor as early as 1996. The focus back then still lay on quantitatively measuring ecological aspects in proportion to a product's cost-effectiveness. It was later on complemented by another analysis that considered the social pillar, which is captured quantitatively as well, thus differing from many other methods, in particular certifications that generally rely on qualitative measurement. Based on those experiences, Tool 1 was developed and has measured and assessed sustainability performances in agriculture for about three years. All three methods are validated by independent assurance companies.

Besides and possibly through this solid backbone of experience, Tool 1 appeared to be the most holistic as well as flexible approach to assessing sustainability in agriculture, based on the findings of a benchmarking of various assessment methods, frameworks, and certifications in the agricultural sector conducted by the author during his previous employment. However, this is not necessarily an advantage. To the best knowledge of the author, there was no other assessment method being that extensive and complex, and consequently time- and resource-intensive, which might limit practitioners' receptiveness for the tool. Furthermore, the nature of the research question required some time to pass between the initial assessment and questions about potential impacts and changes. These circumstances plus the fact that sustainability assessments were still rather new, including Tool 1, significantly limited the number of available studies that were conducted with Tool 1 and contributed to the decision to include another tool.

3.1.2.2 Secondary Case: Tool 2

3.1.2.2.1 Overview

Tool 2 was developed by the agricultural department of a University of Applied Sciences more than ten years ago. The first version was created due to a practical problem of a large food processor that wanted to consider all three dimensions of sustainability on the production level of a farm. As no adequate tool was available, it was developed and applied as a cooperation between the two actors. The second version was then developed by the department, taking a more

structured approach and engaging several stakeholder groups. The method focuses on the farm, meaning that the farm-gates are the system boundaries. However, specific customer demands to include certain off-farm activities can be considered and fulfilled. The necessary data for the assessment is gathered during one 3-to 4-hour interview, the results can usually be presented the next day and are discussed with and explained to the farmer with the objective to trigger continuous improvement of farm sustainability. It is generally used on a fee-for-service basis; it is also possible to purchase a license, which includes training to enable the licensed organization to conduct assessments on its own. More than 1.000 assessments have already been carried out worldwide.

3.1.2.2.2 Rationale

Several reasons led to the decision to not simply include another assessment method, but specifically Tool 2, with the objective of increasing the overall quality of this research. The main reason was certainly that the characteristics of Tool 2 complement those of Tool 1 very well. Whereas the content of both tools in terms of indicators (and thus what gets measured) is quite similar, as is the final presentation and illustration of the results, there are some major differences when it comes to the actual way the methods work. Whereas the data collection phase in Tool 1 is a very time-consuming process that can stretch over several months, the information that is required by Tool 2 can be gathered within one interview session that generally lasts around 3-4 hours and results can often be presented the day after the interview. Another time-relevant aspect was the fact that Tool 2 has a rather fixed scope on the farm, whereas Tool 1 can include the whole supply chain, which in turn requires individual modelling. Furthermore, Tool 2 also has a different business model than Tool 1, as it is possible to buy a license for the tool, allowing the license holder to conduct assessments on their own. Tool 2 has a different background, too, as it was and is still developed by a university and not an industrial player. However, it should be noted that universities of applied sciences need to finance around 80% of research and service activities through third-party funds in the country, and thus, there is also a strong entrepreneurial interest. Unlike the mostly Excel-based Tool 1, the latest version of Tool 2 includes a software, that can either be locally installed and operated offline, or as a web-application. TABLE 5 summarizes the previously mentioned major differences between Tool 1 and Tool 2.

Since Tool 2 already operated since 2000, it had a large body of experience. As the assessments can be conducted faster than with Tool 1, more than 1.000 farms had been analyzed, which also made it very interesting in terms of available data sources. The number of studies that were suitable for this research were strongly limited with Tool 1 and the perspective of clients could only be gathered through the observations of the corresponding advisors. As Tool 2 had more clients, access to two of them could be negotiated, including the multinational corporation that was actively involved in the original development of Tool 2.

TABLE 5 Comparison of Main Criteria of Tool 1 and Tool 2

	Tool 1	Tool 2
Background	Developed by a globally operating chemical provider, based on two precursors. Holistic system for sustainability quantification.	Tool 2 is an indicator- and interview-based method for holistically assessing the sustainability of farm operations, developed by the agricultural department of a University of Applied Sciences.
System Level	Supply chain (consumer, retailer, processor, and farmer) Data mostly provided by client during the data acquisition phase Flexible in terms of scope.	Farm level, environment in general. Can be used by food production industry Assessments are based on surveys (questionnaires) compiled by farmers Flexible System boundaries only regarding the scope of some parameters that can be modified to include off-farm activities.
User/Target Group	Large Producers & Processors, Umbrella Associations	Farmers, Consultants for agriculture and development
Ext. Accreditation	Yes	No
Operating since	2011	2000 (since 2011 version 2.0)
Data Acquisition Form	Excel spreadsheets	Own software, can be locally installed and used offline or online via a web-application
Timeframe	Best case: 2-3 months	Fast (Interviews taking 3-4 hours), no measuring, results can often be presented the next day.

Finally, existing channels to the team of Tool 2 enabled the author to get directly in touch with the project coordinator. That made it possible to negotiate access to relevant sources of Tool 2 rather quickly. Contacting other tools would most likely have proven too time consuming for the limited timeframe of this research.

3.2 Research Design

The following chapter will deal with the design choices of this research. First, the epistemological foundation will be laid, introducing the underlying philosophical perspectives. Second, the global research approach will be presented, including a short overview of the data sources. Afterwards, the main research method that serves to answer the research question is introduced and justified. All subsequent choices were made keeping the research task in mind, how a more strategic impact of external sustainability performance assessment could be achieved in the context of the food sector. In order to find an answer to the main objectives of this research, the earlier presented themes were derived from the literature and the research method designed accordingly to ensure the relevance and completeness of the response to the research question.

3.2.1 Paradigm

Like everything else that involves human activity, this research is no exception to the consequent subjectivity. Since subjectivity clearly affects the research, but also gives researchers a better understanding of their own research design (Easterby-Smith 1991), it is worth giving some thoughts to paradigms that can be defined as "a way of examining social phenomena from which particular understandings of these phenomena can be gained and explanations attempted" (Saunders et al. 2009: 118). Guba & Lincoln (1994) even consider finding a paradigm that is applicable to the research question more important than the selection of the actual research method. Of particular interest for this research were realism and constructivism, two approaches strongly debated within the community of social scientists.

3.2.1.1 Realism

Easterby-Smith (1991) suggests that within this approach, reality exists independently of human influence, however it is interpreted through social values and believes and therefore subjective and often named interpretative approach; following this logic, the observer himself also becomes part of what is observed. The realism approach generally rather addresses a specific phenomenon, trying to comprehend and explain it, instead of investigating external influences or fundamental laws (Easterby-Smith 1991, Remenyi et al. 1998), which can never be completely understood (Guba, 1990). Furthermore, it also acknowledges the bias of the researcher. However, mainstream realism tends to be surprisingly unsuitable for 'real' problems of social science, as the validity of findings and deducted laws come from a closed system (Bhaskar 1978), which can rarely be created in social sciences. Even if it were achievable, other researchers might struggle to reproduce the findings. Acknowledging this aspect, critical realism developed. It opposes offering predictive validity to statements and furthermore takes internal as well as external influences into consideration (Mir & Watson 2001), thus making it highly relevant for this research. Nevertheless, there is still another developed philosophy that should be briefly introduced.

3.2.1.2 Constructivism

Mir & Watson (2001) agree that constructivism and critical realism have much in common and are to some extent compatible. Still, they certainly cannot be used interchangeably and TABLE 6 contrasts realism and constructivism from their point of view. The probably most striking point of the two authors is the opposition to realism's believe in one reality that can be explained once sufficient knowledge is obtained. Constructivists on the other hand trust in the importance of context that can create multiple realities. This philosophy reflects the author's view most adequately and seems most appropriate for this research as context could be identified as one of the main themes during the literature review.

Although they are intrigued by the increasing use of constructivism in the field of performance management, Busi & Bititci (2006) point out two major challenges that come along, as extensive knowledge needs to be required a priori: First of all, regarding the sources of knowledge that are to be selected and secondly, the question which of the many perspectives involved in performance measurement and management are to be considered. However, the author feels confident that both challenges can be overcome, as he not only conducted an extensive literature review, but also gained practical working experience in the field before performing the research. Plus he discussed the topic with several experts from the field. This experience also enabled him to select the disciplinary perspectives that he saw most fitting.

TABLE 6 Realism vs. Constructivism in Strategy (Mir & Watson 2001: 1171)

	Realism	**Constructivism**
Nature of observed reality	Partial, but immutable	Socially constructed
Role of manager	Reactor, information processor	Actor, generator of contexts
Nature of strategic choice	Boundedly rational response to contingencies	Ideological actions of sub-organizational interest groups
Organizational identity	Overt, singular	Multiple, fragmented
Theories of measurement	Replication as a key to accuracy	Context as the key to perspective

3.2.2 General Research Approach

As was discussed in the previous chapter, the strategy considered to provide the best answer to the research question is a case study. Tool 1 was selected as the central case, the findings being supported by the secondary case of Tool 2. The data collection procedure was the same in both cases. This approach enabled the researcher not necessarily to generalize his findings, but rather to benefit from the flexibility and intensive examination (Bryman 2001; Piatt 1988) and thus develop more of a pilot study for this relatively new phenomena with findings that are robust enough to serve as a basis for future research.

This research is explanatory in nature, as it aimed to study the causal relationship between two variables (Saunders et al. 2009), more precisely the relationship between the concepts of performance measurement and a new context, sustainability assessment in the food industry. Nevertheless, when selecting a case study strategy with its ability to generate a very deep understanding (Amaratunga & Baldry 2001), it is also possible to bring forth new discoveries (Shaughnessy & Zechmeister 1990). Therefore, this study is of exploratory nature, too, as it additionally intended to get a better understanding of "what is happening; to seek new insights; to ask questions and to assess phenomena in a new light" (Robson 2002: 59). Whereas this research followed a deductive approach built on existing theory in the beginning by matching identified themes from the literature to the primary data, it was also designed to

leave space for new findings in accordance with its exploratory nature, which was largely relevant when connections between the themes were made.

The data gathered for this empirical research stems from secondary as well as primary data. Secondary data was accessed during the phase of reviewing the literature as it enabled the researcher to extend his mostly practical knowledge by learning about the theoretical aspects of the topic in question. Due to the novelty of the phenomena, the author had to select the relevant aspects of the literature for this research, but received support through discussions with his supervisors. The objective was to provide an overview of the different fields of research involved and create a solid foundation for the following analysis. Furthermore, gathering and organizing the secondary data also helped the researcher to identify the main themes of the research as well as establish new relationships between his ideas (Saunders et al. 2009). Most of the secondary data was found in articles published in academic journals and books, which were accessed either through the library of the University of Jyväskylä or WU (Vienna University of Economics and Business), using the according institution's search engine or Google Scholar. In addition, some secondary data was also found in reports and studies that were freely accessible through the internet, such as general information about the tools on the website of the according organizations, or trends in the agricultural sector or sustainability, analyzed and published by major consulting firms. However, considering that the examined phenomena was relatively new, secondary data alone was not sufficient to provide an adequate answer and required primary data directly linked to the research. This was done conducting semi-structured interviews, which will be explained more in detail subsequently.

3.2.3 Primary Data Collection Method: Semi-structured Interviews

As previously illustrated, this research was more of a preliminary approach to the unexplored phenomenon of the relation between performance measurement and sustainability performance. Several measures were taken and previously discussed to ensure the validity and relevance of this research. In order to further investigate the topic, a qualitative primary data collection was necessary. The method will be presented here, before introducing the interview partners, their role, and their perspective.

3.2.3.1 Semi-structured Interviews: Description and Rationale

Semi-structured and in-depth/unstructured interviews are frequently found in qualitative research (Arthur & Nazroo 2003, Saunders et al. 2009) and therefore often called 'qualitative research interviews' (King 2004). The act of talking to people enables the researcher to comprehend other perspectives (Burgess 1982) and is important for social sciences as language has the power to bring meaning to any aspect (Hammersley & Atkinson 1995).

The understanding whether an interview is semi-structured or unstructured/an in-depth interview differs, which might be explained by the

various models that exist for each type of interview (Arthur & Nazroo 2003) and the related circumstance that researchers individually design the interviews to best suit the need of their work. For this research, collecting qualitative primary data through semi-structured interviews meant conducting the interviews with certain themes derived from the literature review in mind, which were to be covered during the conversation. However, there was no strict sequence in the order the questions were phrased. The questions were also slightly adapted depending on the course of the interview as well as follow-up questions added in order to explore or clarify aspects that arose.

As was discussed previously, the phenomena examined in the course of this research was rather unexplored and the importance of the context was also highlighted. Designing the interviews in a semi-structured manner provided enough structure to ensure relevance to the topic, while leaving enough flexibility to react to the different perspectives and contexts and receive answers that are as complete as possible, achieving sufficient depth (Legard et al. 2003). All interviews were conducted via telephone or Skype, except for one face-to-face meeting, as the interviewees were spread over different countries. The conversation was generally opened with a short introduction about the author and how he came to examine the topic in order to create a sense of familiarity. The initial question referred to the role of the interviewee in the process, further questions were then determined by the participant's response (see ANNEX 5 for more details in the Case Study Protocol). The questions were developed based on the literature review as well as the author's previous experiences and cover general information about the involved organizations, internal and external context, and the content of the tools, as well as the process itself, all of it over the different phases of the projects.

3.2.3.2 Interview Partner

All in all, nine interviews were conducted, five with people having different roles regarding the central case Tool 1, four with participants involved with the secondary case Tool 2. Each interviewee was first contacted via email to ensure their general interest. The initial email always contained a two-pager providing a short overview about the author and the research topic, as well as the subjects that were to be covered during the interview (see ANNEX 6 for an exemplary version); it also ensured the participants about their anonymity in the research report. Therefore all organizations and people involved in the research were anonymized. Each interview lasted about an hour and was recorded with a recording device. In addition, the researcher took some notes during and after the interviews of certain aspects he found to be of particular interest. Afterwards, each interview was transcribed, which resulted in 8-14 pages per interview (see ANNEX 7 for more details on the interviews). The interviews were conducted in the same sequence as listed subsequently. For confidentiality reasons, the original names were replaced with planet names.

3.2.3.2.1 Tool 1

The interviews conducted for the central case involved four people that worked on different levels directly with Tool 1, and one person that worked with it indirectly, parts of Tool 1 being used in the standard her company developed and provided. Unfortunately it was not possible to negotiate direct access to clients, the experiences were captured through the observations of the providers.

- Sun: The interview focused mainly on one specific study that was conducted as one of the first ones with Tool 1 overall and the pilot study in South America. The interviewee worked as advisor in the study and therefore had much contact with the client, a large agricultural producer. Besides his experiences, he also shared an evaluation form filled in by the client at the end of the project.
- Mercury: The interview partner had a similar role as the previous one, acting as an advisor and therefore main contact with clients, but in Western Europe. The conversation was mainly about her experiences with one of the first cases conducted in Europe with the tool.
- Venus: She had a similar role as the first two interview partners, however, she neither worked with Tool 1 directly, nor for the provider, but for a business assurance company that developed and supports the implementation of a standard that includes parts of Tool 1 in the initial assessment. She shared her experiences with clients that implemented that standard.
- Earth: Being the team leader for sustainability assessment, he could provide a broader perspective, being involved in and coordinating most studies, especially in the initial and final stage. Therefore he was also knowledgeable about the evolvement of the tool and its application. He provided support by reviewing the research at several stages, too.
- Mars: Similar to the previous interview, he was also involved in all the latest studies, his main responsibility being the actual modelling. He could therefore share his experiences with clients from a more technical side and also give insight into developments that took place over time.

3.2.3.2.2 Tool 2

The secondary case involved one interview with the main project coordinator of the tool, which allowed a comparison of the results across different tools as previously described. Furthermore, it was also possible to negotiate access to several clients of the tool and compare the experiences that could only be captured indirectly for Tool 1 with firsthand information.

- Jupiter: Working for an organization that focuses on smallholder projects and was one of the first to purchase a second version Tool 2-license, his perspective was broader, as he also had more of a coordinating role.

- Saturn: Working for the same organization as Jupiter, the interview partner was more directly involved in the field work of various projects and thus could provide a more practically oriented perspective.
- Uranus: Being the project coordinator for Tool 2, he had extensive experience and a broad perspective on the topic of sustainability performance assessment in agriculture in general. He furthermore gave an insight into past as well as planned developments of Tool 2.
- Neptune: Having the role of coordinator of activities with Tool 2 within his company, a multinational food processor that was involved in the development of the tool as well, he could provide a general overview of how his company applies the tool and their general experience with. In addition, he also shared some case-specific examples.

3.2.4 Analysis

The primary data from the interviews still required an adequate analysis to provide an answer to the research question that is as complete as possible. This is central to developing an understanding from the case studies (Amaratunga & Baldry 2001), yet, it is the hardest and least systematized aspect (Eisenhardt 1989), keeping in mind that the overall objective of every research should be to allow others to reproduce each step. Indeed, given the unstructured form of and the different perspectives within the qualitative data, this task truly required attention. For this research, a cross-case synthesis was chosen, using NVIVO 10 for Windows – a program for Computer Assisted Qualitative Data AnalysiS (CAQDAS) – for the coding process. Coding – annotating codes to specific words or text passages according to their nature and significance for the research task – is used on the one hand, in order to identify the previously found themes within the qualitative data, and on the other hand to discover new themes and patterns. During the coding process, a coding system was established, which supports the researcher in managing and ordering his data, a task complicated by the non-linear nature of the qualitative analytic process (Silver & Lewins 2014: 16, 158). The development of the coding system will be presented subsequently, after giving a short introduction to CAQDAS.

3.2.4.1 CAQDAS

CAQDAS differs from quantitative content analysis or text mining techniques as it goes beyond mere counting of words and phrases or relations among words (Silver & Lewins 2014: 21). As Yin (2014) points out, CAQDAS tools support the researcher in compiling, disassembling, and/or reassembling data, but the actual work still needs to be done by the researcher. He finds that expectations regarding these tools are strongly inflated due to people's experiences with computer assisted quantitative analysis. Certainly, the use of software can produce a more rigorous analysis, however, this stems from the aspect that the software supports the user in working more attentively and methodically (Bazeley 2007).

Therefore, it makes sense, when the developer team of NVivo, does not promise more than providing researchers with tools that support them in analyzing qualitative data in five principal ways (QSR International 2014):

- managing data;
- managing ideas;
- query data;
- graphically model; and
- report from the data.

Besides offering flexible and sophisticated tools, another reason to select NVivo was of financial nature, as it provides a fully functional free trial of its latest version for 30 days.

3.2.4.2 Developing the Coding System Using Cross-case Synthesis

Even though the final technique used was a cross-case synthesis, several steps were taken beforehand to allow a holistic answer. All steps had in common that the data was first analyzed individually, matching relevant text passages to the previously identified themes in the coding process, followed by a search for relations, alterations, similarities, or anomalies (Silver & Lewins 2014: 158) that enable the creation of a bigger picture.

First of all, each interview was looked at and analyzed individually, using the transcriptions and notes of the researcher, to achieve the highest degree of familiarity possible. This covers with Yin's (2014) demand that each case entity has to be able to form a stand-alone study in its own right. At first, a broad-brush coding (Silver & Lewins 2014: 169) was conducted to see how well the initial set of codes reflecting the theoretical focus applied to the primary data. The themes were found rather frequently as could be expected, since the semi-structured interview was based on the theoretical framework. Creating a code matrix (see ANNEX 8) that provided an overview of overlaps between the themes, provided some interesting findings even without looking at the content of the data by showing how frequently a code was assigned to a section of text on the one hand, and the overlap to the other codes on the other hand. For instance, out of 22 times External Context was assigned, it overlaps 16 times with Motivation; the same goes for Internal Context, which overlaps with Motivation 31 times out of 60 appearances overall. Having formed a first impression of themes that seemed to overlap frequently, the data was analyzed once more to see if sections had been missed or misinterpreted during the broad-brush coding, e.g. sections being coded as Motivation without Internal or External Context were re-analyzed to see if any context-relevant text had not been identified during the first coding round.

Secondly, all data gathered for the central case, including the data from the individual analysis, were compared with each other to identify the relationship of themes across the entities. To get a clearer image, the results of the code matrix for Tool 1 were transferred to Microsoft Excel 2010 for Windows. After

calculating the average overlap for each theme, those that were above-average were highlighted in green (see ANNEX 9). In other words, it became visible which themes overlapped particularly often. The same procedure was then repeated for the secondary case (see ANNEX 10), because the researcher can reduce the risk of drawing premature or even wrong conclusions by looking at the data from many different perspectives (Amaratunga & Baldry 2001). Lastly, the findings of each entity and each case were compared with each other (see ANNEX 8) to see how the appearances and overlaps of themes differed.

It should be noted that the previous process was not done in order to develop arguments for or against certain findings. The amount of interviews conducted and data gathered, especially for the supportive case, was not sufficient to provide any significant results in that direction, e.g. there are much more overlaps with External Context in Tool 2 (four, compared to only two for Tool 1), but it is easily explicable by the low number of overall appearances of External Context in Tool 2 (see ANNEX 9 & 10). However, this process supported the author to highlight certain trends and connections between the themes and thus enabled him to ultimately develop a model depicting the relations and interaction between the themes.

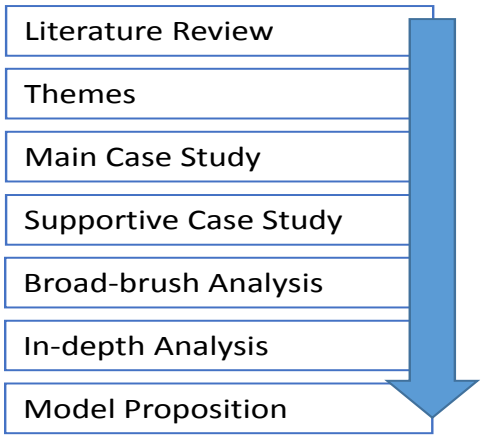


FIGURE 4 Overview of the Research Strategy

In a nutshell, this research started with the literature review and developing a solid knowledge base around the topics relevant for this research question. A set of themes was developed based on the findings from the review and practical experience. Afterwards the central and secondary case studies were conducted. The primary data was then analyzed as described above, coding known themes and identifying their relations. Finally, having reviewed the primary data over all themes, the author's findings are going to be presented in the next chapter. As previously noted, the research was no linear process, but had rather fluid boarders with some going back and forth. Nevertheless, FIGURE 4 provides a rough overview of the research strategy.

Certainly, the methodological choices that were made for this research task also impose certain limitations on the results and their quality. However, a more relevant and critical analysis of these limitations will be possible by pointing them out at the end of this report in the context of the actual findings (Saunders et al. 2009).

4 FINDINGS

This part will present the findings made throughout the gathering and analysis of the primary data, as described previously. As mentioned, it contains information from individuals of the client as well as the provider side with different roles concerning the assessment of sustainability performance and thus different perspectives. Therefore specific information will also be assigned to according individuals to show potential agreement or conflicts across different backgrounds. This chapter will start with an introduction of the findings concerning the previously identified themes. Though there is a strong interconnectivity among the themes which implies sections being coded under several themes, the author will present them as separately as possible at first, putting the focus on providing an overview of the interviewees' perceptions. The second step will then be to really explore and highlight the interconnectivity and relation between different themes.

4.1 Themes

This chapter focuses on introducing the themes identified in the literature review as they could be found across the interviews in order to explain their role in the sustainability assessment process. They are ordered according to their number of appearances, External Context being the code that was assigned least and Process being the code that was assigned most often in total (see ANNEX 8).

4.1.1 External Context

Most of the material that was coded under the theme of External Context was related to external stakeholders and communicating with them in one way or another. This observation could be made across the different tools and all interviews. Interestingly, in case of Tool 1, the external context was also relevant for the provider, which can be explained due to their background and business model. The influence of external context on the client side varied, which might depend on the position in the value chain.

4.1.1.1 Tool 1

For instance, for Tool 1, Sun reported that for their pilot case, they organized a press conference, presenting the results of the case to a larger group of journalist. The interest in the case was quite high from the public side, as *the tool is more holistic and business oriented than the tools or the works that we can see*, usually such projects were rather academic and/or focused on issues related to climate change, such as greenhouse gas emissions. The project was furthermore acknowledged as a success story, amongst others by the national ministry of Agriculture in a publication and won several prices. As the client was one of the country's largest agricultural producers, the strong public interest is understandable. Sun also mentioned that, in general, clients were interested in anything related to sustainability for reputational reasons: *I would say that sustainability is in the media, is in the focus, they have the perception by the public that if they are not sustainable, it would be a problem.* In his opinion, this motivated clients to participate in these kind of projects *[…] even not knowing exactly what it is or understanding sustainability […].* Mercury also pointed out that her client *[…] wished to have the results by October, to be able to add the result to their […] sustainable development report of the company that is produced in December.* She also admitted that it was often more of a communication than an improvement tool, especially if the client was not a producer but further downstream the food value chain. Earth confirmed that also his company as provider, had – at least initially, when developing the tool – this interest to use the tool in order to communicate to the public and consequently to influence opinions. He explained, however, that later on it was more and more the demand for a decision support system that drove the development of the tool. In the interview, it was also highlighted that clients perceived such an assessment as something special, and therefore wanted to communicate about it, sometimes before even starting the assessment. He furthermore stated that a lot of pressure to produce more sustainably came from the clients' clients along the Food Value Chain (FVC) that have their own internal and external context.

4.1.1.2 Tool 2

The findings for Tool 2 were quite similar, especially when it came to communication for reputational reasons. For instance, Uranus pointed out that the marketing orientation was still dominating for many clients when it comes to sustainability assessment projects, even though companies might not say it that way. Neptune agreed that his company publicly communicated about sustainability assessment projects, nevertheless, it was not their main purpose and that such projects were not intended to create product specific claims. An interesting point that was still brought up by Uranus, was that the national regulatory context often strongly limits the possibility of farmers to act, in particular in central Europe. In his opinion, the strongest initiators were usually public aid programs or pressure coming from other actors of the FVC.

4.1.2 Implementation

Regarding aspects that seemed to directly affect the implementation of recommendations, the findings across both tools largely agreed and complemented another. On both sides it was reported that there were cases in which the implementation failed, often because of a lack of knowledge and wrong expectations regarding the outcome of the tools. Furthermore, the strong relevance of different contexts was named repeatedly.

4.1.2.1 Tool 1

Sun and Mercury both reported that in their specific projects, they had some doubts whether there was a clear understanding regarding the outcome of the process. In other words, the clients, at least in the initial projects seemed disappointed or at least puzzled about the final recommendations. It seemed they were expecting something more concrete. According to Mercury this might relate to the complexity of the tool and the topic, as well as the fact that the communication for her project was not in the customer's native tongue. As Earth confirmed *there is often a certain surprise, also a certain helplessness, and therefore we have to focus much more on the topic of implementation.* He added that it was also an issue of the respective agricultural system, which made it extremely complex and therefore sustainability assessments could hardly provide easy solutions, especially to organizations that already produce at a very high level: *If I want to improve my CO2-footprint, there are several opportunities, reducing fertilizers is possibly one of them. But if I reduce to a certain degree and my yield goes down, it can be that this ultimately has a negative feedback on my CO2-footprint […].* Mars also pointed out that much of the actual recommendations that the clients received were not really new to them, especially if they have an agricultural education. However, what they could learn from the tool and the sustainability assessment was how their agricultural system is interconnected, which should ultimately only be the starting point for an agricultural consultancy. The consulting aspect is also what they wished to focus on in the near future, including the development of a separate version of the tool that supported users by providing more insight into aspects the farmer can address firsthand. Venus, whose organization is actively involved in the implementation process, also highlighted the importance of an established management culture and in particular a clear allocation of responsibilities, as well as choosing realistic targets for the implementation.

4.1.2.2 Tool 2

Besides the importance of communicating in the native tongue, which seemed to be increasingly important the smaller the client organization was, the necessity of embedding the results into a consulting package was mentioned repeatedly. For instance, all interview partners from the demand/client-side used the tool and the sustainability assessment in several representative cases of a region in

order to develop a consulting strategy that suited best to the regional needs. According to Uranus the regional context was also significant, as his experience showed that organizations from emerging nations appeared to be most likely to implement measures as they had a certain amount of management skills, a solid knowledge basis concerning sustainable agriculture, as well as financial possibilities on the one hand, and sufficient room for improvement on the other hand, also in terms of regulatory context. He pointed out that the chances of implementation could be increased in developed countries as well by conducting the assessment at the right moment in time when existing organizations were going through transformations anyway, such as a transfer of business. Uranus furthermore reported that he witnessed implementations fail because organizations overestimated their (knowledge) capacities and efforts were designed too broadly to make sense for all their farms and organizations. In the interview with Neptune, this gap of knowledge regarding local circumstances between the consulting entity and the farm were still often considered the main reason why implementations (partially) fail. Similar statements came from Jupiter and Saturn, stating that insufficient resources for the projects then prevented the implementation. In general, Uranus emphasized that there was still an unwillingness or a lack of awareness concerning the understanding of sustainability and sustainability assessment, meaning that it is a highly complex topic, where the assessment and development of solid strategies required proper funding and patience as it was a long-term process. Therefore he also considered real management commitment as the most important factor for a successful implementation.

4.1.3 Concepts

In terms of concepts, the gathered information was quite diverse and covered different areas, still it evolved mainly around the tools themselves, which concepts were directly applied within, e.g. LCA, an holistic approach to sustainability, or more indirectly, such as the purpose of the tools. Regarding the latter, there was a dominating interest perceivable to communicate about the results, whereas the strategic management opportunities often went unrecognized, at least initially.

4.1.3.1 Tool 1

One part that seemed to create at least vagueness for clients, was to which degree the tool could and should be used for marketing purposes. As Sun and Mercury both reported for their projects, using the tool for reputational reasons and gaining public legitimization was a major objective, as became clear e.g. when Sun said *[…] they already had a lot of certification of their product process, so we didn't have impact in terms of certification or to help them with this kind of thing* or *it was important […], they had this expectation in terms of communication.* When it came to actually using the tool in order to improve their sustainability performance, it was often met with a certain lack of understanding that resulted in unrealistic

expectations, as Mercury put it: *And they were all waiting that […] like a magic stick, it answers all their questions.* At the same time, what often created great interest was the quantitative measurement over the life cycle of the product and the ability to identify improvement opportunities in the process, in particular the graphical visualization illustrating the performance. The vagueness was also present on the provider side, as the business model for the tool was rather unclear, at least to some of the interviewees, such as Mercury: *What is the market, what is the objective, what is the positioning, is it just a communication tool to give a flavor of sustainability in our communication, in our discussion with politicians? […] And where is the link with the business?* On a higher level this was clear, as Earth and Mars explained that the purpose of the tool for the supplier was the development of scientific based arguments backing sustainable solutions for agricultural products; the tool itself should initiate the continuous improvement of the client's sustainability process. Consequently, limiting sustainability to CSR-activities was not in their interest. Earth also addressed the question, what sustainability meant for the tool, that it did not distinguish between sustainable and unsustainable, but simulated less or more sustainable outcomes: *[…] can a car in private transport be sustainable, we believe yes, and we believe, that if I inherently improve my system, it improves my sustainability.* One of Mars' answers added to this when he explained that it was important for the success of the tool to have a similar logic of sustainability with the customer or at least, that the customer should not have an understanding of sustainability that conflicts with the tool's concept of it, e.g. when only organic agriculture is considered sustainable. Though there could be agreement on what sustainability means, Earth observed either a strong interest in ecological impact aspects with only little concern for the social side, or a very open mindset in case clients participated in order to get a broader perspective on their business.

4.1.3.2 Tool 2

Also the clients of Tool 2 showed great interest and appreciation for the quantitative measurement and the graphical visualization, as both Jupiter and Saturn confirmed, as well as the holistic approach of the Tool covering all dimensions of sustainability. Neptune found the tool well suited to stimulate an increase of the sustainability performance, too, by furthermore including the management aspect of the farm. Uranus on the other hand felt like overall, there was a stronger interest in the ecological and social dimension, whereas the economic well-being of the farm tended to be neglected or less valued. Concerning the way organizations intended to utilize Tool 2, he observed similar events as for Tool 1: though he found a general increased interest in sustainability assessments over the last years – illustrated by the fact that meanwhile, they were rather addressed by clients instead of campaigning for them – for many organizations, the defensive marketing-oriented approach dominated, even if the companies might not say it that way. However, those organizations that actually had the assessment conducted seemed to use it more and more strategically and in their own interest, meaning issues like securing of resources. He furthermore

also perceived a strong existing sustainability management of clients in terms of sustainability departments critically as it created too much distance, sustainability should rather be deeply anchored in every department. Neptune highlighted that his organization made a clear and conscious distinction between sustainability assessments and certification as the assessment required trust, an aspect that was mentioned for both tools. The fear of not qualifying for the certificate was counterproductive and would prevent trust and thus a full identification of weak spots. Consequently, activities with tool 2 were communicated in the organization's CSR-report, but not in detail or product-specific. As he said, *[…] therefore we are against using Tool 2 in terms of certifications, it is a development tool and shall remain one, otherwise we lose a very crucial element.* This development aspect was also highlighted by Uranus when he told the author that his organizations still intended to focus further on orienting Tool 2 towards consulting and education in the future.

4.1.4 Content

Aspects that were identified as content-related within the interviews circled for both tools largely around the question, whether and how strongly adaptations took place within the tool in order to best fit it to the needs of the clients. Both tools relied on their established content and adaptations were generally only made to better reflect regional or production system-specific circumstances.

4.1.4.1 Tool 1

The main adaptation that took place for each project individually was the determination of the project scope, which became necessary due to the LCA-backbone of the method. Within the scope, the boundaries of the project were fixed, meaning at which point of the value chain the assessment starts and where it stops. In one of Sun's projects, e.g. the production of fertilizers was the starting point and it ended with the transport to a port. In many cases, it was initially still open which product was to be analyzed, too. Furthermore, there were several alterations done so that the assessment would better reflect the regional circumstances. Sun reported that there was a too strong orientation of the tool towards the situation in which it was developed, which required some changes, e.g. concerning local agri-environmental schemes, which were then incorporated in cooperation with and based on feedback from the client. This was similar in Mercury's project. Furthermore, both stated that there was a certain insecurity and lack of knowledge in terms of LCA. Earth acknowledged that *there are few players – those from the high-end side, I would say – who can judge whether the way we go, life cycle assessment, three dimensions, scenario-based, that this is the right way for them.* Consequently, clients often struggled to clearly formulate what they wished to analyze with the study, which made it complicated to determine the best scope, as Mars explained, but clients also relied on and trusted in the expertise of his team to figure this out. Finally, Earth pointed out that even though there were certain adaptations, these were generally only owned to a

small extent to client wishes in order to avoid any greenwashing of the final results. There are cases where the team might have decided to use generic datasets, but only because past experiences enabled them to anticipate that some aspects were not directly impacted within particular project scopes. So overall, adaptations were generally due to regional differences and/or related to the extensiveness to which some aspects were analyzed.

4.1.4.2 Tool 2

A similar way of adapting could be found for Tool 2, even though it had a fixed scope that focused on the farm-level, thus discussions about the scope of a project were largely unnecessary in the past. But as Uranus revealed, there were ongoing projects where they intended to evolve Tool 2 into a more open method in terms of incorporating certain off-farm aspects, e.g. post-harvest, according to a particular client demand, and data acquisition. Unlike Tool 1, Tool 2 gathered its required data through qualitative interviews, which were intended to be more flexible concerning the extensiveness and depth to which an aspect is covered in the future, following the wishes that were expressed by Jupiter and Saturn in their respective interviews. In all the interviews for Tool 2, it was mentioned that alterations where mostly done concerning the questions of the interviews, either to adapt and interpret them to fit regional social customs, or simply drop them in case they were not applicable, e.g. questions regarding husbandry when it was a purely agricultural farm. But as for Tool 1, the idea was rather to simplify topics of the data acquisition instead of completely dropping them. Uranus pointed out that whereas he considered this less standardized option to be the better solution within a consulting and educational context, it was problematic in terms of monitoring and statistics, as the comparability across projects decreased. Neptune also brought up the topic that using a single, standardized version of the tool somehow compromised the applicability and resulted in falsified graphical illustrations, which could be misinterpreted by people that are not as intensively involved with the project or the tool.

4.1.5 Externalization

Sections coded under this theme explore what led the later clients to have the sustainability performance assessment conducted by an external provider and which consequences arose from this decision, positive as well as negative. The major finding for both cases were that clients sought out externals in the field of sustainability assessment for their expertise as they considered the topic to be highly complex. In the best case, this could lead to a more productive process, in which both sides combined their knowledge – about sustainability assessment on the one side and the respective context on the other side. However, the gap of knowledge between the actors and their consequent mutual dependence also led to complications, such as problems in communication, insufficient resource allocations, or differing objectives.

4.1.5.1 Tool 1

Several reasons were stated in relation to clients' rationales for outsourcing the sustainability performance assessment to an external provider, but most important was always the expert knowledge necessary to run such a method. Sun's client for example answered him that their main reason was *the knowledge you offered, that could measure the sustainability in an organization in a comparative way and along the life cycle of the product.* But Earth shared that although the expertise was valued, being an external could result in a lack of credibility. In Sun's pilot project, it was actually his side, the tool provider, that approached the client as they were looking for an organization that also had expert knowledge in their respective field, in that case crop growing, a partner with the ability to understand and judge the process as well as the result of the assessment. Earth explained that this was the usual way in most of the early studies, but that in the meantime they were generally approached by prospective partners. The reasons being that, first of all, the topic itself gained momentum, and secondly, the reputation of their tool was growing. This moreover plays again into the aspect why clients picked that specific company and thus Tool 1, where, according to Earth, the overall company brand reputation was decisive. As noted earlier, most partners did not have an overview of tools on the market or the understanding which tool suited best to them and thus trusted in a well-known actor to find the optimal solution, as Mars confirmed. Therefore the team of Tool 1 always had to develop an adequate knowledge base on the client side in order to figure out the proper scope without confusing the partner or creating too high expectations, which happened several times. Another problem of working with an external partner, at least in case of Tool 1, was the dependence on the other party. Especially in the phase of the data acquisition that needed to be largely provided by the client, this repeatedly led to issues: it simply took too long as the clients were overstrained and did not allocate sufficient resources to the task. This extended the duration of the whole process and led to overlaps with other projects. Another potential problem in choosing such a solution was that the providing organization had its own interests as well and as Mercury said, when there were requests for other projects from within the organization, *depending on who sends a request you cannot say no. If it's coming from the top you don't say no.* Besides, several interview partners pointed out that although they had the expertise in sustainability assessment, they were no experts on the client's particular field of operation and could only highlight certain topics, but not provide concrete solutions. Therefore a lot of communication was required, which took time as well, plus, there were cases where different languages as well as different levels of knowledge in terms of the assessment prevented an efficient exchange to some extent. On the other hand, the differing perspectives could also lead to productive discussions between the actors, as Mars mentioned, and having a broader perspective furthermore enabled the supplier to provide a benchmarking for their clients, so their results could be compared, e.g. with other regional performances. Such a benchmarking possibility was planned in the new spin-off version of Tool 1. Finally, a point that was already brought up earlier,

the problem of adapting the system to best fit to each individual client, which was only possible to some extent. Mercury and Sun both found that some of the method's indicators simply did not fit to the regional context. At the same time Earth warned that client wishes could only be fulfilled to a certain point in order to avoid greenwashing and the process becoming too resource intensive.

4.1.5.2 Tool 2

Unlike described for Tool 1, Jupiter told during his interview that his organization compared several tools available and decided to select Tool 2 as it best suited their requirements. Uranus spoke of another large client that conducted the same selective process and ultimately decided in favor of the tool that was closest to their own systems. Additionally, Neptune reported that his organization was involved in the founding phase of Tool 2. The reason to develop such a tool with an external provider was, first of all, the lack of an adequate method at the time to holistically capture all three dimensions of sustainability. Plus, like many other large companies in the food chain, his organization was not an agricultural producer but further down the value chain and therefore lacked the expertise of an agricultural university of applied science. Still, Uranus stated that many clients did not compare tools before and approached his team simply because they had heard of the tool and were interested in the topic in general. One issue came up during the interview with Jupiter: the team around Tool 2 was embedded in a significantly smaller organization with less resources than the team of Tool 1. Providing the license solution could therefore on the one hand safe resources, but might damage the quality and reputation of the tool, if the people using it were not properly trained. By keeping the exclusive training rights for Tool 2, its organization tried to avoid that danger, but then again had to deal with resource shortages. This limitation also made it more difficult to include customizations, even though these were consciously restricted anyway, as noted earlier. However, such a shortage of resources could also be witnessed by Uranus in client organizations, when they were not adequately prepared and misjudged the required means for the assessment process, often resulting from the gap of knowledge between the two actors.

4.1.6 Internal Context

The findings concerning the role of the internal context refer to a large part to the circumstances within the organizations such as their structure and management, their products, but also their level of development in terms of experience with sustainability. Although it was already indicated for External Context, here, it became really obvious that, due to the externalization, the internal context was relevant for the client as well as the supplier side.

4.1.6.1 Tool 1

Client:
The organizations in which Tool 1 was applied had strongly varying backgrounds from smallholders to some of the largest national agricultural producers, stock traded and with a long history. Nevertheless, almost all experiences derived from dealings with well-organized and structured enterprises, as projects with smallholders were coordinated through umbrella associations as well. Both Mars and Earth also saw this as an important criteria, due to the extensive amount of data required by Tool 1 that needed to be provided by the organizations. Therefore, documents and records had to be accessible. Plus, they also explained that an advanced level of agricultural knowledge simplified the discussions around the scoping. As noted earlier, Venus highlighted that such large bodies with a well-developed management culture were most likely to be able to stem the effort of implementing her organization's standard. All of them agreed that more experienced enterprises also had clearer strategic objectives to conduct a sustainability assessment, e.g. in Sun's case, a better understanding of the processes to optimize them. But as he brought up, this required sufficient resources. Nevertheless, it could also happen that clients struggled with clear objectives. Mercury indicated that in some cases *they don't exactly know why they are doing this, and they don't know what the results can be. So there is a gap between the knowledge […]* of the actors. When it comes to past experiences with sustainability and sustainability assessment, varying backgrounds existed, from clients that only just started out or, as in Mercury's project, with a lot of knowledge from previous consulting and impact assessments. What the majority of cases had in common, was a lack of understanding concerning LCA, which, however, was not that relevant, according to Mars. What he experienced to be more challenging were existing sustainability philosophies and their respective management, which often led to clients questioning the tool's philosophy instead of aiming for understanding and improving the analyzed agricultural system. He found that this was particularly prevalent when dealing with associations, such as the FAO. Earth answered that internal factors that stimulated the cooperation with clients were often related to the leading person's age and attitude: being rather young and having a certain pioneer character was advantageous as sustainability and its assessment was still a relatively new topic. Most of the clients' activities were top-driven.

Supplier:
An aspect that was also brought up, particularly by Mercury, was the role of the supplying organization's internal context: as a large stock corporation, it naturally wished to generate business value. Projects needed to be coordinated worldwide and though this provided a lot of knowledge, conflicts of interest could arise, as past experiences had shown. When several projects ran parallel, those of more interest to (the upper management of) the tool provider were favored, which resulted in failing to stick to other projects' schedules due to

resource shortages. Earth pointed out that as they are selling chemical products, the tool also served to create legitimization for the company, which was a major factor for the initial development of the tool.

4.1.6.2 Tool 2

Client:
Uranus considered the organizational background of their clients as a decisive element as well, because it strongly affected the way the client-organization intended to utilize the tool. Yet, in the other interviews it still became obvious that although their organizations had different backgrounds, the common ground of Jupiter and Neptune was the fact that they operated globally and therefore required a tool that could be applied to assess sustainability performance worldwide. Although Tool 2 did not require quantitative data in the data acquisition phase of the interviews, the way client organizations, more precisely how their sustainability (assessment) efforts were managed, still played a role: driving such measures from within, e.g. the purchasing department, was more fruitful than from a higher management level, as the responsible people could better justify their actions, e.g. in terms of sourcing, in the opinion of Uranus. But he also warned of conflicting interests, particularly in larger organizations with different departments. Therefore, a company-wide commitment to and understanding of sustainability as a long-term, complex process was identified as important as well. As noted previously, the timing to do such an assessment could be crucial as well, since an organization that is in the process of transition could be more receptive to additional changes.

Supplier:
Being a much smaller group and part of a university of applied science, the internal context of the supplying organization of Tool 2 was highly relevant, too, as Uranus confirmed. Like Tool 1, they also had an entrepreneurial interest as most of their funding comes from the projects, which regularly conflicts with the educational background. As a small team they also had to face resource shortages on their side.

4.1.7 Knowledge Development

An important aspect throughout the interviews was the development of knowledge or, as one of the interviewees called it, *capacity building*. Sections relating to this theme could be found throughout all interviews and were relevant not only for the clients, but also for the providers of the tools that go through a constant learning process as well. Whereas the final results often only created little new knowledge and were considered as not concrete enough several times, the clients learned a lot about the practical implications of sustainability for their activities, a process that generally already started at the beginning of the project.

4.1.7.1 Tool 1

One objective of the tool was the creation of knowledge regarding the actual sustainability situation on the field and along the value chain, and as was noted earlier, the expertise in sustainability assessment that tool-suppliers can provide was one of the main reasons for organizations to externalize the project. However, this also means that there was a gap of knowledge between the demand- and supply-side in the beginning, which could lead to problems, as Mercury pointed out. Mars also noted that this imbalance made it difficult for clients to clearly formulate their objectives for the tool and for his team to properly identify the clients' needs. Thus, a lot of groundwork needed to happen in the beginning, including long discussions and preparatory work with the client side, as Mars and his team discovered, which then strongly simplified the initial meetings. Moreover, knowledge was also developed in the final phase, when the results were presented and recommendations given. Although this final presentation may not always have delivered new insights, it gave the clients a good overview of the situation and concretized sustainability in the according context, creating knowledge that way. Sun confirmed that this aspect was appreciated by the clients, especially the clarification of the impacts of different actors along the value chain. Still, they often expected more concrete solutions. However, both Earth and Mars pointed out that the current tool's task was to highlight certain topics and to create awareness as well as a better understanding of what sustainable agriculture meant, what its main drivers and impacts were, and how these could be influenced. Earth said that *many partners and clients see an assessment project at the end [...], we see it at the very beginning and afterwards there is the topic of implementation and that is often difficult to anticipate for our partners […].* Concrete solutions could then be developed with the partner and the team could simulate the results, to make it more tangible. Even though the more practical aspects regarding the assessment of sustainability led to issues, such as delaying projects because the ISO guidelines for critical reviewing were not considered, this knowledge had been established meanwhile, according to Earth, also due to a mutual learning process with clients in pilot projects. What he still perceived as a challenge was the field of expectation management and implementation. To solve the latter one and provide more concrete support and solutions for agricultural consulting, the development of a spin-off version of Tool 1 was initiated.

4.1.7.2 Tool 2

As Tool 1, Tool 2 also led to knowledge creation in two areas: about the actual farm situation and about the meaning of sustainable agriculture in general. Uranus observed that, in general, clients appreciated and acknowledged the results, but were expecting more concrete solutions at times, the major benefit to them was to understand sustainability in a farm context. Neptune and Saturn also confirmed that already the interview stage of the tool initiated a thought process and was often an eye-opener to the farmers, like the results. Another

distinction has to be made here due to the licensing solution of Tool 2: whereas the last observations came from the actual usage of the tool, there was also a strong knowledge development when people received training to use the tool. For instance, Neptune's organization actively used it as an educational tool for their employees, e.g. from the purchasing department, to close the growing knowledge gap to farming activities that become increasingly complex. Uranus pointed that aspect out as well and furthermore added that it also required a certain level of knowledge, to have the same vocabulary, in order to efficiently narrow that gap, which could complicate the work, e.g. with smallholders. But he criticized that too many participants in training session still suppressed the inherent conflict in sustainability, e.g. short-term gains that required barely any resources vs. strategic-long term investments that 'real' sustainability often needed. Tool 2 also went through a process of continuous improvement and evolved based on acquired knowledge from projects as well as feedback from partners, as Jupiter told.

4.1.8 Motivation

The motivation theme sought to discover what drives the different actors, what is their motivation and their intention. As for other themes, it became clear that there is motivation on the client as well as on the supplier side that did not always cover. Clients were still largely driven by external and consequently reputational factors, but in particular larger organizations recognized the strategic potential of the assessment.

4.1.8.1 Tool 1

Client:
Earlier it was noted that many clients were interested in improving their knowledge and understanding of sustainability in the context of agriculture. Sun reported that the primary motivation of the client in his case was about really *[…] identifying improvements in the process, to understand really, what were the impacts that they were causing and to understand how to deal with that […]*, also because they had a strategic objective to produce more sustainably and were looking for consultation and solutions. It became obvious in the previous themes that reputational aspects were a major factor as well, which Sun pointed out for the same project: *[…] as we had a press conference scheduled, so also it was important in terms of, they had this expectation in terms of communication. And I would say that this is something that really worked well for them, because the reputation of this case was so high.* Mercury spoke of similar observations but introduced another aspect: the intentions of how the results of the tool should be put to use also depended on the position in the value chain: *[…] when you are on a food chain level, people are more looking for communication things. So it's not a question of implementation, you have nothing to implement […],* but *the more you go to the producer level, to the farmer level, the more you have to go into the tool […]*. Moreover, pressure coming from the value chain was named as an important driver by Earth, especially for sub-

contractors of retailers for instance. This pressure was also considered a primary motivation of Venus' clients who sought to differentiate themselves from their competition when they implemented the standard. It should be noted that her clients received a certificate at the end and thus the situation slightly differed from the other findings. Earth and Mars found that general interest in sustainability activities had increased over the past years and thus for a tool that could holistically identify and illustrate impacts to support companies' continuous improvement, but as both said, so far not for monitoring purposes, even though a re-evaluation after a certain period would be the best case. Moreover, Earth saw a relation between the current level of a company and how well they succeeded in strategically deploying the sustainability assessment, e.g. clients on the high-end side expected to receive a financial benefit by finding ways to become more efficient.

Supplier:
Even though the sustainability assessment was provided free of charge, the supplier still had an entrepreneurial interest to generate business from their service as was confirmed in all interviews. Earth clarified this and explained that their idea was to create a link between their own and clients' strategies, and also between sustainability and their current and future products. This was well illustrated by Sun, when he told that the cooperation with the original partner for the pilot was cancelled, because the client went through a reorganization and their new business model was no longer attractive for the project. Mercury pointed out that ultimately, their business was still the sales of chemical products. Whereas the tool was originally intended for advocacy, to back up arguments for certain agricultural systems, the main intention shifted to decision support. But it still was important to the provider-side from a communicational aspect, because they used it to connect and discuss about the meaning and implications of sustainable agriculture with different stakeholders, as Mars described.

4.1.8.2 Tool 2

Client:
One of the main drivers was the creation of a better understanding of sustainability as well. Uranus perceived a generally increased interest in sustainability in agriculture and consequently sustainability assessment, too. Particularly for organizations that were only beginning to deal with sustainability questions, the tool allowed to conduct a baseline study serving as a scientific basis for further actions, e.g. developing regionally focused development strategies, as Jupiter and Saturn mentioned. Almost none of the client organizations had used the tool for a latter re-evaluation and monitoring, though it had been planned in a few cases. In terms of communication, Neptune also explained that his organization mainly focused on developing a dialogue for capacity building and creating awareness for the topic with other value chain actors - mainly producers – as well as company intern, and that the tool was only used for CSR-purposes to a very small extent. What was much more relevant to

him, was the aspect of securing the supply for the chain and managing related risks, e.g. ensuring long-term, adequate supply before investing into a new processing factory in an area. Uranus pointed that out as well and added that the monetary aspect was one of the biggest drivers, which is not an issue per se, but he criticized that the conflict of interests were largely ignored, constant growth was expected although resources within the system were limited. Furthermore, he also saw a much larger impact of internally rooted motivation and questioned the long-term success of external factors such as pressure from other value chain actors.

Supplier:
Two motivations to supply such a sustainability assessment tool could be found. First of all, as Uranus openly admitted, there was an entrepreneurial interest involved. This derived from the fact, that as part of a university of applied science, the team had to cover around 80 percent of their expenses through third-party funds. At the same time, they were also motivated to improve the overall sustainability performance in agriculture and provide a high-quality educational tool, thus they created the licensing model with exclusive training rights to reach a broader audience, as pointed out by Jupiter.

4.1.9 Process

The purpose of the subsequent theme was to get a better understanding of the sustainability assessment's process in general and, more particularly, steps that were problematic and actions that were taken to create a more individual alignment to the clients' needs.

4.1.9.1 Tool 1

Overall, the process usually started with 2-3 initial meetings, where the clients were introduced to the tool and the scope of the project was determined. The number of people involved on the client side varied depending on the company size and also the phase of the process. Whereas Venus reported of up to 20 people during the training, Sun's project included people from several areas (farms, HR, management) in the training and presentation phase and only two main contacts throughout the process. When the tool was still relatively new, the clients were selected by the supplier from existing business partners, which required internal discussions before approaching the identified candidates, especially for pilot projects. One of the most critical phases was usually during the data acquisition, which was considered as quite challenging by most clients, no matter their background. Consequently, the part often turned out to be quite lengthy, and as the provider depended on the data for the actual modelling, it then delayed the whole schedule, e.g. Sun's project lasted eight months and Mercury failed to stick to her original project schedule, too. Therefore the supplier side attempted to be more involved, and Earth and Mars announced intentions to focus more on steering and pushing the customers actively through the process. This was

common practice for the standard of Venus' organization. The project length should ideally not exceed three month, as the motivation and interest in the project suffered otherwise, as Earth explained. Therefore close contact was required and since the assessment team was a central unit, the regional supplier units had to be strongly embedded, which Mercury perceived that way, too. In one instance, the project was also cancelled as a client went through a reorganization and could not provide sufficient resources for the data acquisition. Once the data was assembled, the modelling began, in which phase there was usually barely any communication between the actors, which was a point of critique by Sun's client. The results then first went through an internal review before being presented to and reviewed with the client. Based on their feedback, smaller alterations were often made to best fit to the analyzed system, e.g. adaptations to regional circumstances, such as specific soil properties that the client had better knowledge of. As Earth pointed out, the presentation was the next critical step, as they first of all had to breakdown an around 100-page report to only a few essential aspects to communicate to the client. Secondly, it had to be discussed how to implement the results on the farm-level, creating a link between analysis and consulting. This caused problems in the past, when the partners were not adequately prepared during the initial phase. That was a learning process for the Tool 1 team and consequently much more focus was laid on discussions and preparatory work in the beginning of the project regarding the management of the clients' expectations. Earth observed that as a result more time was then also spent on discussing the implementation than the actual results. Mars still saw room for improvement in the initial discussions to figure out what the clients need, which was not always what they wanted, at least in the beginning, having little or no knowledge about the tool's capabilities. To further close this gap between analysis and implementation the spin-off version was initiated to provide more flexibility regarding the topics in focus and thus a more concrete consulting. In general the efficiency of the process improved with increasing experience that allowed e.g. to use generic datasets for certain aspects of production systems, where results could be anticipated, or a better steering of the process overall.

4.1.9.2 Tool 2

All in all, the process of Tool 2 was very similar to Tool 1 with one or two initial meetings, the training of people, the use of the tool, and the presentation of the results. What was different, was, first of all, the license option. Before being certified and receiving an institutional license, which was e.g. in case of Jupiter's organization valid for one year, people had to go through a theoretical and practical training. Furthermore, this training right was exclusive to the supplier organization, meaning licensed bodies could not certify new people. This created situations, where the provider did not have sufficient resources to satisfy the demand for more training. Another major difference was the method of data acquisition, which happened through qualitative interviews in a single session of 3-4 hours. All interviewed clients agreed that this close contact with the assessed

body during the mutual collection of the data had an extensive effect in terms of knowledge development for both sides. However, they also agreed that this effect was strongly dependent on the social skills of the interviewer in terms of developing trust and also reading the context and between the lines. The face-to-face data acquisition then also allowed more flexibility in how the questions were asked to achieve a picture as complete as possible. Since licensed institutions could conduct their own assessments, they often chose a similar approach as Tool 1, to work with umbrella organizations and thus have a farther outreach.

This first part showed that all themes that had been previously identified based on existing literature could be discovered within the interviews and how they are connected to the practices of sustainability assessment across both tools. For the largest part, the findings between the two tools do not contradict each other but point to similar directions, either supporting or complementing each other. One major aspect that had not been addressed in literature, was to distinguish between the supply and demand sides. This part therefore complements the first research objective, identifying the essential elements of an external sustainability performance assessment in the context of the food value chain and getting a clearer understanding of the process. It also largely fulfills the second objective as it illustrates how the actors are involved and affected during the process.

4.2 Interconnectivity

The previous part provided an overview of the different themes and how they emerged during the primary research, in order to develop a better understanding of each of them in the context of this work. As mentioned, every theme was introduced individually; still, it became apparent that there are relations and overlaps between them. Subsequently, the connections between different themes will be analyzed more in detail, to further fulfill the research objectives, a better understanding of the processes of and around sustainability assessments and the way they impact an organization. As was noted previously, the number of cases as well as the amount of interviews per case were limited. Therefore, this chapter will rather focus on illustrating the interconnectivity between the themes in the context of the cases than debating the meaning and significance of the actual numbers.

In order to getter a better overview of the overlap between the different themes, meaning that a section of text from an interview was identified to fulfill the criteria of more than one theme, a query was run using NVivo, more precisely its Matrix coding process, which allows to cross-tabulate the coding intersections for all interviews. Several connections become apparent when taking a closer look at the overlaps between the themes as illustrated in ANNEX 8 - 10. It showed that many significant theme-relations – meaning they appear more often than the average – can be found for both tools, whereas some are unique to one. Themes that are significant in the combined results will be addressed in more detail for each tool, whereas those that are only highlighted in one of tables are only

examined for the respective case. If themes overlap twice, they will be addressed in more detail in the context of the theme that appears more often overall, assuming that the information is more significant, e.g. the overlap between Concepts and Motivation will be more closely examined in the context of Motivation, which was coded 71 times compared to Concepts appearing only 43 times in total.

4.2.1 External Context

Having an average overlap of 27% for both tools combined, themes that appear at least that often are Concepts (45%), Internal Context (27%) and Motivation (73%). Whereas the connection to Concepts and Motivation is relevant for each tool individually, Internal Context (60%) is only highlighted in Tool 2, as is Implementation (40%), which is not highlighted in the combined results.

+ Concepts

Tool 1: Sections that include both themes, External Context and Concepts, generally referred to the interest of organizations to use the tool for marketing purposes to communicate with their external stakeholders, such as in Sun's interview, when he explained about his client that *[…] they had expectations in terms of communication […],* and *[...] the reputation for this case was so high.*

Tool 2: The same goes for Tool 2, where Uranus found e.g. that especially in case of organizations that inquire about the tool and then do not go through with it after all, the defensive or marketing-oriented aspect dominates.

+ Motivation

Tool 1: In this combination, it could be seen how the external context affects the motivation, which was similar in all the cases on the demand side. The organizations, particularly those further down the food chain, according to Mercury, felt a strong need to justify and defend their business actions against other stakeholders, such as their own clients or the public. Consequently, there was a general interest in ways that could be used for external communication. Earth pointed out that this applied to their side as well considering their background as a major chemical seller. He further added that the growing interest of external partners also influenced the development of such a method.

Tool 2: Whereas Uranus did not show any motivation to provide such a tool based on their external context, he agreed that clients often wish to use the tool for communication with external stakeholders. Neptune confirmed that his organization reports about general activities with Tool 2 to the society, yet, as he pointed out, only to a small degree, as they are not interested in using the tool for product-specific claims that would require compliance.

+ Internal Context

Tool 1: Internal and External Context overlapped several times, when clients were insecure about their objectives to conduct such an assessment and/or wanted everything at once and expected the tool to work like a 'magic stick', as Mercury called it. She illustrated that the objective was not completely clear on the supplier side neither.

Tool 2: The results for Tool 2 broadened the picture, Uranus observed that internal and external context were interacting in a different way: the more external regulation a client faced, the more important it was to conduct the assessment at the right moment, e.g. transfer of business.

+ Implementation

Tool 2: According to Uranus, the External Context also affected the Implementation, more precisely he found the currently largest driver to be either financial aid programs or pressure from the value chain. However, he doubted the long-term efficiency of such externally-driven projects.

4.2.2 Implementation

Looking at the combined results, four themes stick out in connection to Implementation as they have an overlap larger than the average (20%). They are Internal Context (24%), Knowledge Development (34%), Motivation (32%), and Process (24%), which is identical to the individual results of Tool 2. The individual matrix of Tool 1 differs, as it does not include Motivation and Process, but additionally highlights Concepts (55%) and Externalization (55%).

+ Internal Context

Tool 1: This connection refers to the importance of customizing the solutions to be implemented according to the needs of the partner and his internal context in a close cooperation. As Sun pointed out, the provider is not an expert regarding the precise circumstances of the client and his processes. Realizing this over time, led to the decision to develop the spin-off version to be able to better incorporate this aspect, Mars and Earth explained. Venus also considered it crucial to develop achievable objectives for each client individually to promote long-term success, as well as considerable experience with management systems.

Tool 2: Uranus confirmed the positive influence of integrating the objectives into existing management systems, a certain level of management experience, but also financial opportunities. He particularly pointed out the importance of a general commitment of the client's management to sustainability initiatives.

+ Knowledge Development

Tool 1: As noted beforehand, the provider learned through a growing number of projects that clients expected more concrete solutions to implement, which led to an increased focus on the consulting activity. This was illustrated through a stronger emphasis on discussing potential solutions on the one hand, and the development of a spin-off of Tool 1 that should be able to provide concrete solutions for each project on the other hand.

Tool 2: Saturn mentioned that with smaller organizations, it was important to be able to test the recommendations on a smaller scale. Neptune also observed that though not all recommendations might be implemented, the increased awareness that developed through the assessment changed the perspective and thus had a general positive influence on the sustainability performance.

+ Motivation

Tool 1: The main aspect that was addressed in the relation between Motivation and Internal Context was that the final result often led to confusion and disappointment, thus negatively affecting the implementation phase, as expressed by Mercury and Sun. Venus stressed an overall high level of motivation regarding the implementation of the standard.

Tool 2: A different factor that was addressed by Neptune: the previously mentioned awareness that sustainability issues exist within the company, and as Uranus also observed, once the organization could be convinced of this fact, the motivation to implement strongly improves.

+ Process

Tool 1: Sections that covered this relation referred to the way disruptions during the process influenced the implementation. Especially delays, which occurred several times, particularly during the data acquisition phase, or in Mercury's case due to problems with the external review, decreased the motivation. So far the sustainability assessment process had not been repeated by any client, no matter the result of the implementation, although Earth found that some clients included certain aspects in their own systems.

Tool 2: A similar experience regarding an assessment repetition was made by Saturn and Jupiter, even though some plans existed to do a second round in some projects, however, those were dropped. Uranus remembered that in the very few cases where it had been done, it was to optimize consultancy services. Saturn furthermore noted that the general process of Tool 2 had a positive influence on the implementation, as long as a certain degree of flexibility was provided.

+ Concepts

Tool 1: One case where Concepts and Implementation overlapped was Sun's project. He assumed that one of the reasons why the customer struggled to get to the stage of implementation, was that the assessment did not provide any support concerning their existing certifications, as certifications have a different concept than performance assessment. Mercury experienced confusion on the client side how to deal with the results, too, and also related it to a lack of knowledge and awareness, also on the supplier side, regarding the concepts: *What is the aim? Why do we do that? We say it's an improvement tool, where often it's more a communication tool.*

+ Externalization

Tool 1: As was already described previously, to get to the point of implementation, solutions need to be developed in close cooperation, as the external provider lacks the client's expertise concerning their context and processes. Earth observed that in several cases the client was not prepared for that, leading to problems in the implementation phase.

4.2.3 Concepts

The average overlap quota being 25%, three themes scored higher: Internal Context (35%), Knowledge Development (26%), and Motivation (44%), which covers exactly with the results of Tool 1. Whereas Tool 2 does not contain Knowledge Development, it does highlight Content (23%). The connections to Internal Context and Motivation will both be addressed more in detail in their respective context.

+ Knowledge Development

Tool 1: The relation between Concepts and Knowledge Development dealt with aspects of sustainability assessment and how the tool concretized sustainability. For instance Earth pointed out that there was still a lack of experience and understanding of sustainability regarding the time-horizon for them as provider as well as their clients, since customers always wanted to see quick results. On the other hand, the tool can improve the comprehension of the spatial impact of sustainability, e.g. Sun described that in his case, the client really appreciated to see that the responsibility was distributed between the different actors of the value chain. Mars and Earth furthermore described how gathering experience led to the decision to develop a spin-off version of Tool 1 in order to concentrate more on the consulting aspect and provide more concrete solutions.

Tool 2: Uranus also noted that the inability or unwillingness of clients to comprehend sustainability as a long-term essential business investment was a major reason to prevent efficient knowledge development, too often clients did not provide adequate resources, e.g. came unprepared to training sessions.

+ Content

Tool 2: Similar to the previous point, Uranus observed that although people were looking for a sustainability assessment tool, they are mainly interested in the environmental dimension and resource-efficiency, putting e.g. the long-term effect of social evils aside. Neptune found that the concept of Tool 2 adequately served his firm's need to capture a sustainability baseline of regions, however, he still saw potential for future development, in particular in terms of a more flexible content, to better suit to the context of the region and the production system.

4.2.4 Content

Noteworthy relations larger than the average (18%) exist to Knowledge Development (30%), Motivation (23%), as well as Process (34%). Regarding variances in the individual tables, Tool 1 highlights Externalization (26%) instead of Motivation. The relation to Process will be addressed later on.

+ Knowledge Development

Tool 1: The relation between Content and Knowledge Development deals on the one hand with the results of the assessment that informed clients about the origin of impacts, as Sun observed. Mars also found that the final overview supports the comprehension of concrete sustainability. On the other hand, a lot of knowledge could also be transferred and developed during the discussions about the scope and thus the content of the projects, as Mars explained. Furthermore, Earth named examples, where clients tried to adapt their own systems to at least part of the content of the assessment for further monitoring.

Tool 2: Similarly to Tool 1, Neptune also reported that the results could be used by the assessed body to continuously learn about and improve his processes in terms of sustainability performance. Saturn and Jupiter agreed that much knowledge was already transferred during the data acquisition interviews, stimulating new trains of thoughts. Regarding the knowledge to create an adequate content, in case of Tool 2 it developed from practical experience gathered over the years. A systematic evolvement approach was adopted only later on, before releasing the second version.

+ Motivation

Tool 1: In order to generate motivation, the content of the tool has to be suitable to the client's needs, for instance in Mercury's case, the product scope was determined to be maize in accordance with previous studies of the customer. Another example was given by Earth, where an external value chain actor demanded the client to prove that he produces in a way that optimizes biodiversity. Earth added that unexpected results could generate more motivation, if it is in terms of the impact-result of the content, as long as the partner was well aware of the content itself.

Tool 2: The client-side of Tool 2 reported similar opinions regarding the suitability of the tool to their needs and how it motivated them. Both Jupiter and Neptune pointed out that an important aspect for them was a certain degree of adaptation of the content according to their wishes, simply to have a tool that was or could be adapted depending on the context they operated in.

+ Externalization

Tool 1: Here it became apparent that the Externalization has consequences on the Content. First of all, Mercury pointed out that Tool 1 was very wide, which could be a benefit, because it could be applied in many contexts. At the same time, she and Sun added that it was also a problem, as only the clients were experts in their respective fields and processes. As Earth clarified, the content was generally only adapted to the production system, but to best understand the system, the supplier required the support of the client to adapt the content accordingly. His acknowledgment that only few actors could determine whether Tool 1 and its content, capturing all three dimensions of sustainability, was suitable to their needs, further complicated the optimal adaptation of the content.

4.2.5 Externalization

Outstanding overlaps with Externalization (average = 21%) are Internal Context (34%), Knowledge Development (29%), Motivation (29%), and Process (32%). Whereas the relation to Motivation is not highlighted in Tool 1, Tool 2 does not point out Knowledge Development and Process. Only the relation between Externalization and Motivation is examined here, the others will be analyzed later on.

+ Motivation

Tool 1: This overlap addresses the question why organizations chose the tool of an external provider to conduct the sustainability performance assessment. The main motivation was the expertise delivered by an entity that specialized on that topic. What furthermore played an important role was the reputation of the supplier brand and increasingly also the reputation of Tool 1 itself.

Tool 2: Using a tool by an external provider was also mainly motivated by the expertise and knowledge that came along with it. Jupiter and Neptune both agreed that they lacked in particularly the agricultural expertise to develop such a tool on their own, as their organizations were no producers. Whereas most clients had heard of the tool before establishing contact, only few systematically analyzed which tool would best fit to their needs.

4.2.6 Internal Context

Within the Internal Context (average = 24%), the connections to Concepts (25%), Externalization (33%), Knowledge Development (28%), and Motivation (52%)

stick out. In Tool 1, Process (26%) is highlighted as well, whereas in Tool 2 Concepts and Knowledge development did not appear more often than the average. The relation to Knowledge Development and Motivation will be addressed at a later point.

+ Externalization

Tool 1: In the situations where Externalization and Internal Context overlapped, it was either regarding the internal context of the supplying or the demanding organization. The first case came up when Mercury criticized that the tool as an external method was somehow lacking a clear strategy to pursue beyond reputational benefits. The client's internal context could also affect the externalization: the assessment required trust, as sensitive data were gathered. Earth illustrated that therefore people representing sustainability within the customer's ranks often positively influenced the awareness for the upcoming process.

Tool 2: Jupiter and Uranus reported that the supplying side had a strong interest to give access to preferably many organizations, while ensuring the quality of the tool. Therefore, and because of their limited resources, the licensing solution with exclusive training rights was introduced.

+ Concepts

Tool 1: This connection is about how existing organizational structures influence the applicability of the concepts built into the tool, e.g. Sun expressed that in his project the client was already active in terms of sustainability, *[…] they already have some certifications and they are in different forums to discuss sustainability [...]*, and were then looking for a quantitative measurement of their sustainability performance. Yet, they struggled to find sufficient resources for the data acquisition. As noted, Mars pointed out that if the client had an established conflicting sustainability philosophy, chances for a fruitful cooperation were limited.

Tool 2: As for Tool 1, the main issue described by Uranus was the earlier noted lack of adequate resources that were not provided by clients, hoping that sustainability could simply be added on top of the everyday business, instead of deeply rooting it within the organizational core business.

+ Process

Tool 1: Regarding the connection between internal context and process, two aspects stuck out. First of all, a lack of proper organization on the client-side further complicated the data acquisition and thus delayed the whole process. Second, the larger the commitment from the customer's top-management, the better were the chances for a successful project, and as Earth furthermore added, one or two people could drive a whole organization and thus positively influence the process.

4.2.7 Knowledge Development

In the combined overview with an average overlap of 25%, these themes appeared more often: Externalization (28%), Internal Context (28%), Motivation (28%), and Process (46%). Motivation is not highlighted in Tool 1's table. Externalization is not colored in the overview of Tool 2, but in addition Implementation (25%, examined previously) and Content (25%, examined previously) are. The connection to Process is going to be analyzed later.

+ Internal Context

Tool 1: Although it was previously noted that existing sustainability competencies inside an organization could be helpful, it also already prevented or at least limited the learning experience in some cases. Mars pointed out that it could lead to issues when the internal perception of sustainability of the client was not compatible with the philosophy of the external provider. A more pragmatic view on sustainability positively impacted on some client's better comprehension of their systems. Furthermore, it showed that even though high-end clients sought more complex solutions as they mistrusted too easy systems, they were equally helpless dealing with the results, thus teaching Earth and the team of Tool 1 to further stress the implementation aspect.

Tool 2: Comparable to previous relations it was pointed out that a commitment to and – like in Tool 1 – a similar philosophy regarding sustainability was elementary. Furthermore, Saturn mentioned that many different levels of knowledge existed, which needed to be taken into consideration and reacted upon when starting and conducting the sustainability assessment.

+ Externalization

Tool 1: Externalization promoted knowledge development in case of the client as well as the provider. Sun pointed out that in his project, the mutual learning aspect was stressed, and *[…] the idea was really to have a partner to understand and also to judge the results.* Mars mentioned that although much of the information itself was often not new to the clients, seeing how the indicators and impacts are connected in a concrete context gave them a whole new perspective and also regularly resulted in discussions that were informative for the tool provider as well. Still the externalization could also lead to difficulties, e.g. in Mercury's project, when different languages created a barrier. The afore mentioned knowledge gap between provider and customer about the characteristics and abilities of the tool was hindering in some cases, too, when it was not clear what could be expected from the tool.

Tool 2: The findings were the same as for Tool 1, both sides could benefit and learn from the external position of the other one through discussions and sharing the expertise in different contexts.

+ Motivation

Tool 1: Particularly clients that showed more interest in continuous improvement of their sustainability performance were motivated to expand their knowledge and get a more concrete understanding of their processes and operations.

Tool 2: A major aspect that was addressed for Tool 2, was its educational character, which was especially highlighted by the client-side, e.g. Neptune reported that one of their main objectives was to train not only their suppliers but also people from within the own organization to get a better and holistic understanding of and awareness about sustainability in the context of agriculture. Mercury also told of training sessions they held with the tool, which had the purpose of educating not only farmers, but also actors from the public sector about sustainability in agriculture.

4.2.8 Motivation

The themes that have an above-average overlap (24%) with Motivation are Concepts (27%) and Internal Context (44%). In the overview of Tool 1, External Context (46%, examined previously) and Externalization (36%, examined previously) stick out as well. Concepts is not part of Tool 2's significant overlaps, but Implementation (21%, examined previously), Knowledge Development (21%, examined previously), and Process (23%) are. The interaction with Process will be addressed later on.

+ Internal Context

Tool 1: The motivation to strategically improve the sustainability performance was mentioned several times, more precisely the wish of several clients to develop a better understanding of sustainability in their context and an according analysis of their processes. Especially highly developed larger actors perceived the opportunity to become more efficient and save resources, which motivated them to seek external advice.

Tool 2: Although the strategic idea of process improvement was mentioned by Uranus, too, the motivation to utilize it developed more out of the desire to establish baseline studies for certain regions, as Jupiter and Neptune pointed out. These results would then be used to optimize regional strategies. Additionally, capacity building within the own organization was brought up by Neptune as an important reason to use Tool 2.

+ Concepts

Tool 1: Sections in which those two themes overlapped dealt with the aspect whether organizations actively sought for certain concepts within the tool. Several cases across all interviews of Tool 1 were given in which the significance of communicating about the sustainability assessment and its results were

highlighted. There was furthermore a strong interest in the power to quantitatively capture the sustainability performance across the value chain, which was connected to clients' strategic objectives, e.g. by Sun. Earth saw the customer's ability to differ between CSR activities and sustainability performance measurement as provided by the tool as a prerequisite for a successful cooperation. He and Mars clarified that preparing and educating clients about the concepts and philosophy of the tool in the initial phase was essential for a more successful reception in the final presentation phase.

Tool 2: Uranus observed that the majority of firms contacting his organization for a sustainability assessment still were attracted due to reputational reasons. Yet, he assumes that those organizations have the biggest chances for a long-term success and continuous improvement that strategically integrate sustainability into their core business for strategic purposes.

4.2.9 Process

When looking at Process, three themes are above the average (17%): Content (20%), Externalization (23%), Knowledge Development (34%), and Motivation (18%). Whereas Tool 2 highlights Implementation (18%, examined previously) instead of Externalization, Motivation is not marked as above average in the table of Tool 1.

+ Content

Tool 1: During several interviews it became clear that the content was flexible to a certain degree and adapting it to the client's needs is a process that mainly takes place during the scoping-discussions. Both, Mars and Earth, stressed the importance of this phase to clearly establish the objective of the client before actually starting to plan the project. Furthermore, during and towards the end of the project, feedback from the client, which generally concerned regional and production system-specific aspects e.g. in Sun's case, was implemented and considered for future projects.

Tool 2: For Tool 2, clients observed that the content covered the most essential aspects and offered some flexibility during the data acquisition, which depended on the skills of the interviewer. Jupiter and Neptune reported that the content had been adapted in newer versions according to their feedback, yet, they were sometimes missing a stronger adaptation to regional contexts and production systems. Uranus revealed that they intended to offer such options in the future.

+ Knowledge Development

Tool 1: Regarding this relation, it was mostly about the learning process that took place throughout and after the projects. This was true for the client-side, where a lot of capacity and understanding was already built in the initial phase and during the process, and not just through the results. But it also applied to the

supplier organization, and not just concerning assessment related 'hard' facts, such as extending the database or being more efficient in the critical review, but even more so for 'soft' aspects concerning the dealings with clients. Earth for instance reported that a strong shift of focus occurred from discussing the results to finding ways to concretize the recommendations. Mars, too, expressed his intentions to concentrate more on discussing the optimal project scope.

Tool 2: As previously mentioned, a lot of knowledge and awareness on the client side developed through training sessions and the conducting of the interviews for the data acquisition. Similar to Tool 1, also the team of Tool 2 could gather a lot of experience through cooperation with and feedback from partners and implement it into improved versions of their tool.

+ Externalization

Tool 1: When Externalization appeared in the Process-context, it mainly concerned the communication between the actors, which had main contacts on either side, e.g. in Sun's case him on the supplier- and two people on the client-side, one from Human Resources and one from Operations. The farms were included in the initial phase, but not in the final phase, a point that was criticized by the client. The externalization and consequently the dependence on each other created issues for the process, as information had to be transferred between the actors, leading to delays, e.g. in the data acquisition phase. In Sun's project it was also brought up that the communication could have been a bit more regular.

Tool 2: The process was sometimes complicated through the externalization, when the provider struggled to coordinate his limited resources to the needs of the client, e.g. for training sessions of new projects, as pointed out by Jupiter.

+ Motivation

Tool 1: This relation dealt with past experiences made by Earth and Mercury that a delay in the project schedule negatively impacts on the client's initial motivation. E.g. in Mercury's project, the disturbance in the project progress meant that the client could not integrate the results into the sustainability report, which was expressed to be one of the client's main objectives.

Tool 2: The interview sections revealed that except for two or three projects, the sustainability process had not been repeated after the first run. Although there had sometimes been intentions to do so, and Saturn agreed that it might be interesting to properly assess the success of the developed strategies, Neptune pointed out that he declined a repetition, because he feared it would move the assessment too close to a compliance check, negatively impacting the trust required to run an efficient measurement in the first place.

The previous parts provided an overview of the primary findings. It could be demonstrated how the themes occurred during the interviews and which role they play within sustainability assessment. It was also illustrated, how the themes are connected among each other in the practical context. Besides

complementing the first and second research objective, getting a clearer understanding of the process of an assessment and the way actors are affected, it supports the subsequent development of a model that aims to facilitate an answer to the research question, how external sustainability assessments might achieve a more significant strategic impact.

5 DISCUSSION

In Chapter 2, an extensive literate review was conducted that covered performance measurement and management as well as sustainability, in particular its effect on performance and its adequate assessment. The review identified several overlaps across the topics which resulted in a group of nine themes: *External Context, Internal Context, Content, Knowledge Development, Motivation, Implementation, Process, Concepts,* and *Externalization*. In doing so, the first research objective, developing a clear understanding of external sustainability assessment and its essential elements, could largely be met. However, the literature lacked material that covered the context of this research, the food value chain. Therefore, primary research was conducted, interviewing experts from the field, to test the validity of the themes in this context. The findings from these interviews as presented in the first part of Chapter 4 emphasized the significance of the themes. They furthermore provided an insight into the role of the involved actors and how they can be affected, thus also fulfilling the second research objective to a large extent. Beyond these results, a more detailed analysis demonstrated a strong interconnectivity between the themes, which was illustrated in the second part of Chapter 4, *Motivation, Knowledge Development,* and *Internal Context* standing out in particular. This could further clarify the understanding of the process and the way actors are affected. However, it also showed that sustainability assessments are a complex topic due to the involved elements and the interconnectivity among them possibly creating conflicts. Additionally, previous models were only partially useful in depicting the process, a consequence of the *Externalization.* In order to provide further clarification and facilitate a better response to the research question, how external sustainability assessments might achieve a more significant strategic impact, a model illustrating the overall process and the most essential impacts along the way is proposed and discussed subsequently. Afterwards, potential conflicts are highlighted.

5.1 Model Proposition

The themes identified during the literature review all played a role in sustainability assessment. Building on the theory and to a large part on the practical expertise of the interview partners, a model illustrating the relations within the sustainability performance assessment process could be developed, as seen in FIGURE 5. It should be noted that the model is simplified and focuses on the critical aspects of the process, instead of providing a detailed overview that includes every single step.

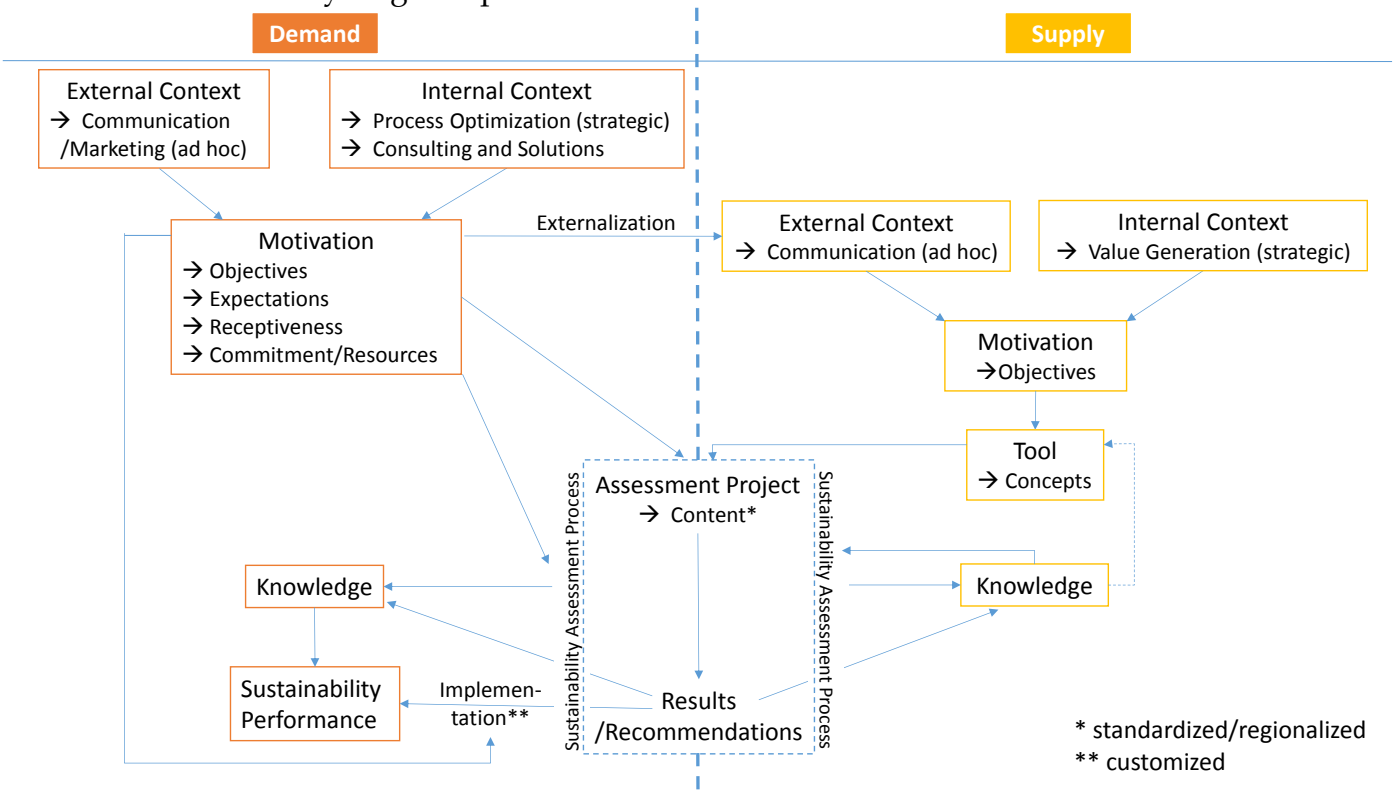


FIGURE 5 External Sustainability Assessment Model

The general starting point is the client's motivation to conduct a sustainability assessment. As the topic is rather new and considered as complex, the actual assessment task is externalized. The reasons that led to the demand of an assessment could be found in the external and/or internal context and created specific objectives and expectations regarding the assessment. From outside the company boarders, clients perceived pressure from the value chain and required a tool for communication and marketing purposes, which Schrettle et al. (2014) categorized as ad hoc measures. Coming from within the organization, there was usually the wish to optimize the organization's processes and to receive consulting and solutions from experts, a more strategic approach. It should be noted that objectives derived from both contexts are not mutually exclusive; in those cases, where the tools were used in a more strategic fashion, the efforts were also communicated to external stakeholders. The interviews indicated the external context to be the driving force, while assuming the internal context to be

more efficient in terms of performance improvements which matches to the findings of Llopis & Tarí (2003).

A characteristic that previous models lacked, was the externalization of the assessment, which means that another actor was involved, thus creating the demand and supply situation. As the empirical evidence showed, external and internal context were relevant for the tool suppliers as well. First of all, the external demand was a motivation to develop and provide an assessment tool. Second, due to their background, particularly the provider of Tool 1 also had an interest in communicating their sustainability efforts to external stakeholders. At the same time, there was also an internal entrepreneurial interest involved. Both contexts led to the motivation to provide an assessment tool, with certain objectives from the suppliers, e.g. value generation.

Within the tools, several concepts are integrated, which were determined by the providers. This step involved a stakeholder survey, but is not individually adapted to each client. Both tools chose a holistic, quantitative approach to sustainability considering the economic, environmental, and social dimension. Both suppliers had the objective to provide a tool that enables the measurement and continuous improvement of sustainability performance in agriculture and thus a strategy- rather than a marketing-support.

As there is demand (Motivation) as well as supply for a sustainability assessment tool, both sides agree on a mutual assessment project, which at the end leads to results and recommendations for the client, together forming the starting- and end-point of the sustainability assessment process. The Content theme is embedded within the assessment project and would require a strong adaptation to the client-organization, according to Bourne et al. (2005) and Schrettle et al. (2014). However, it is mostly standardized and usually only adapted to the regional context and/or the production system. After initial meetings that generally were used to familiarize the client with the tool and discuss the goal of the assessment, the required data are gathered, and the final results calculated and presented. One of the main objectives named by clients to conduct such an assessment was knowledge development, which seemed to be generally achieved. Interestingly, new knowledge did not only emerge from the results, but could be developed throughout the process and not only on the demand- but also on the supply-side. The provider then applies the new experiences to later projects – mostly considering 'soft' issues, such as expectation management, which accords with Hoffman & Henn (2008), who found that barriers to change often emerge from social and psychological circumstances that were not adequately considered – or to updating the tool, e.g. more flexibility and precision according to regions and production systems through newly acquired data. This ongoing development of the tool is important, as Biticti et al. (2000) point out that whereas performance measurement can initiate change, changes in the internal and external environment need to be considered and require according adaptations. Insufficiently allocated resources on the client side generally were the most challenging factor during the assessment process.

The recommendations that are given to the client in order to improve their sustainability performance were often considered as too general, which seemed to prevent their implementation. The research's evidence matches with the findings of de Waal & Counet (2009) that one of the main reasons for implementation failures was still a lack of management commitment that does not sufficiently prioritize the implementation. So it might be argued, which side bears more responsibility. The providers seem to have drawn the right conclusions as both announced plans to focus more on the consulting aspect to provide more customized recommendations and solutions. This is in agreement with the results of Amarantunga & Baldry (2002) who consider the ability to actually utilize the delivered results as crucial to drive change and thus move from measurement to performance management. Although different levels of receptiveness for change after the recommendations were brought up by the participants, none of them indicated that they attempted an active identification and consideration of barriers, as recommended by Lozano (2013), which can be found in the internal (managerial, historical, organizational, and supportive) and external context (stakeholder behavior).

Whether or not the client's sustainability performance ultimately improves is difficult to answer as it depends on how sustainability performance is defined. The indirect improvement of the sustainability performance through a better understanding of sustainability in the client's own context matches with the proposed framework of Schrettle et al. (2014). Regarding the direct impact on the assessment, one needs to question how efficient the tools are at assessing the sustainability performance. Considering that the collected experience is also applied by the provider side to the development of the tools, it can be assumed that they increasingly succeed at this task. Should this assumption be accepted, it can then be said that implementing the recommendations based on the tool's results would directly improve the client's sustainability performance. However, the only empirical evidence found for this reasoning is based on a few expert opinions. A more reliable statement will only be possible, once several projects have been re-assessed. However, it can be concluded that ultimately, the content of the performance assessment is not the major barrier as is illustrated by the fact that the results are generally well received, even though organizational alignment is named as one of the crucial elements in the literature. In this scenario however, the customization seems to be sufficient and more decisive at the stage of the solution development, as has been recognized by the provider side. The main challenge rather seems to lie in different conceptions of sustainability and the consequent misunderstanding regarding the tool's purpose, which Otley (2006) considered as one of the main reasons for failure. Using the categorization of Franco-Santos et al. (2007), the interviews showed that the provider side saw the main role of the tools to be Strategy management, whereas the majority of their clients was still focused on Communication. This also matches with Lozano's (2009) findings that organizations often lack a long-term strategy when they engage in sustainability. Several other critical points could be identified, which are going to be highlighted in the next section.

5.2 Conflicts

The presentation and analysis showed that the identified themes were strongly interwoven and though it allowed the author to develop the above model, it also brought up conflicts that the model consequently inherited. Furthermore, some of the themes also contain conflicts on their own. In order to increase the awareness about those problems, the most striking will be addressed subsequently.

Motivation had the largest average of theme overlaps and thus also had the most connections within the model. It makes sense, as it was the respective motivation that led to the externalization as well as the development of assessment tools. It furthermore initiates the beginning, influences the course of the mutual project, and ultimately it is elementary to the success of the assessment. Both, primary and secondary findings agree on its importance and considered a lack of commitment as one of the major reasons for failure (e.g. de Waal & Counet 2009). The primary research found that the involvement of two actors meant two motivations, which could lead to different objectives with the tool, as suggested by Aramyan et al. (2007). Particularly the background and the business model of the provider of Tool 1 could thus easily result in differing objectives, as illustrated by a cancelled project for a pilot study. It also seems that motivation might strongly be affected by the internal context in terms of the existing culture: practitioners and literature (de Waal & Counet 2009) pointed out that if an organization lacked a general management culture, the implementation was rather likely to fail.

Knowledge Development was named as one of the key motivations to conduct a sustainability assessment and also has the second highest average of theme overlaps. As could be seen, both sides benefit from it. However, as the assessment is externalized, so are many benefits of a measurement system, such as the ability to learn and the creation of a learning culture (Amarantunga & Baldry 2002). As the interviews showed, knowledge was transferred throughout the process, yet it could be questioned, whether it achieved the same results as an internal assessment system, especially considering differing motivations of the actors. This is supported by Aramyan et al. (2007), who found that such conditions might hinder the sharing of information between the actors. On the other hand, it was also reported that in some cases, the mutual cooperation could also reinforce the learning effect. In any case, the initial knowledge gap between the actors, concerning sustainability assessment as well as the client's organization, created issues, also in terms of alignment as demanded by the literature.

One of those issues that was mentioned earlier as well, are different conceptions regarding the role of the tool, which Otley (2006) considers a major source of failure. Whereas the provider side saw it as a tool for strategy management and decision support, the main intention on the client side was still largely to utilize it for communication. It also became clear that differing conceptions of sustainability could create issues, either in terms of

underestimating the necessary means for a long-term effort, or in mismatched definitions, which meant that there was no common ground for discussions and thus mutual learning. This is supported by Searcy (2009b), who warns of the complexity of sustainability and the struggle of organizations to find a common understanding of it.

Finally, being a sustainability assessment tool, it also contains the inherent system-conflicts and trade-offs between the three dimensions, as pointed out by Dyllick & Hockerts (2002). Even though such an approach solves the ongoing problem of traditional performance measurement, the neglecting of non-financial dimensions (Drucker 1954; Johnson & Kaplan 1987; Fitzgerald 1988; Goold & Quinn 1990; Ghalayini & Noble 1996; Choong 2013), the interest of clients in the performance measurement capabilities seemed to be limited. It might increase, if the trade-offs could be expressed in financial units, however this would lead to optimizing the economic dimension once more, instead of the overall sustainability performance and adopt the criticism against performance measurement.

This part introduced a model based on the previous findings and highlighted the conflicts that can arise when conducting such an assessment. Answering the research question, how a more significant strategic impact could be achieved, it can be summarized that an extensive and efficient cooperation between the supply and demand side is necessary. The model can support such a collaboration in so far, as it illustrates the process and can be used to identify and solve conflicts at an early stage.

6 CONCLUSION

The corporate sector increasingly seeks to implement sustainability into their strategic management, yet it struggles to implement solutions beyond marketing efforts. This research illustrated that at least partial answers to this dilemma exist, such as the tools of the case studies. Many organizations fail, though, to realize the opportunities provided in the assessment tools on a strategic level, as they do not succeed to create the link to performance measurement and management, a connection that has also mostly been ignored by previous research. However, this research demonstrated that sustainability performance assessment and performance measurement and management have a lot of common ground that can help to clarify the strategic role of external sustainability performance assessments.

Nevertheless, it also became obvious that combining sustainability and performance measurement and management creates new challenges, particularly in the examined cases, where the sustainability assessment was provided by an external actor, creating a demand and supply situation. As agriculture and sustainability are both extremely complex and multifactorial, such a setup can be mutually beneficial and informative when both sides manage to contribute their expertise along the process, but it might also happen that the cooperation suffers from the knowledge gaps between the actors.

This study furthermore proposed a model that depicts not only the involved themes from previous works, but also how these are connected among each other. Besides getting a better understanding of the overall process, the model allows to highlight potential conflict areas. In doing so, it facilitates an answer to the research question, how such assessments might achieve a more significant impact: both sides need to be aware of the process and its most essential elements (*External Context, Internal Context, Content, Knowledge Development, Motivation, Implementation, Process, Concepts,* and *Externalization*), as well as the interconnectivity among them; understanding the process and which conflicts may arise can enable a more efficient collaboration with shared expertise, thus ultimately increasing the chances of a stronger impact.

6.1 Implications

It needs to be highlighted once more that sustainability in the corporate and particularly the food context is essentially a complex problem and that it will not be possible to find an approach that globally applies (Searcy 2012). Still this research has practical implications.

First of all, it demonstrated the potential of two modern tools to integrate sustainability into the managerial and strategic level. Even though its background can be found in environmental impact assessment, the research showed the strong relation to the field of performance measurement and management. This is an opportunity that client organizations were still often not or only partially aware of.

Second, the proposed model can clarify the process of sustainability assessment and thus increase the awareness of interconnectivities between different aspects. Being more sensible to potential conflicts could support both sides to prepare accordingly and cooperate more smoothly and efficiently. Even though certain aspects are unlikely to change, such as the respective contexts and the deriving objectives, knowing their importance in the process enables both actors to identify barriers early on and develop strategies to overcome them.

Finally, this paper showed that in practice many of the challenges to sustainability assessment no longer come from finding a way to measure sustainability, but from finding the adequate way for each organization and developing strategies to implement it successfully on the long-term. By drawing from the lessons of performance measurement and management, the research could show that providers are on the right track by supplying flexible standardized tools and customized solutions based on the assessment results, interpreted according to the client's context. Yet, this research has its limitations and further research will have to be conducted to provide more concrete answers.

6.2 Limitations and Future Research

This study underlies several limitations that need to be acknowledged. First of all, having selected a case study strategy, this research contains the characteristics that are inherent to the approach, as discussed in the methodology section. In a nutshell, a case study is not suited to provide globally applicable findings, instead it focuses on exploring and understanding unique situations. Therefore, the proposed model cannot claim to be universally valid. However, it can serve as a starting point for future research, to be further tested. Quantitative data could be gathered through questionnaires and the validity of this research's results could be verified against more available tools. It might also be intriguing to apply the model in other contexts and sectors beside the food industry that might find themselves in a comparable initial situation.

Secondly, the collected primary data has a certain regional focus on Europe and America, even though some of the interviewees also gathered experiences in Africa and Asia. Considering the importance of the context, it might be that

cultural differences as well as legislative frameworks play a much larger role than was assumed in this paper. Future research could go into more detail about the findings in a specific region, or experiences with and from actors at specific points of the value chain. Keeping that in mind, it also needs to be emphasized once more that even though the sustainability assessments generally find the farms to be the most important stage to act upon, no farmers participated in this study. The reason for this is simple: too many variables are involved, no two farms are the same, and the available resources for this research did neither allow for an extensive gathering of data, nor did it seem to make sense for the objective of this research, which required broader perspectives due to its preliminary character. Therefore future research could also explore the role and behavior of the farmers.

Last, the timeframe for this research was limited and the examined phenomena rather new. Additionally, sustainability and the food value chain are extremely complex subjects on their own, and even more so when combined. Although one of the tools existed since almost 15 years, the topic of sustainability assessment in agriculture only started to gain momentum within the last years. Therefore the tools themselves, but in particularly the way they are applied still leaves much space in terms of gathering experience and witnessing medium- and long-term impacts of the assessment. Furthermore, the assessment has almost never been repeated, thus this research had to rely on expert opinions regarding the impact of the assessments. Future research could therefore address the question, whether such assessments actually improve the sustainability performance. It would also be interesting to see, if the type of motivation, internal or external, really influences the overall outcome.

Despite the above limitations, this research significantly contributes to the existing literature, as it is the first work that connects sustainability and performance measurement in the context of the food value chain. In doing so, it becomes possible to get a clearer perspective on the process of sustainability assessment, thus allowing a better preparation and more efficient cooperation between the actors. Nevertheless the findings of this study need to be further examined using quantitative as well as qualitative methods so that their validity can be determined.

REFERENCES

Agger, A. 2010. Involving citizens in sustainable development: evidence of new forms of participation in the Danish agenda 21 schemes. Local Environment 15(6), 541-552.

Ahn, H. 2001. Applying the balanced scorecard concept: an experience report. Long Range Planning 34(4), 441-461.

Amaratunga, D. & Baldry, D. 2001. Case study methodology as a means of theory building: performance measurement in facilities management organisations. Work Study 50(3), 95-105.

Amaratunga, D. & Baldry, D. 2002. Moving from performance measurement to performance management. Facilities 20(5/6), 217-223.

Anstine, J. 2000. Consumers' willingness to pay for recycled content in plastic kitchen garbage bags: a hedonic price approach. Applied Economics Letters 7(1), 35–39.

Appelbaum, S.H., Habashy, S., Malo, J.L. & Shafiq, H. 2012. Back to the future: revisiting Kotter's 1996 change model. Journal of Management Development 31(8), 764-782.

Aramyan, L., Oude Lansink, A., van der Vorst, J. & van Kooten, O. 2007. Performance Measurement in agri-food supply chains: a case study. Supply Chain Management: An International Journal 12(4), 304-315.

Arthur, S. & Nazroo, J. 2003. Designing Fieldwork Strategies and Materials. In J. Ritchie & J. Lewis (Eds) Qualitative research practice: A guide for social science students and researchers. London: Sage, 109-137.

Atkinson, A.A. 1998. Strategic performance measurement and incentive compensation. European Management Journal 16(5), 552-561.

Babakri, K.A., Bennett, R.A. & Franchetti, M. 2003. Critical factors for implementing ISO 14001 standard in United States industrial companies. Journal of Cleaner Production 11(7), 749-752.

Baker, W.E. & Sinkula, J.M. 1999. The synergistic effect of market orientation and learning orientation on organizational performance. Journal of the academy of marketing science 27(4), 411-427.

Banks, R.L. & Wheelwright, S.C. 1979. Operations versus strategy – trading tomorrow for today. Harvard Business Review 57(3), 112-20.

Bansal, P. & Roth, K. 2000. Why companies go green: a model of ecological responsiveness. Academy of Management Journal 43(4), 717–736.

Bazeley, Pat 2007. Qualitative Data Analysis with NVivo. London: Sage.

Becker, E., Jahn, T, Stieß, I. & Wehling, P. 1997. Sustainability: A Cross-Disciplinary Concept for Social Transformations. Report on the results of the first phase of the MOST-project: To-wards Sustainable Development Paradigms and Policies. Paris: UNESCO.

Becker, H.S 1986.Writing for social scientists: how to start and finish your thesis. Chicago: University Press of Chicago.

Berg, B. L. 2001. Qualitative research methods for the social sciences. Boston: Pearson.

Bhaskar, R. 1978. A Realist Theory of Science. Hassocks: Harvester Press.

Bititci, U.S., Carrie, A.S. and Mcdevitt, L. 1997. Integrated performance measurement systems: a development guide. International Journal of Operations & Production Management, 17(5), 522-534.

Bititci, U.S., Mendibil, K., Nudurupati, S., Turner, T. & Garengo, P. 2004. The interplay between performance measurement, organizational culture and management styles. Measuring Business Excellence 8(3), 28-41.

Bititci, U.S., Nudurupati, S.S., Turner, T.J. & Creighton, S. 2002. Web enabled performance measurement systems – management implications. International Journal of Operations & Production Management 22(11), 1273-1287.

Bititci, U.S., Turner, U. & Begemann, C. 2000. Dynamics of performance measurement systems. International Journal of Operations & Production Management 20(6), 692-704.

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

Bourne, M., Kennerley, M. & Franco-Santos, M. 2005. Managing through measures: a study of impact on performance. Journal of Manufacturing Technology Management 16(4), 373-395.

Bourne, M., Neely, A.D., Platts, K.W. & Mills, J.F. 2002. The success and failure of performance measurement initiatives: the perceptions of participating managers. International Journal of Operations & Production Management 22(11), 1288-1310.

Bourne, M., Pavlov, A., Franco-Santos, M., Lucianetti, L. & Mura, M. 2013. Generating organisational performance: The contributing effects of performance measurement and human resource management practices. International Journal of Operations & Production Management 33(11/12), 1599-1622.

Bowen, D.E. & Ostroff, C. 2004. Understanding HRM-Firm Performance Linkages: The Role of the "Strength" of the HRM System. Academy of Management Review 29(2), 204–221.

Bowen, H. 1953. Social Responsibilities of the Businessman. New York: Harper & Row.

Bowersox, D.J. & Closs, D.J. 1996. Logistical Management: The Integrated Supply Chain Process. New York, NY: McGraw-Hill.

Brignall, S. (1992). Performance measurement systems as change agents: a case for further research. Warwick Business School Papers No. 72. Warwick: University of Warwick, Warwick Business School Research Bureau.

Brockman, B.K. & Morgan, R.M. 2003. The role of existing knowledge in new product innovativeness and performance. Decision Sciences 34(2), 385-419.

Bromley, D.B. 1986. The Case Study Method in Psychology and Related Disciplines. Chichester: Wiley.

Bryman, A. 2001. Social Research Methods. Oxford: Oxford University Press

Burgess, R.G. 1982. Elements of sampling in field research. In R.G. Burgess (Ed.) Field Research: A Source Book and Field Manual. London: Allen & Unwin.

Busi, M., & Bititci, U. S. 2006. Collaborative performance management: present gaps and future research. International Journal of Productivity and Performance Management 55(1), 7-25.

Camp, R.C. 1989. Benchmarking – the Search for Industry Best Practices that Lead to Superior Performance. Milwaukee, WI: ASQS Quality Press.

Carlowitz, von, H.C, 1713. Sylvicultura oeconomica, oder haußwirthliche Nachricht und Naturmäßige Anweisung zur wilden Baum-Zucht. Leipzig: Braun.

Carroll, A. 1999. Corporate social responsibility. Business and Society 38(3), 268-295.

Catasús, B., Ersson, S., Gröjer, J.-E. & Wallentin F.Y. 2007. What gets measured gets... on indicating, mobilizing and acting. Accounting, Auditing & Accountability Journal 20(4), 505-521.

Choong, K. K. 2013. Are PMS meeting the measurement needs of BPM? A literature review. Business Process Management Journal 19(3), 535-574.

Christopher, M. 2011. Logistics and Supply Chain Management, 4th edition. Harley, UK: Pearson Education.

Clark, B. 2006. Measuring Performance: The marketing perspective. In A. Neely (Ed.) Business performance measurement. Cambridge: Cambridge University Press, 22-40.

Clark, J. M. 1916. The changing basis of economic responsibility. The Journal of Political Economy 24(3), 209-229.

Cocca, P. & Alberti, M. 2010. A framework to assess performance measurement systems in SMEs. International Journal of Productivity and Performance Management 59(2), 186-200.

Cohen, H.B. 1998. The performance paradox. The Academy of Management Executive 12(3), 30-40.

Colsman, B. 2013. Nachhaltigkeitscontrolling. Wiesbaden: Springer.

Cresswell, J. 1998. Qualitative Inquiry and Research Design: Choosing Among Five Traditions. Thousand Oaks: Sage.

Daly, H. E. 1990b. Sustainable development: from concept and theory to operational principles. Population and Development Review 16, 25-43.

Daly, H.E. 1990a. Toward Some Operational Principles of Sustainable Development. Ecological Economics 2(1), 1-6.

Darnall, N., Gallagher, D.R., Rigling, D., Andrews, R.N.L. & Amaral, D. 2000. Environmental management systems: opportunities for improved environmental and business strategy? Environmental Quality Management 9(3), 1-9.

Darroch, J. & McNaughton, R. 2003. Beyond market orientation: knowledge management and the innovativeness of New Zealand firms. European journal of Marketing 37(3/4), 572-593.

Day, G.S. 1994. The capabilities of market-driven organizations. Journal of marketing 58(4), 37-52.

De Toni, A. & Tonchia, S. 2001. Performance measurement systems – models, characteristics and measures. International Journal of Operations & Production Management 21(1/2), 46-70.

De Vries, H.J., Bayramoglu, D.K. & van der Wiele, T. 2012. Business and environmental impact of ISO 14001. International Journal of Quality & Reliability Management 29(4), 425-435.

De Waal, A.A. & Coevert, V. 2007. The effect of performance management on the organizational results of a bank. International Journal of Productivity and Performance Management 56(5/6), 397-416.

De Waal, A.A. & Counet, H. 2009. Lessons learned from performance management systems implementations. International Journal of Productivity and Performance Management 58(4), 367-390.

De Waal, A.A. 2002. The role of behavioural factors in the successful implementation and use of performance management systems. In A. Neely, A. Walters & R. Austin (Eds.) Performance Measurement and Management: Research and Action. Cranfield: Centre for Business Performance, Cranfield School of Management, 157-164.

Delmas, M. 2001. Stakeholders and competitive advantage: the case of ISO 14001. Production and Operations Management 10(3), 343-358.

Demartini, C. 2014. Performance Management Systems - Design, Diagnosis and Use. Berlin Heidelberg: Springer.

Denzin, H. 1984. The Research Act. Englewood Cliffs: Prentice-Hall.

Diener, E. 2006. Guidelines for national indicators of subjective well-being and Ill-being. Applied Research in Quality of Life 1(2), 151-157.

Donaldson, T. & Preston, L.E. 1995. The stakeholder theory of the corporation: concepts, evidence, and implications. Academy of Management Review 20(1), 65–91.

Drucker, P.F. 1954. The Practice of Management. New York, NY: Harper and Row.

Dyllick, T., & Hockerts, K. 2002. Beyond the business case for corporate sustainability. Business strategy and the environment 11(2), 130-141.

Easterby-Smith, M. 1991. Management Research: An Introduction. London: Sage.

Eccles, R.G. & Pyburn, P.J. 1992. Creating a comprehensive system to measure performance. Management Accounting 74(4), 41-44.

Eccles, R.G. 1991. The Performance Measurement Manifesto. Harvard Business Review 69(1), 131-137.

Eisenhardt, K.M. 1989. Building theories from case study research. Academy of Management Review 14(4), 532-555.

Elkington, J. 1994. Towards the suitable corporation: win-win-win business strategies for sustainable development. California management review 36(2), 90-100.

Etzion, D. 2007. Research on organizations and the natural environment, 1992 – present: a review. Journal of Management 33(4), 637–664.

Evans, J.R. 2001. An exploratory study of performance measurement systems and relationship with performance results. Paper presented at the Decisions Science Institute 32nd Annual Conference Digitizing Decisions & Markets, San Francisco, CA, 1-27.

FAO 2013. Statistical Yearbook 2013. World Food and Agriculture. Rome: Food and Agricultural Organization of the United Nations.

Ferreira, A. & Otley, D. 2009. The design and use of performance management systems: An extended framework for analysis. Management Accounting Research 20(4), 263-282.

Figge, F., Hahn, T., Schaltegger, S. & Wagner, M. 2002. The sustainability balanced scorecard–linking sustainability management to business strategy. Business strategy and the Environment, 11(5), 269-284.

Fiol, C.M. & Lyles, M.A. 1985. Organizational learning. Academy of management review 10(4), 803-813.

Fitzgerald, L. 1988. Management performance measurement in service industries. International Journal of Operations & Production Management 8(3), 109-116.

Flamholtz, E.G. 1983. Accounting, budgeting and control-systems in their organizational context – theoretical and empirical-perspectives. Accounting, Organisations and Society 8(2), 153-169.

Flamholtz, E.G., Das, T.K. & Tsui, A.S. 1985. Towards an integrative framework of organisational control. Accounting, Organisations and Society 10(1), 35-50.

Flynn, B.B., Huo, B. & Zhao, X. 2010. The impact of supply chain integration on performance: A contingency and configuration approach. Journal of Operations Management 28(1), 58-71.

Folan, P. & Browne, J. 2005. A review of performance measurement: Towards performance management. Computers in Industry 56 (7), 663-680.

Forza, C. & Salvador, F. 2000. Assessing some distinctive dimensions of performance feedback information in high-performing plants. International Journal of Operations & Production Management 20(3), 359-385.

Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D. & Neely, A. 2007. Towards a definition of a business performance measurement system. International Journal of Operations & Production Management 27(8), 784 – 801.

Frank, M., Fischer, K., & Voeste, D. 2014. BASF: Messbarkeit gemeinsamer Wertschöpfung in der Landwirtschaft. In M D'heur (Ed.) CSR und Value Chain Management. Heidelberg: Springer.

Fraser, E. D. G., Dougill, A. J., Mabee, W. E., Reed, M. & McAlpine, P. 2006. Bottom up and top down: analysis of participatory processes for sustainability indicator identification as a pathway to community empowerment and sustainable environmental management. Journal of Environmental Management 78 (2), 114-127.

Freeman, R.E. 1984. Strategic Management: A Stakeholder Approach. Boston: Pitman.

Gates, S. 1999. Aligning Strategic Performance Measures and Results. New York, NY: The Conference Board.

Ghalayini, A.M. & Noble, J.S. 1996. The changing basis of performance measurement. International Journal of Operations & Production Management 16(8), 63-80.
Goldratt, E.M. & Cox, J. 1986. The Goal: Beating the Competition. Hounslow: Creative Output Books.
Goold, M. & Quinn, J.J. 1990. The paradox of strategic controls. Strategic Management Journal 11(1), 43-57.
Goyder, G. 1961.The Responsible Company. Oxford: Blackwell.
Guba, E, G. 1990. The Alternative Paradigm Dialog. In E. G. Guba (Ed.) The Paradigm Dialog. London: Sage.
Guba, E. & Lincoln, Y. 1994. Competing paradigms in qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds) Handbook of Qualitative Research. London: Sage, 105–117.
Hacking, T. & Guthrie, P. 2008. A framework for clarifying the meaning of Triple Bottom-Line, Integrated, and Sustainability Assessment. Environmental Impact Assessment Review 28(2), 73-89.
Hahn, T., Figge, F., Pinkse, J. & Preuss, L. 2010. Trade-offs in corporate sustainability: you can't have your cake and eat it. Business Strategy and the Environment 19(4), 217-229.
Hakim, C. 2000. Research Design: Successful Research Designs For Social And Economic Research, 2nd edition. London: Routledge.
Hall, R.W. 1983. Zero Inventories. Homewood, IL: Dow-Jones Irwin.
Hammersley, M. & Atkinson, P. 1995. Ethnography: Principles in Practice, 2nd edition. London: Routledge.
Hart S. 1995. A natural-resource-based view of the firm. Academy of Management Review 20(4), 986-1014.
Hart S. 1997. Beyond greening: strategies for a sustainable world. Harvard Business Review 75(1), 66-76.
Hart, S. & Ahuja, G. 1996. Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance. Business Strategy and the Environment 5(1), 30–37.
Hayes, R.H. & Abernathy, W.J. 1980. Managing our way to economic decline. Harvard Business Review 58(4), 67-77.
Hendricks, J., Defreitas, D. & Walker, D. 1996. Changing performance measures at caterpillar Management Accounting 78(6), 18-23.
Henri, J. 2006. Organizational culture and performance measurement systems. Accounting, organizations and society 31(1), 77-103.
Heras, I., Dick, G. & Casadésus, M. 2002. ISO 9000 registration's impact on sales and profitability: a longitudinal analysis of performance before and after accreditation. International Journal of Quality & Reliability Management 19(6), 774-791.
Hoffman, A.J. & Henn R. 2008. Overcoming the social and psychological barriers to green building. Organization & Environment 21(4), 390-419.
Holloway, I. & Wheeler, S. 1996. Qualitative Research for Nurses. Oxford: Blackwell Science.
Hoque, Z. & James, W. 2000. Linking balanced scorecard measures to size and market factors: impact on organizational performance. Journal of Management Accounting Research 12(1), 1-17.
Huang, P.-S. & Shih, L.-H. 2009. Effective environmental management through environmental knowledge management. International Journal of Environmental Science and Technology 6(1), 35–50.
Huber, G.P. 1991. Organizational learning: The contributing processes and the literatures. Organization science 2(1), 88-115.
Hudson, M., Lean, J. & Smart, P.A. 2001a. Improving control through effective performance measurement in SMEs. Production Planning and Control 12(8), 804-13.

Hudson, M., Smart, A. & Bourne, M. 2001b. Theory and practice in SME performance measurement systems. International Journal of Operations & Production Management 21(8), 1096-115.

Hussain, M. & Hoque, Z. 2002. Understanding non-financial performance measurement practices in Japanese banks. Accounting, Auditing & Accountability Journal 15(2), 162-83.

Iraldo, F., Testa, F. & Frey, M. 2009. Is an environmental management system able to influence environmental and competitive performance? The case of the eco-management and audit scheme (EMAS) in the European union. Journal of Cleaner Production 17(16), 1444–1452.

ISO 2014a. ISO Standards. Retrieved May 30, 2014 from: http://www.iso.org/iso/home/standards.htm.

ISO 2014b. Certification. Retrieved May 30, 2014 from: http://www.iso.org/iso/home/standards/certification.htm.

ISO 2014c. Environmental Management. Retrieved June 1, 2014 from: http://www.iso.org/iso/home/standards/management-standards/iso14000.htm.

Ittner, C., Larcker, D. & Randall, T. 2003. Performance implications of strategic performance measurement in financial service firms. Accounting, Organizations and Society 28(7), 715-741.

Jacobs, B.W., Singhal, V.R. & Subramanian, R. 2010. An empirical investigation of environmental performance and the market value of the firm. Journal of Operations Management 28(5), 430–441.

Jiménez-Jiménez, D. & Sanz-Valle, R. 2011. Innovation, organizational learning, and performance. Journal of Business Research 64(4), 408-417.

Johnson, H.T. & Kaplan, R.S. 1987. Relevance Lost: The Rise and Fall of Management Accounting. Boston, MA: Harvard Business School Press.

Johnston, R., Brignall, S. & Fitzgerald, L. 2002. Good enough performance measurement: a trade-off between activity and action. Journal of the Operational Research Society 53(3), 256-262.

Kaplan, R.S. & Norton, D.P. 1992. The balanced scorecard – measures that drive performance. Harvard Business Review 70(1), 71-79.

Kaplan, R.S. & Norton, D.P. 1996. Using the balanced scorecard as a strategic management system. Harvard Business Review 74(1), 75-84.

Kaplan, R.S. & Norton, D.P. 2001. The Strategy Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment. Boston, MA: Harvard Business School Press.

Kaplan, R.S. 1994. Devising a balanced scorecard matched to business strategy. Planning Review 22(5), 15-19.

Kennerley, M. & Bourne, M. 2003. Assessing and maximising the impact of measuring business performance. Proceedings of the POMS/EurOMA Conference, Cernobbio, Lake Como, Italy, 16-18 June, 493-502.

Kennerley, M.P. & Neely, A.D. 2002. A framework of the factors affecting the evolution of performance measurement systems. International Journal of Operations & Production Management 22(11), 1222-1245.

Kerssens-Van Drongelen, I.C. & Fisscher, O.A.M. 2003. Ethical dilemmas in performance measurement. Journal of Business Ethics 45(1), 51-63.

Keskin, H. 2006. Market orientation, learning orientation, and innovation capabilities in SMEs: An extended model. European Journal of Innovation Management 9(4), 396-417.

King, A.A. & Lenox, M.J. 2001. Lean and green? An empirical examination of the relationship between lean production and environmental performance. Production and Operations Management 10(3), 244–256.

King, N. 2004. Using interviews in qualitative research. In C. Cassell & G. Symon (Eds) Essential Guide to Qualitative Methods in Organizational Research. London: Sage, 11-22.

Klassen, R.D. & McLaughlin, C.P. 1996. The impact of environmental management on firm performance. Management Science 42(8), 1199–1214.

Kolk, A. 2003. Trends in sustainability reporting by the Fortune Global 250. Business strategy and the environment 12(5), 279-291.

Kotler, P. 1984. Marketing Management Analysis, Planning and Control. Englewood Cliffs, NJ: Prentice-Hall.

Kotter, J.P. 1995. Leading change: Why transformation efforts fail. Harvard business review 73(2), 59-67.

KPMG/WIMM. 1999. KPMG International Survey of Environmental Reporting 1999. The Hague: KPMG.

Krut, R. & Moretz A. 2000. The state of global environmental reporting: lessons from the Global 100. Corporate Environmental Strategy 7(1), 85–91.

Kueng, P., Meier, A. & Wettstein T 2001. Performance measurement systems must be engineered. Communications of the Association for Information Systems 7(3), 1–27.

Labuschagne, C., Brent, A.C. & Van Erck, R.P.G. 2005. Assessing the sustainability performances of industries 13(4), 373–385.

Lagi, M., Bertrand, K. Z., & Bar-Yam, Y. 2011. The food crises and political instability in North Africa and the Middle East. arXiv preprint arXiv:1108.2455.

Lebas, M.J. & Euske, K. 2002. A conceptual and operational delineation of performance. In A. Neely (Ed.) Business performance measurement: theory and practice. Cambridge: Cambridge University Press, 65–79.

Lebas, M.J. 1995. Performance measurement and performance management. International Journal of Production Economics 41(1-3), 23-35.

Legard, R., Keegan, J. & Ward, J. 2003. In-depth interviews. In J. Ritchie & J. Lewis (Eds) Qualitative research practice: A guide for social science students and researchers. London: Sage, 138-169.

Lewis, J. 2003. Design issues. In J. Ritchie & J. Lewis (Eds) Qualitative research practice: A guide for social science students and researchers. London: Sage, 47-76.

Line, M., Hawley, H. & Krut, R. 2002. The development of global environmental and social reporting. Corporate Environmental Strategy 9(1), 69–78.

Lingle, J.H. & Schiemann, W.A. 1996. From balanced scorecard to strategy gauge: is measurement worth it? Management Review 85(3), 56-62.

Lipe, M.G. & Salterio, S.E. 2000. The balanced scorecard: judgmental effects of common and unique performance measures. Accounting Review 75(3), 283-298.

Lipe, M.G. & Salterio, S.E. 2002. A note on the judgmental effects of the balanced scorecard's information organization. Accounting Organizations and Society 27(6), 531-540.

Llopis, J. & Tarí, J. 2003. The importance of internal aspects in quality improvement. International Journal of Quality & Reliability Management 20(3), 304-324.

Lockamy, A. & Cox, J.F. 1995. An empirical study of division and plant performance measurement systems in selected world-class manufacturing firms – linkages for competitive advantage. International Journal of Production Research 33(1), 221-236.

Lockamy, A. 1998. Quality-focused performance measurement systems: a normative model. International Journal of Operations & Production Management 18(8), 740-766.

Lokman, M. & Clarke, B. 1999. Market competition, management accounting systems and business unit performance. Management Accounting Research 10(2), 137-58.

López-Gamero, M.D., Molina-Azorín, J.F. & Claver-Cortés, E. 2009. Environmental regulation as a potential to change managerial perception, environmental

management and financial performance. 16th International Annual EurOMA Conference, Göteborg, Sweden.

Lozano, R. 2009. Orchestrating Organisational Change for Corporate Sustainability. Strategies to Overcome Resistance to Change and to Faciliate Institutionalization, PhD Doctoral dissertation. Cardiff: University of Cardiff.

Lozano, R. 2013. Are companies planning their organizational change for corporate sustainability? An analysis of three case studies on resistance to change and their strategies to overcome it. Corporate Social Responsibility and Environmental Management 20(2013), 275-295.

Luchs, M.G., Walker Naylor, R., Irwin, J.R. & Raghunathan, R. 2010. The sustainability liability: potential negative effects of ethicality on product preference. Journal of Marketing 74, 18–31.

Lynch, R.L. & Cross, K.F. 1991. Measure Up – The Essential Guide to Measuring Business Performance. London: Mandarin.

Magee, L., Scerri, A. & James, P. 2012. Measuring social sustainability: a community-centred approach. Applied Research in Quality of Life 7(3), 239-261.

Maisel, L.S. 2001. Performance Measurement Practices Survey Results. New York: AICPA.

Malina, M.A. & Selto, F.H. 2001. Communicating and controlling strategy: an empirical study of the effectiveness of the balanced scorecard. Journal of Management Accounting Research 13, 47-90.

Manoochehri, G. 1999. Overcoming obstacles to developing effective performance measures. Work Study 48(6), 223-229.

Marr, B. & Schiuma, G. 2003. Business performance measurement – past, present and future. Management Decision 41(8), 680-687.

Martinez, V., Kennerley, M., Harpley, R., Wakalen, R., Hart, K. & Webb, J. 2010. Impact of Performance Measurement and Management Systems. Management Services 54(3), 8-15.

Martínez-Costa, M. & Martínez-Lorente, A.R 2007. A triple analysis of ISO 9000 effects on company performance. International Journal of Productivity and Performance Management 56(5/6), 484-499.

Martins, R.A. 2002. The use of performance measurement information as a driver in designing a performance measurement system. Performance Measurement and Management: Research and Action, Centre for Business Performance, Boston, MA.

McAdam, R. & Bailie, B. 2002. Business performance measures and alignment impact on strategy – the role of business improvement models. International Journal of Operations & Production Management 22(9), 972-996.

McCrea, R., Shyy, T.-K. & Stimson, R. 2006. What is the strength of the link between objective and subjective indicators of urban quality of life? Applied Research in Quality of Life 1(1), 79-96.

McGee, J.V. 1992. What is strategic performance measurement? Center for Business Innovation. Arizona: Ernst & Young.

Meadows, D. H., Goldsmith, E. I. & Meadow, P. 1972. The limits to growth. London: Earth Island.

Melnyk, S.A., Sroufe, R.P. & Calantone, R. 2003. Assessing the impact of environmental management systems on corporate and environmental performance. Journal of Operations Management 21(3), 329–351.

Melnyk, S.A., Sroufe, R.P., Calantone, R. & Montabon, F.L. 2002. Assessing the effectiveness of US voluntary environmental programmes: an empirical study. International Journal of Production Research 40(8), 1853-1878.

Melville, N.P. 2010. Information systems innovation for environmental sustainability. MIS Quarterly 34(1), 1–21.

Mendoza, C. & Saulpic, O. 2002. Strategic management and management control: designing a new theoretical framework. In M.J. Epstein & J.F. Manzoni (Eds.) Performance Measurement and Management Control: A Compendium of Research. Oxford: Elsevier Science, 131-158.

Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. 2001. Defining supply chain management. Journal of Business logistics 22(2), 1-25.

Mir, R., & Watson, A. 2001. Critical realism and constructivism in strategy research: toward a synthesis. Strategic Management Journal 22(12), 1169-1173.

Mitchell, J.C 1983. Case and situation analysis. Sociological Review 31(3), 186-211.

Mohammed, M. 2000. The ISO 14001 EMS implementation process and its implications: a case study of Central Japan. Environmental Management 25(2), 177-188.

Moon, P. & Fitzgerald, L. 1996. Delivering the goods at TNT: the role of the performance measurement system. Management Accounting Research 7(4), 431-457.

Morrow, D. & Rondinelli, D. 2002. Adopting corporate environmental management systems: motivations and results of ISO 14001 and EMAS certification. European Management Journal 20(2), 159-171.

Neely, A. 1998. Measuring Business Performance: Why, What and How. London: Economist and Profile Books.

Neely, A. 2005. The evolution of performance measurement research: Developments in the last decade and research agenda for the next. International Journal of Operations & Production Management 25(12), 1264-1277.

Neely, A., Adams, C. & Kennerley, M. 2002. The Performance Prism: The Scorecard for Measuring and Managing Business Success. London: Prentice Hall.

Neely, A., Gregory, M. & Platts, K. 1995. Performance measurement system design – a literature review and research agenda. International Journal of Operations & Production Management 15(4), 80-116.

Neely, A., Richards, A.H., Mills, J.F., Platts, K.W. & Bourne, M.C.S. 1997. Designing performance measures: a structured approach. International Journal of Operations & Production Management 17(11), 1131-1152.

Newbold, J. 2006. Chile's environmental momentum: ISO 14001 and the large-scale mining industry – case studies from the state and private sector. Journal of Cleaner Production 14(3/4), 248-261.

Otley, D.T. 1999. Performance management: a framework for management control systems research. Management Accounting Research 10(4), 363-382.

Otley, D.T. 2006. Measuring performance: The accounting perspective. In A. Neely (Ed.) Business performance measurement. Cambridge: Cambridge University Press, 3-21.

Ou, C.S., Liu, F.C., Hung, Y.C. & Yen, D.C. 2010. A structural model of supply chain management on firm performance. International Journal of Operations & Production Management 30(5), 526-545.

Park, J. & Seaton, R. A. F. 1996. Integrative research and sustainable agriculture. Agricultural Systems, 50(1), 81-100.

Pavlov, A. & Bourne, M. 2011. Explaining the effects of performance measurement on performance. International Journal of Operations & Production Management 31(1), 101-122.

Pettigrew, A., Whipp, R. & Rosenfield, R. 1989. Competitiveness and the management of strategic change processes. In A. Francis & P.K.M. Tharakan (Eds.) The Competitiveness of European Industry: Country Policies and Company Strategies. London: Routledge.

Piatt, J. 1988. What can case studies do? In R.G. Burgess (Ed.) Conducting Qualitative Research. Greenwich: JAI Press.

Porter, M. E. & Kramer, M. R. 2011. Creating shared value. Harvard business review, 89(1/2), 62-77.

Porter, M.E. 1985. Competitive Advantage. New York, NY: The Free Press.

Poveda, C. A. & Lipsett, M. G. 2011. A Review of Sustainability Assessment and Sustainability/Environmental Rating Systems and Credit Weighting Tools. Journal of Sustainable Development 4(6), 36-55.

QSR International 2014. NVivo 10 for Windows. Retrieved August 29, 2014 from: http://www.qsrinternational.com/products_nvivo.aspx.

Remenyi, D., Williams, B., Money, A., & Swartz, E. 1998. Doing Research in Business and Management. London: Sage.

Robson, C. 2002. Real World Research, 2nd edition. Oxford: Blackwell.

Rondinelli, D. & Vastag, G. 2000. Panacea, common sense, or just a label? The value of ISO 14001 environmental management systems. European Management Journal 18(5), 499-510.

Roome NJ. 1998. Sustainability Strategies for Industry: the Future of Corporate Practice. Washington; Island.

Sampaio, P., Saraiva, P. & Monteira, A. 2012. ISO 9001 certification pay-off: myth versus reality. International Journal of Quality & Reliability Management 29(8), 891-914.

Sampaio, P., Saraiva, P. & Rodrigues, A.G. 2009. ISO 9001 certification research: questions, answers and approaches. International Journal of Quality & Reliability Management 26(1), 38-58.

Sampaio, P., Saraiva, P. & Rodrigues, A.G. 2010. A classification model for prediction of certification motivations from the contents of ISO 9001 audit reports. Total Quality Management and Business Excellence 21(12), 1279-1298.

Santos-Vijande, M.L., López-Sánchez, J.Á. & Trespalacios, J.A. 2012. How organizational learning affects a firm's flexibility, competitive strategy, and performance. Journal of Business Research 65(8), 1079-1089.

Saunders, M., Lewis, P., & Thornhill, A. 2009. Research methods for business students. Essex: Pearson.

Schaltegger, S. & Burritt, R., 2000. Corporate Sustainability. The International Yearbook of Environmental and Resource Economics, 185–222.

Schaltegger, S. & Hasenmüller, P. 2005. Nachhaltiges Wirtschaften aus Sicht des "business case of sustainability". Fachdialog des Bundesumweltministeriums (BMU), CSM, Lüneburg.

Schmenner, R.W. 1988.Escaping the black holes of cost accounting. Business Horizons 31(1), 66-72.

Schrettle, S., Hinz, A., Scherrer-Rathje, M. & Friedli, T. 2014. Turning sustainability into action: Explaining firms' sustainability efforts and their impact on firm performance. International Journal of Production Economics, 147(PA), 73-84.

Searcy, C. 2009a. Corporate sustainability performance measurement: Lessons from system of systems engineering. In Proceedings of the 2009 IEEE international conference on systems, man, and cybernetics, San Antonio, October 2009.

Searcy, C. 2009b. Setting a course for corporate sustainability performance measurement. Measuring Business Excellence 13(3), 49–57.

Searcy, C. 2012. Corporate sustainability performance measurement systems: A review and research agenda. Journal of business ethics 107(3), 239-253.

Sharma, S., Pablo, A.L. & Vredenburg, H. 1999. Corporate environmental responsiveness strategies: the importance of issue interpretation and organizational context. Journal of Applied Behavioral Science 35(1), 87–108.

Shaughnessy, J.J. & Zechmeister, E.B. 1990. Research Methods in Psychology. New York: McGraw-Hill.

Sheate, W., Dagg, S., O'Donnell, K., Richardson, J., Aschemann, R. & Steen, U. 2001. SEA and Integration of the Environment into Strategic Decision-Making. Final Report to the European Commission. Retrieved July 20, 2014 from: http://ec.europa.eu/environment/archives/eia/sea-studies-and-reports/pdf/sea_integration_main.pdf.

Sheate, W., Dagg, S., Richardson, J., Aschemann, R., Palerm, J. & Steen, U. 2003. Integrating the environment into strategic decision-making: conceptualizing policy SEA. European Environment 13(1), 1-18.

Shepherd, C. & Gunter, H. 2005. Measuring supply chain performance: current research and future directions. International Journal of Productivity and Performance Management 55(3/4), 242-258.

Silver, C. & Lewins, A. 2014. Using Software in Qualitative Research a step-by-step guide. London: Sage.

Simon, H. A., Guetzkow, H., Kozmetsky, G., & Tyndall, G. 1954. Centralization vs decentralization in organizing the controllers department. New York: Controllership Foundation.

Simons, R. 1991. Strategic orientation and top management attention to control systems. Strategic Management Journal 12, 49-62.

Skinner, W. 1974. The decline, fall, and renewal of manufacturing. Industrial Engineering 6(10), 32-38.

Skjott-Larsen, T., Schary, P.B. 2007. Managing the global supply chain. Copenhagen: Copenhagen Business School Press.

Slack, N. 1991. The Manufacturing Advantage: Achieving Competitive Manufacturing Operations. London: Mercury.

Slater, S.F. & Narver, J. C. 1995. Market orientation and the learning organization. The Journal of Marketing 59(3), 63-74.

Smith, P.C. & Goddard, M. 2002. Performance management and operational research: a marriage made in heaven? Journal of the Operational Research Society 53(3), 247-55.

Stake, R. 2000. Case studies. In N.K. Denzin & Y.S. Lincoln (Eds) Handbook of Qualitative Research, 2nd edition. Thousand Oaks: Sage.

Steger, U., Ionescu-Somers, A. & Salzmann, O. 2007. The economic foundations of corporate sustainability. Corporate Governance 7(2), 162-177.

Stern, N. (Ed.). 2007. The economics of climate change: the Stern review. Cambridge: Cambridge University Press.

Supply-Chain Council 2014. Our frameworks. Retrieved June 5, 2014 from: https://supply-chain.org/our-frameworks.

Szymanski, M. & Tiwari, P. 2004. ISO 14001 and reduction of toxic emissions. Journal of Economic Policy Reform 7(1), 31-42.

Tan, L.P. 2005. Implementing ISO 14001: is it beneficial for firms in newly industrialized Malaysia? Journal of Cleaner Production 13(4), 397-404.

Taticchi, P. & Balachandran, K.R. 2008. Forward performance measurement and management integrated frameworks. International Journal of Accounting and Information Management 16(2), 140-154.

Taticchi, P., Asfalti, A. & Sole, F. 2010a. Performance Measurement and Management in Smes: Discussion of Preliminar Results from an Italian Survey. In P. Taticchi (Ed.) Business Performance Measurement and Management. Heidelberg: Springer, 3-12.

Taticchi, P., Tonelli F. & Cagnazzo, L. 2010b. Performance measurement and management: a literature review and research agenda. Measuring Business Excellence 14(1), 4-18.

Taticchi, P., Tonelli, F. & Cagnazzo, L. 2009. A decomposition and hierarchical approach for business performance measurement and management. Measuring Business Excellence 13(4), 47-57.

Taticchi, P., Tonelli, F. & Pasqualino, R. 2013. Performance Measurement of sustainable supply chains – A literature review and a research agenda. International Journal of Productivity and Performance Management 62(8), 782-804.

Tien, S.W., Chung, Y.C. & Tsai, C.H. 2005. An empirical study on the correlation between environmental design implementation and business competitive advantages in Taiwan's industries. Technovation 25(7), 783-794.

Tippins, M.J. & Sohi, R.S. 2003. IT competency and firm performance: is organizational learning a missing link? Strategic Management Journal 24(8), 745-761.

Tonelli, F., Evans, S. & Cainarca, G. C. 2013. Industrial Sustainability. In P. Taticchi, P. Carbone & V. Albino (Eds.) Corporate Sustainability. Heidelberg: Springer, 27-58.

Turk, A.M. 2009. The benefits associated with ISO 14001 certification for construction firms: Turkish case. Journal of Cleaner Production 17(5), 559-569.

Turney, P.B.B. & Anderson, B. 1989. Accounting for continuous improvement. Sloan Management Review 30(2), 37-48.

Tyndall, G., Gopal, C., Partsch, W. & Kamauff, J. 1998. Supercharging Supply Chains: New Ways to Increase Value Through Global Operational Excellence. New York: John Wiley & Sons.

UN & Accenture 2013. CEO Study on Sustainability – Architects of a better world. Retrieved June 29, 2014 from: http://www.accenture.com /SiteCollectionDocuments/PDF/Accenture-UN-Global-Compact-Acn-CEO-Study-Sustainability-2013.PDF.

United Nations, Department of Economic and Social Affairs, Population Division 2013. World Population Prospects: The 2012 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP.227. New York.

Vallance, S., Perkins, H.C. & Dixon, J.E. 2011. What is Social Sustainability? A Clarification of Concepts. Geoforum 42(3), 342-348.

Vastag, G. & Melnyk, S.A. 2002. Certifying environmental management systems by the ISO 14001 standards. International Journal Of Production Research 40(18), 4743-4763.

Waggoner, D., Neely, A. & Kennerley, M. 1999. The forces that shape organizational performance measurement systems: an interdisciplinary review. International Journal of Production Economics 60(1), 53-60.

WCED 1987. Our Common Future. Report of the World Commission on Environment and Development. New York: United Nations (General Assembly, 42nd session, document A43/427 of 4 August).

Wheeler, D., Colbert, B. & Freeman, R.E. 2003. Focusing on value: reconciling corporate social responsibility, sustainability and a stakeholder approach in a network world. Journal of General Management 28(3), 1–28.

Wilson, M. 2003. Corporate sustainability: what is it and where does it come from? Ivey Business Journal 67(6), 1-5.

Yin, H.T. & Ma, C.B. 2009. International integration: a hope for a greener China? International Marketing Review 26(3), 348-367.

Yin, R.K. 1993. Applications of Case Study Research. Newbury Park: Sage.

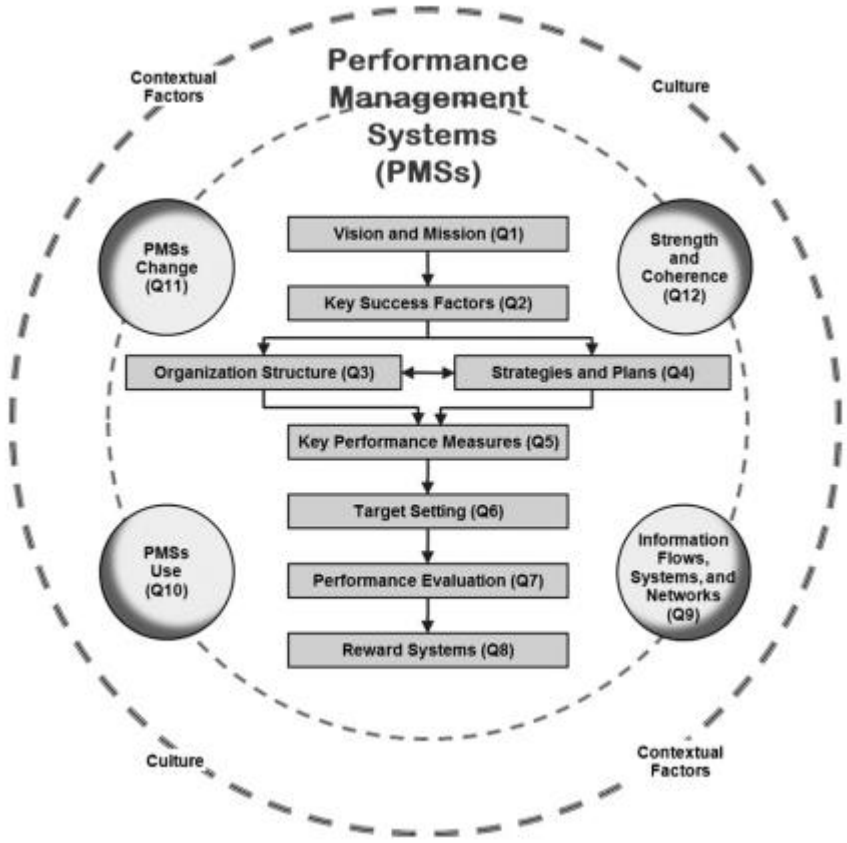
Yin, R.K. 2014. Case Study Research: Design and Methods, 5th edition. Thousand Oaks: Sage.

Zheng, W., Yang, B. & McLean, G.N. 2010. Linking organizational culture, structure, strategy, and organizational effectiveness: Mediating role of knowledge management. Journal of Business Research 63(7), 763-771.

Zutshi, A. & Sohal, A. 2004. Environmental management system adoption by Australasian organisations: Part 1: reasons, benefits and impediments. Technovation 24(4), 335-57.

APPENDICES

ANNEX 1 Performance Management Systems Framework (Ferreira & Otley 2009)



ANNEX 2 Original 5-Questions Performance Management System Framework (Otley 1999: 365-366)

Q1. What are the key objectives that are central to the organization's overall future success, and how does it go about evaluating its achievement for each of these objectives?

Q2. What strategies and plans has the organization adopted and what are the processes and activities that it has decided will be required for it to successfully implement these? How does it assess and measure the performance of these activities?

Q3. What level of performance does the organization need to achieve in each of the areas defined in the above two questions) and how does it go about setting appropriate performance targets for them?

Q4. What rewards will managers (and other employees) gain by achieving these performance targets (or, conversely, what penalties will they suffer by failing to achieve them)?

Q5. What are the information flows (feedback and feed-forward loops) that are necessary to enable the organization to learn from its experience) and to adapt its current behaviour in the light of that experience?

ANNEX 3 12-Questions Performance Management System Framework (Ferreira & Otley 2009: 266-267)

Q1. What is the vision and mission of the organization and how is this brought to the attention of managers and employees? What mechanisms, processes, and networks are used to convey the organization's overarching purposes and objectives to its members?

Q2. What are the key factors that are believed to be central to the organization's overall future success and how are they brought to the attention of managers and employees?

Q3. What is the organization structure and what impact does it have on the design and use of performance management systems (PMSs)? How does it influence and how is it influenced by the strategic management process?

Q4. What strategies and plans has the organization adopted and what are the processes and activities that it has decided will be required for it to ensure its success? How are strategies and plans adapted, generated and communicated to managers and employees?

Q5. What are the organization's key performance measures deriving from its objectives, key success factors, and strategies and plans? How are these specified and communicated and what role do they play in performance evaluation? Are there significant omissions?

Q6. What level of performance does the organization need to achieve for each of its key performance measures (identified in the above question), how does it go about setting appropriate performance targets for them, and how challenging are those performance targets?

Q7. What processes, if any, does the organization follow for evaluating individual, group, and organizational performance? Are performance evaluations primarily objective, subjective or mixed and how important are formal and informal information and controls in these processes?

Q8. What rewards — financial and/or non-financial — will managers and other employees gain by achieving performance targets or other assessed aspects of performance (or, conversely, what penalties will they suffer by failing to achieve them)?

Q9. What specific information flows — feedback and feed-forward —, systems and networks has the organization in place to support the operation of its PMSs?

Q10. What type of use is made of information and of the various control mechanisms in place? Can these uses be characterised in terms of various typologies in the literature? How do controls and their uses differ at different hierarchical levels?

Q11. How have the PMSs altered in the light of the change dynamics of the organization and its environment? Have the changes in PMSs design or use been made in a proactive or reactive manner?

Q12. How strong and coherent are the links between the components of PMSs and the ways in which they are used (as denoted by the above 11 questions)?

PMS = Performance Management System

ANNEX 4 Top10 PMSs Implementation Problems Ranking according to Practitioners and Academics (de Waal & Counet 2009)

Rank	Practitioners' Perspective	Academics' Perspective
1	The organization does not have a performance management culture	The current ICT system does not support the PMS adequately
2	Lack of management commitment	The organization is in an unstable phase
3	Management puts low priority on the PMS implementation	The PMS has a low priority or its use is abandoned after a change of management
4	The organization does not see (enough) benefit from the PMS	Lack of management commitment
5	The PMS has a low priority or its use is abandoned after a change of management	The organization does not have a performance management culture
6	Organizational members are not adopting the right management style	The organization does not have a clear and understandable strategy
7	There is resistance from organizational members towards the new PMS	The PMS is not regularly updated and maintained after implementation
8	There is too much focus on the results of the PMS implementation, while the change process of the organization is ignored	There is resistance from organizational members towards the new PMS
9	The organization does not have a clear and understandable strategy	The system lacks cause and effect relations or is over-complex due to too many causal relation
10	There are too many KPIs defined	The PMS is not used for the daily management of the organization

ANNEX 5 Case Study Protocol

For Tool 1/Tool 2 Interviews

Purpose

To obtain the perspective of the respondent managing the process of sustainability assessment on its potential impact on the ordering organization's performance. The questions are designed to elicit information about:

1. The role of the interviewee
2. Information about provider organization
3. Information about client organizations
4. The process of the sustainability assessment in general and for specific cases
5. Adaptations that have been made to customize the assessment to better fit to organizations
6. Particularities of the case, especially in relation to dealing with sustainability

Before-Phase (From initial idea to the actual beginning of the assessment)

What led to the Sustainability Assessment, what was the motivation?

Have there been any (strategic) objectives? What were they? Has it been clear which role the measurement should have (Measure Performance/Strategy Management/ Communication/Influence Behavior/Learning and Improvement)

What led to the decision to use Tool 1/Tool 2, have any alternative solutions been considered? Why assess Sustainability (three Dimensions: Economy, Environment, and Society)?

Have there been any alterations to Tool 1/Tool 2 to align the measurement to the objectives, the strategy, the organizational size/structure (type)/culture (values) of the company?

Have there been any other Measurement or Management Systems in place within the organization?

What was the level of understanding regarding sustainability and sustainability assessment?

During-Phase (From beginning the assessment to final presentation of the results and recommendations)

Were there any alterations? For what reason?

Were there any complications?

After-Phase (From receiving the final results and recommendations to now)

Were the results communicated? To which stakeholders?

Were the recommendations successfully implemented? How many of the recommendations? Why and/or why not? What would have to happen to successfully implement the recommendations?

Is there a system in place to continue the assessment of the company's sustainability performance?

General

How much time passed between the initial idea to ordering the assessment to receiving the recommendations to implementation?

How was the project managed? How many people were actively involved at the before- and during-stage from the company-side? How many people were involved to implement the recommendations? Are there still people actively involved? Who took/takes the decisions?

What went well/not so well?

Was the communication in the customer's native language?

ANNEX 6 Exemplary Two-Pager Sent with First Email

Philipp Müller
Company Tool 1, Country Tool 1
University of Jyväskylä, Finland
Email: philipp.mueller@company-tool1.com

Supervisor: Earth
Email: earth@company-tool1.com

Thesis: Strategic Role of External Sustainability Assessment in the Food Value Chain

About the author

After completing my previous studies in Business and Administration, I went on to pursue the topic of Sustainability and started a Master Degree in Corporate Environmental Management in Jyväskylä, Finland. One part of my studies was a voluntary internship at Company Tool 1, where I mainly dealt with Sustainability Assessment. Seeing how much effort it takes to conduct a study while the results often remain untouched, sparked the idea to closer investigate this phenomenon.

What is the thesis about?

By conducting interviews with different actors along the Food Value Chain who had a Sustainability Assessment carried out, this research intends to identify barriers and enablers around the process of the assessment. The questions that should be covered during the interview can be found below. The motivation stems from the fact that despite the resource intensity of carrying out a Sustainability Assessment, experience and previous research shows that recommendations based on the results are often not or only partially implemented, leaving much of the potential unrealized.

What does it take to participate?

Only about 1 hour of the participants' time for the interviews to talk about their experiences before, during, and after the Sustainability Assessment. The interviews will mostly be carried out via telephone in English or German. The interviews and results will be anonymized to ensure a high level of confidentiality.

How do participants benefit?

After the research is successfully completed, participants will receive a feedback on the results that they are free to use. The clearer understanding of typical barriers and enabling measures will support a more efficient design and execution of future performance related processes.

(continues)

ANNEX 6 (continues)

Subjects to be covered during the interview

General Information

- Information about the organization, its vision, mission and business objectives, the organizational structure and its environment, and sustainability practices

Before-Phase (From initial idea to the actual beginning of the assessment)

- What led to the Sustainability Assessment, what was the motivation? Why use Tool 1/2? Why assess Sustainability (all three Dimensions: Economy, Environment, and Society)?
- Were there any (strategic) objectives? What were they? Has it been clear which role the assessment should have (Measure Performance/ Communication (Marketing)/ Strategy Management/ Influence Behavior/ Learning and Improvement)?
- Were there any alterations to Tool1/2 to align the measurement to the objectives, the strategy, the organizational size/structure (type)/culture (values) of the company?
- Were any other Measurement or Management Systems in place within the organization? Any impact assessments that have been conducted previously?
- Was there a clear understanding of what could be expected from the assessment, what the end result would look like?

During-Phase (From beginning the assessment to final presentation of the results and recommendations)

- Were there any alterations? For what reason?

After-Phase (From receiving the final results and recommendations to now)

- Were the results communicated? To which stakeholders?
- Were the results what you expected? Did any changes occur based on the results? What would have to be different for changes to take place? What lacked?
- Is there a system in place to continue the assessment of the company's sustainability performance?

General questions about the assessment process

- How was the project managed? How many people were actively involved from the company-side? Who takes decisions?
- What went well/ not so well? Which factors were beneficial/ hindering to the process?
- Was the communication in your native language?

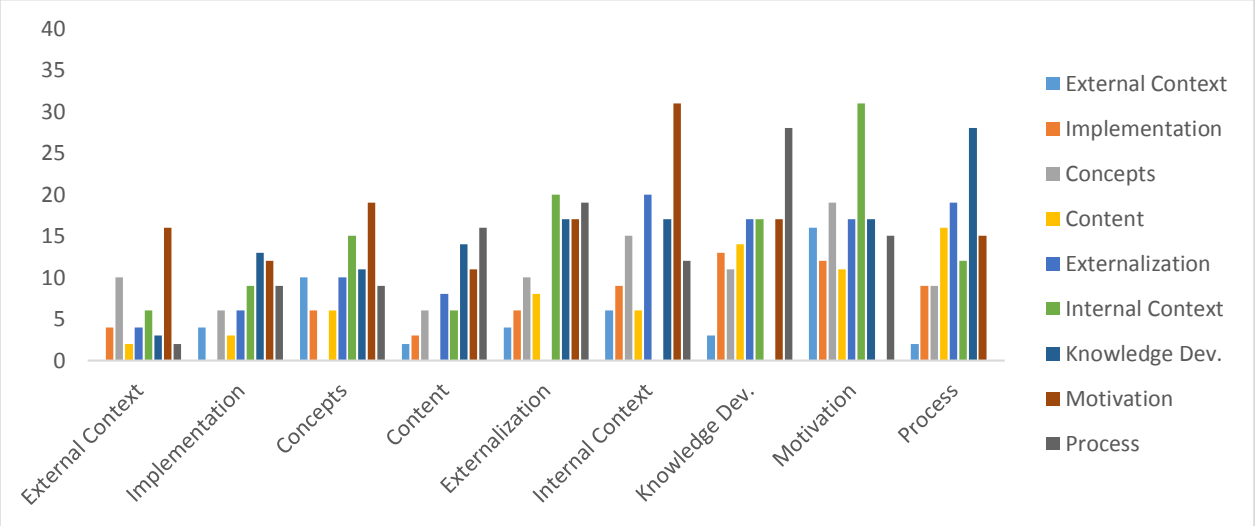
ANNEX 7 Interviews Overview

	Participant	Mode	Duration	Pages	References
Tool 1	**Sun**	Telephone	1h 31min	13	578
	Mercury	Telephone	1h 3min	9	387
	Venus	Telephone	48min	8	212
	Earth	Face-to-face	57min	14	491
	Mars	Telephone	52min	10	386
Tool 2	**Jupiter**	Telephone	57min	11	264
	Saturn	Telephone	55min	10	144
	Uranus	Telephone	1h8min	12	486
	Neptune	Telephone	57min	11	295

ANNEX 8 Combined Code-Matrix Results (Values greater-than-or-equal to average overlap highlighted in green)

Tool 1 + 2 (abs.)	***Total***	**External Context**	**Implementation**	**Concepts**	**Content**	**Externalization**	**Internal Context**	**Knowledge Dev.**	**Motivation**	**Process**	**Ø**
External Context	*22*		4	10	2	4	6	3	16	2	*5,88*
Implementation	*38*	4		6	3	6	9	13	12	9	*7,75*
Concepts	*43*	10	6		6	10	15	11	19	9	*10,8*
Content	*47*	2	3	6		8	6	14	11	16	*8,25*
Externalization	*59*	4	6	10	8		20	17	17	19	*12,6*
Internal Context	*60*	6	9	15	6	20		17	31	12	*14,5*
Knowledge Dev.	*61*	3	13	11	14	17	17		17	28	*15*
Motivation	*71*	16	12	19	11	17	31	17		15	*17,3*
Process	*82*	2	9	9	16	19	12	28	15		*13,8*

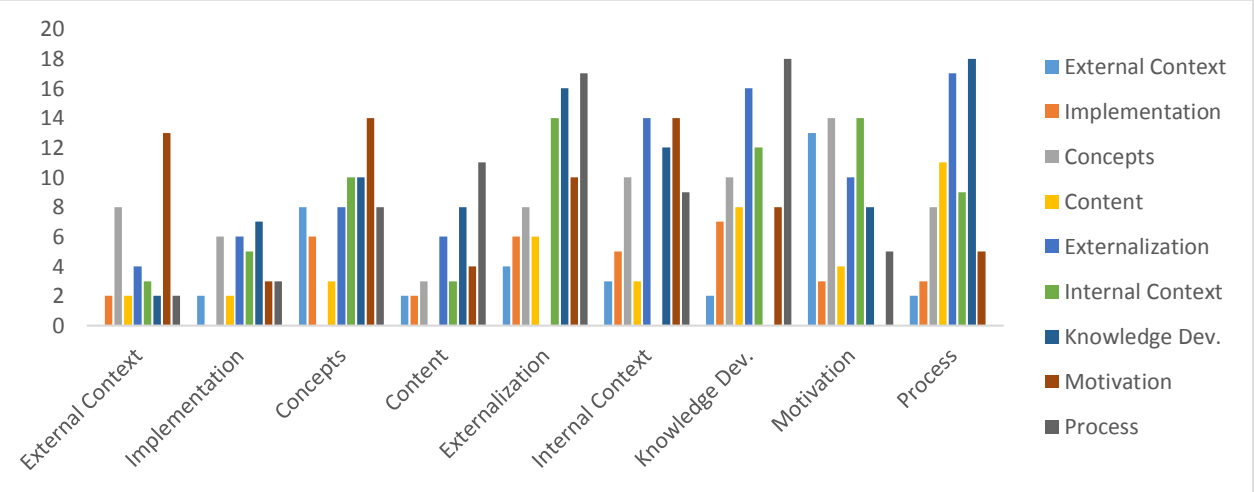
Tool 1 + 2 (%)	***Total***	**External Context**	**Implementation**	**Concepts**	**Content**	**Externalization**	**Internal Context**	**Knowledge Dev.**	**Motivation**	**Process**	**Ø**
External Context	*22*		18%	45%	9%	18%	27%	14%	73%	9%	*27%*
Implementation	*38*	11%		16%	8%	16%	24%	34%	32%	24%	*20%*
Concepts	*43*	23%	14%		14%	23%	35%	26%	44%	21%	*25%*
Content	*47*	4%	6%	13%		17%	13%	30%	23%	34%	*18%*
Externalization	*59*	7%	10%	17%	14%		34%	29%	29%	32%	*21%*
Internal Context	*60*	10%	15%	25%	10%	33%		28%	52%	20%	*24%*
Knowledge Dev.	*61*	5%	21%	18%	23%	28%	28%		28%	46%	*25%*
Motivation	*71*	23%	17%	27%	15%	24%	44%	24%		21%	*24%*
Process	*82*	2%	11%	11%	20%	23%	15%	34%	18%		*17%*



ANNEX 9 Code-Matrix Results for Tool 1 (Values greater-than-or-equal to average overlap highlighted in green)

Tool 1 (abs.)	***Total***	**External Context**	**Implementation**	**Concepts**	**Content**	**Externalization**	**Internal Context**	**Knowledge Dev.**	**Motivation**	**Process**	**Ø**
External Context	*17*		2	8	2	4	3	2	13	2	*4,5*
Implementation	*11*	2		6	2	6	5	7	3	3	*4,25*
Concepts	*30*	8	6		3	8	10	10	14	8	*8,38*
Content	*23*	2	2	3		6	3	8	4	11	*4,88*
Externalization	*42*	4	6	8	6		14	16	10	17	*10,1*
Internal Context	*34*	3	5	10	3	14		12	14	9	*8,75*
Knowledge Dev.	*37*	2	7	10	8	16	12		8	18	*10,1*
Motivation	*28*	13	3	14	4	10	14	8		5	*8,88*
Process	*49*	2	3	8	11	17	9	18	5		*9,13*

Tool 1 (%)	***Total***	**External Context**	**Implementation**	**Concepts**	**Content**	**Externalization**	**Internal Context**	**Knowledge Dev.**	**Motivation**	**Process**	**Ø**
External Context	*17*		12%	47%	12%	24%	18%	12%	76%	12%	*26%*
Implementation	*11*	18%		55%	18%	55%	45%	64%	27%	27%	*39%*
Concepts	*30*	27%	20%		10%	27%	33%	33%	47%	27%	*28%*
Content	*23*	9%	9%	13%		26%	13%	35%	17%	48%	*21%*
Externalization	*42*	10%	14%	19%	14%		33%	38%	24%	40%	*24%*
Internal Context	*34*	9%	15%	29%	9%	41%		35%	41%	26%	*26%*
Knowledge Dev.	*37*	5%	19%	27%	22%	43%	32%		22%	49%	*27%*
Motivation	*28*	46%	11%	50%	14%	36%	50%	29%		18%	*32%*
Process	*49*	4%	6%	16%	22%	35%	18%	37%	10%		*19%*



ANNEX 10 Code-Matrix Results for Tool 2 (Values greater-than-or-equal to average overlap highlighted in green)

Tool 2 (abs.)	***Total***	**External Context**	**Implementation**	**Concepts**	**Content**	**Externalization**	**Internal Context**	**Knowledge Dev.**	**Motivation**	**Process**	***Ø***
External Context	*5*		2	2	0	0	3	1	3	0	*1,38*
Implementation	*27*	2		0	1	0	4	6	9	6	*3,5*
Concepts	*13*	2	0		3	2	5	1	5	1	*2,38*
Content	*24*	0	1	3		2	3	6	7	5	*3,38*
Externalization	*17*	0	0	2	2		6	1	7	2	*2,5*
Internal Context	*26*	3	4	5	3	6		5	17	3	*5,75*
Knowledge Dev.	*24*	1	6	1	6	1	5		9	10	*4,88*
Motivation	*43*	3	9	5	7	7	17	9		10	*8,38*
Process	*33*	0	6	1	5	2	3	10	10		*4,63*

Tool 2 (%)	***Total***	**External Context**	**Implementation**	**Concepts**	**Content**	**Externalization**	**Internal Context**	**Knowledge Dev.**	**Motivation**	**Process**	***Ø***
External Context	*5*		40%	40%	0%	0%	60%	20%	60%	0%	*28%*
Implementation	*27*	7%		0%	4%	0%	15%	22%	33%	22%	*13%*
Concepts	*13*	15%	0%		23%	15%	38%	8%	38%	8%	*18%*
Content	*24*	0%	4%	13%		8%	13%	25%	29%	21%	*14%*
Externalization	*17*	0%	0%	12%	12%		35%	6%	41%	12%	*15%*
Internal Context	*26*	12%	15%	19%	12%	23%		19%	65%	12%	*22%*
Knowledge Dev.	*24*	4%	25%	4%	25%	4%	21%		38%	42%	*20%*
Motivation	*43*	7%	21%	12%	16%	16%	40%	21%		23%	*19%*
Process	*33*	0%	18%	3%	15%	6%	9%	30%	30%		*14%*