

Towards an international Corporate Sustainability Performance Measurement System

Applied to the Dairy Sustainability Framework by means of an internship at FrieslandCampina

Sietske Drost 5533511



Master's Thesis

Sustainable Development – track Environmental Governance

45 ECTS

June 29, 2016

Towards an international Corporate Sustainability Performance Measurement System

Applied to the Dairy Sustainability Framework by means of an internship at FrieslandCampina

Confidential

Student

Sietske Drost
Student number: 5533511
Weerdsingel O.Z. 51A
3514 AE Utrecht
Phone number: 0621494966
Email address: s.e.drost@students.uu.nl

Supervisor Utrecht University

Dr. Walter J.V. Vermeulen
Heidelberglaan 2
3584 CS Utrecht
Room 1126a
Phone number: + 31 30 253 2989
Email address: W.J.V.Vermeulen@uu.nl

Internship Organization

FrieslandCampina
Stationsplein 4
3818 LE Amersfoort

Supervisor Internship Organization

Jaap Petraeus
Manager Corporate Environmental affairs & Sustainability
E-mail address: Jaap.Petraeus@frieslandcampina.com





Towards an international Corporate Sustainability Performance Measurement System

Sietske Drost



Summary

This research focuses on the relevance of a Performance Measurement System (PMS) for the improvement of Corporate Sustainability (CS). The actual design of a PMS is taken into consideration and applied to the dairy industry by using the Dairy Sustainability Framework (DSF) and is supported by an internship at FrieslandCampina.

In the last years, environmental issues have played an increased role for businesses. Due to this, CS has become more important in order to describe the sustainable contribution of organizations. One of the sectors that face challenges regarding sustainability is the food industry. The food industry has to deal with these due to the increasing world population, increased welfare, and in turn, the higher consumption of food. In response to this, the food industry has been intensified and concerns about the long-term maintenance of agricultural production have been raised. As a result, many organizations are shifting towards a more CS based production strategy to ensure sustainable food production. One of the industries that is moving to more sustainable practices is the dairy industry, which this thesis will focus on.

The objective of the research conducted in this thesis is to make sustainability in the dairy industry measurable and to develop a PMS within the DSF. This should provide information about the performance of dairy producers regarding continuous improvement of sustainability, as well as how to bring the entire dairy industry to a higher sustainability level. In order to answer the research questions, the research is subdivided in three parts. In the first part, literature research is executed in order to build a scientific background. In the second part, empirical research is performed by conducting 26 in-depth interviews and an analysis of relevant comparable cases. Finally, in the third part, knowledge derived from the first two parts is used to apply a design of a PMS to the dairy industry.

In this study the PMS contains 5 levels of maturity, takes into account long-term sustainability predictions and should continuously be kept up to date. Empirical data shows that strong communication networks and in turn transparency are crucial to improve CS. Additionally, stakeholder- and societal engagement are important elements to consider when measuring organizational performance. The system applied to dairy will be a self-assessment, so that organizations can measure their own sustainability practices and get insights in how to improve these.

Taking into account the conclusions and discussion of the research, recommendations for further research and to the internship organization are given in order to be able to set further steps towards the implementation of the PMS for sustainable dairy production.



Preface

This thesis is written for the Master program Sustainable Development, following the track Environmental Governance at the Faculty of Geosciences at Utrecht University. In order to fulfil the research provided in this thesis, an internship was carried out at the Central Office of dairy producer Royal FrieslandCampina NV, located in Amersfoort. The period that was reserved for this research was approximately 8 months, i.e. 32 weeks. Within this time-frame desk research was executed, interviews were conducted and internship observations were made in order to answer the research questions in this thesis.

During the two years master's program I have learned a lot and with this research all knowledge that I have obtained during my study can be applied to a specific issue. I have always been interested in sustainable food production and consumption and in this final thesis I have been able to do research in this specific subject.

Acknowledgements

This thesis would not have been possible to complete without the help and support of various people and organizations. Here I would like to give a special thanks to several people who have helped me over the past few months. First of all, I would like to thank my intern supervisor, Jaap Petraeus, for the great opportunities he has given me to get insights in the company and get experience in this working area. This internship has brought me unique opportunities to gain experience and to execute the research for my thesis. Also, a big thanks goes out to my 'colleagues' during the internship for their support, contributing their ideas and their advice giving role in the project.

Additionally, I would like to thank my supervisor at the University, Walter Vermeulen, for his feedback and supervision during the entire process. Furthermore, I am very thankful for the collaboration of the twenty-six interviewees. Without their help I could not complete this research. Subsequently, a big thank you goes to all information providers that I have not mentioned specifically. Everyone's input was extremely helpful.

Last but not least, I would like to thank my family and friends for all the support and help they gave me during these months and the final reading through of my thesis. Thanks for everything.

Sietske Drost

*Utrecht, the Netherlands
June 29, 2016*

List of figures

Figure 1	Dairy Sustainability Framework (DSF, 2014)	3
Figure 2	Research Framework	8
Figure 3	Corporate sustainability, the triple bottom line (Azapagic, 2003; Elkington, 2004)	12
Figure 4	DSF Demonstration Scheme (FrieslandCampina, 2015)	45
Figure 5	Plan-do-check-act cycle (Moen & Norman, 2006)	46
Figure 6	Regional Classification	50
Figure 7	Artificial impression of the results of one of the columns of DSF	53
Figure 8	Bar chart that shows results	54

List of tables

Table 1	Types of interviews	10
Table 2	Brief overview of main articles for maturity levels	17
Table 3	First set of criteria for a Performance Measurement System	20
Table 4	Levels of maturity of main articles	21
Table 5	Levels of maturity in the category Approach & Vision	24
Table 6	Levels of maturity in the categories Cooperation & Engagement and Communication	25
Table 7	Levels of maturity in the categories Data collection & Measuring and Goals achievement & Commitment	27
Table 8	Levels of maturity in the category Reporting	28
Table 9	Overview of the comparable cases used in this study	30
Table 10	Continuous improvement of the system	37
Table 11	Second set of criteria for a Performance Measurement System	38
Table 12	Levels of maturity in the category Approach & Vision	39
Table 13	Levels of maturity in the category Communication	39
Table 14	Levels of maturity in the category Materiality Assessment	41
Table 15	Levels of maturity in the categories Reporting, Output & Results	41
Table 16	Regional division	49

List of abbreviations

BSC	Balanced Scorecard
CO ₂ PL	CO ₂ Prestatieladder/ CO ₂ Performance Ladder
CDP	Carbon Disclosure Project
CMM	Capability Maturity Model
CS	Corporate Sustainability
CSR	Corporate Social Responsibility
DJSI	Dow Jones Sustainability Index
DSF	Dairy Sustainability Framework
GDA	Global Dairy Agenda for Action
GRI	Global Reporting Initiative
ISO	International Organization for Standardization
KPI	Key Performance Indicator
MVOPL	MVO Prestatieladder (Maatschappelijk Verantwoord Ondernemen)
PMS	Performance Measurement System
PDCA	Plan-Do-Check-Act
SCOR	Supply Chain Operations Reference Model
SPMS	Sustainability Performance Measurement System
SQ	Sub-question



Contents

Summary	VII
Preface	VIII
Acknowledgements	IX
List of figures	X
List of tables.....	X
List of abbreviations	XI
1. Introduction.....	1
1.1 Societal background.....	1
1.2 Scientific background and previous studies	2
1.3 Sustainability challenges for dairy businesses.....	2
1.4 Knowledge gap.....	4
1.5 Scientific relevance	4
1.6 Societal relevance.....	4
1.7 Aim and research questions	5
1.8 Outline of the thesis.....	5
2. Methodology	7
2.1 Research strategy.....	7
2.2 Data collection	9
2.3 Data analysis	10
2.4 Reliability and validity of the research	11
3. Results part one: Theoretical background & Literature review	12
3.1 Corporate Sustainability	12
3.2 Sustainability standards.....	13
3.3 The relevance of a Performance Measurement System	14
3.4 Introduction of leading articles	17
3.5 Towards an intermediate conclusion.....	19
3.5.1 First set of requirements and conditions for a good working PMS.....	19
3.5.2 Characteristics of maturity levels from literature	21
4. Results part two: Comparable cases & Lessons from practice	30
4.1 Comparable cases.....	30
4.1.1 Introduction of the cases.....	30
4.1.2 Scope of the cases	32
4.1.3 Reliability of the cases	32



- 4.1.4 Methodology of the cases 33
- 4.1.5 Process vs performance 34
- 4.1.6 Main advantages and disadvantages 35
- 4.2 Towards an intermediate conclusion 36
 - 4.2.1 Second set of requirements and conditions for a good working PMS.. 36
 - 4.2.2 Characteristics of maturity levels from empirical research 38
- 5. Results part one and two: Synthesis & PMS Design for dairy 43
 - 5.1 Main differences and similarities 43
 - 5.2 Towards an intermediate conclusion 44
 - 5.2.1 The structure of the PMS for the dairy industry 44
 - 5.2.2 Entry level..... 47
 - 5.2.3 Differentiation in regions 48
 - 5.2.4 Differentiation in types of organizations 50
 - 5.2.5 Continuous improvement of the PMS 51
 - 5.2.6 Process vs performance 51
 - 5.2.7 Questionnaire for self-assessment 52
 - 5.2.8 Display of results 53
- 6. Results part three: Further application to the dairy industry..... 56
 - 6.1 The situation of FrieslandCampina 56
 - 6.2 Pilot findings 56
 - 6.3 Towards an intermediate conclusion..... 58
- 7. Discussion 59
- 8. Conclusions 61
- 9. Recommendations to business 63
- References 65
- Appendices..... 71
 - Appendix A. List of interviews 71
 - Appendix B. List of observations..... 72
 - Appendix C. Format of the interview 75
 - Appendix D. Interview presentation 77
 - Appendix E. Brief explanation BSC and SCOR 82
 - Appendix F. Characteristics derived from literature 83
 - Appendix G. Characteristics derived from interviews and comparable cases .. 85
 - Appendix H. Quickscan **Fout! Bladwijzer niet gedefinieerd.**



1. Introduction

1.1 Societal background

The role of businesses in society has made large changes over the last decades (Steering Committee of the State-of-Knowledge Assessment of Standards and Certification, 2012). One of the main drivers for companies to change their production behavior is the changing environment and increase in world population, which is leading to higher consumption (GDAA, 2011). Currently the world faces many challenges. These include meeting the increased food demand due to an increased and wealthier population and producing in a more environmental and societal sustainable manner (Charles et al., 2010).

An important change in the role of businesses in society, and in relation to these challenges, is the shift from rather economic company values towards more social oriented company values. From this change in company values, corporate social responsibility (CSR)² emerged. Through CSR, not only economic profits are seen as important, but also involving society and taking into account environmental impacts (Vermeulen & Witjes, 2016). These are known as the three dimensions of people, planet and profit (three P's). This triple bottom line is created by Elkington (1997) and was a breakthrough in sustainability issues. Moreover, CSR supports cooperation between companies, resulting in more knowledge exchange (Watanatada & Mak, 2011) and thus positively contributing to the business functioning.

In the subsequent years, environmental issues have played an increasing role and therefore CSR has become limited (Vermeulen & Witjes, 2016). In turn, this increased attention led to the development of corporate sustainability (CS) (Baumgartner & Ebner, 2010). CS is defined as the implementation of strategies by individual companies to achieve long-term sustainable development, considering all three P's (Baumgartner & Ebner, 2010; Vermeulen & Witjes, 2016). CS focuses more on sustainability of the company rather than all possible contributing actions a company can do besides their own business activities. For this reason, the study conducted in this thesis will focus on CS instead of CSR.

One of the industries currently struggling with these challenges is the food industry. Due to an increasing world population, increase in welfare and higher consumption, an increase in global food demand takes place (GDAA, 2011). To meet the increased food demand, agricultural practices have been intensified over the past years. However, competition for land, water and energy is growing (Charles et al., 2010). As a result of the negative effects that result from intensified food production, concerns about the long-term maintenance of agricultural production have arisen (Matson et al., 1997). Many companies are shifting towards a more CS based production strategy to ensure sustainable food production now and in the future (Baumgartner & Ebner, 2010).

One of the branches within the food industry that is working towards becoming more sustainable is the dairy industry. The growing Dutch dairy industry aims to produce dairy products in a climate neutral manner, taking into account animal welfare, outdoor grazing and preservation of biodiversity and the environment (Roland Berger Strategy Consultants, Nederlandse Zuivel Organisatie, & ZuivelNL, 2015). One of the biggest dairy producers that has sustainability high on the agenda is Royal FrieslandCampina NV (FrieslandCampina, n.d.²). Due to the fact that FrieslandCampina is moving towards sustainable production and processing, and that it makes use of CS strategies, this company has been chosen for the internship for this research. A more specific

²CSR could be defined as *“business’ commitment to contribute to sustainable economic development, working with employees, their families, the local community, and society at large to improve their quality of life.”* (Kotler & Lee, 2005, p.3).

description of FrieslandCampina and its sustainability strategies will be given in section 6 of this thesis.

1.2 Scientific background and previous studies

Sustainable development in businesses and, more specifically, CS are widely investigated in previous studies. The long-term sustainability strategies related to CS consist of five key themes: towards (1) climate control and a renewable based society, (2) a circular economy in which less materials will be used, (3) prevention of negative impacts on biodiversity and animal welfare, (4) an economy that is not only based on monetary values, but also includes community well-being, and (5) the creation of a fair market (Vermeulen & Witjes, 2016). To respond to the sustainability challenges organizations face today, there must be a method to measure their progress towards sustainable development (Azapagic & Perdan, 2000). However, the question is, how do organizations deal with these sustainability issues and how will they determine their progress? In order to determine organization's progress and define strategies towards sustainability, certain performance measurement systems (PMS)³ have been developed (Epstein & Roy, 2001).

In relation to this, studies are being conducted into PMS's. PMS as a concept has been growing since the eighties and early nineties (Folan & Browne, 2005). Subsequently, different definitions of a PMS have developed and a division is made between traditional PMS's and innovative PMS's. According to Bourne et al. (2000), traditional PMS's have been criticized by scientific research for various reasons. Most commonly mentioned were; the systems are only focusing on short-term results, lack the focus on the business' strategies and aim only to draw attention to variance rather than continuous improvements. In this study an innovative PMS will be used, because this type has a long-term orientation and it monitors improvement as opposed to only making a comparison with the standard (McCormack et al., 2008). Within PMS's, maturity models are widely used to compare businesses in their performance and development stages by dividing several levels of maturity (McCormack et al., 2008). These types of systems support companies to move towards sustainable development⁴ (Cagnin et al., 2005). Certain levels should therefore be included in the PMS⁵. Useful reports regarding sustainability strategies in practice are found in i.e. Baumgartner and Ebner (2010), Cagnin, et al. (2005), Rietbergen and Blok (2013) and McCormack et al. (2008). These reports describe how levels of maturity can be divided, how certain systems work and assess existing performance systems.

1.3 Sustainability challenges for dairy businesses

For some specific sectors or issues, various measurement systems already exist (e.g. CO₂ performance ladder in construction industry). However, the dairy industry does not yet have a definitive measurement system with levels of maturity. But, since the dairy industry in the Netherlands, and on a wider international scale, is moving towards more sustainable production (Roland Berger Strategy Consultants, Nederlandse Zuivel Organisatie, & ZuivelNL, 2015), there is a need to measure this improvement in a solid manner. This shows that there is a certain demand for a corporate sustainability PMS.

In addition, the dairy industry has recognized the environmental and sustainability challenges that come along with the production of dairy products and has set up the Global Dairy Agenda for Action (GDAA). The GDAA was launched in 2009 and originated

³ "A successful performance measurement system is a set of performance measures that provides a company with useful information that helps manage, control, plan, and perform the activities undertaken by the company". (Searcy, 2012, p.241).

⁴ Continuous improvement is a process in which a company moves from its current maturity level towards a higher maturity level (Lockamy & McCormack, 2004).

⁵ For this study, the term PMS will be used for a performance measurement system which contains levels of maturity

from the cooperation of several large dairy organizations⁶. The goal of the GDAA is to reduce the greenhouse gas emissions and negative environmental impacts from the whole dairy value chain. This is done by providing governance for the global dairy industry's achievements in addressing its sustainability challenges (DSF, 2014). Important obligations⁷ were made and signed for during the World Dairy Summit in Berlin in 2009 (GDAA, 2011). By following these obligations, the GDAA supports efficient dairy production to contribute to global sustainability (GDAA, 2011).

The GDAA, together with the consultancy company SustainAbility, set up a framework, the Dairy Sustainability Framework (DSF) (Figure 1), to address sustainability in the dairy industry. The DSF is a result of over 100 individual interviews, several global meetings and workshops, and over 100 worldwide dairy- and non-dairy sustainability initiatives.

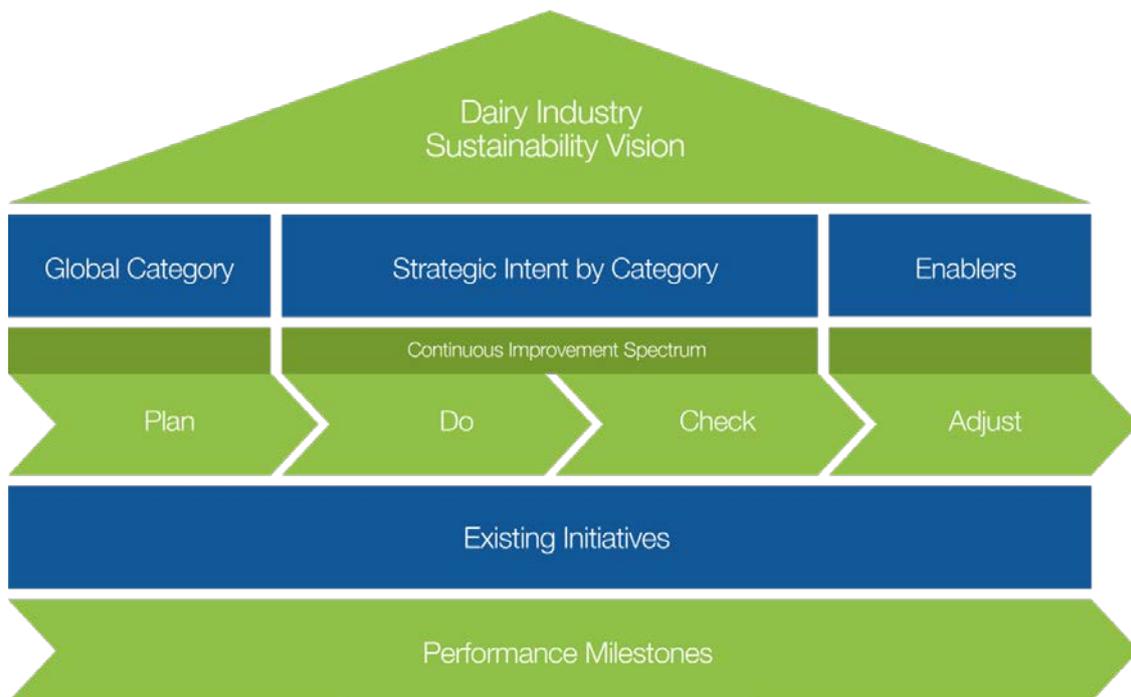


Figure 1. Dairy Sustainability Framework (DSF, 2014)

This framework presents a continuous improvement approach on how to implement sustainability strategies within organizations in the dairy industry (DSF, 2014). Although not all companies perform uniformly on the sustainability issues, the continuous improvement approach of the DSF empowers a global increase of sustainability within the dairy industry (DSF, 2014). However, the DSF has not yet defined levels for sustainability performance.

The DSF is a useful foundation in the process towards designing a corporate sustainability PMS. However, the problem definition is: how can the DSF be used in order to demonstrate sustainability in maturity levels in accordance with scientific literature

⁶ GDAA consists of the European Dairy Association (EDA), Eastern and Southern African Dairy Association (ESADA), Pan-American Dairy Federation (FEPALE), Global Dairy Platform (GDP), International Dairy Federation (IDF) and Sustainable Agricultural Initiative Platform (SAI).

⁷ The five obligations made and signed in Berlin: (1) promotion of the development of a standard methodology framework for evaluating the carbon footprint of dairy products, (2) promotion of the adoption of world's best practices within the dairy sector, (3) providing tools to measure and monitor emissions from production and processing, (4) improvement of farmers' understanding of their emissions and opportunities to reduce them on farm level and (5) promotion of the share of information and the alignment of research efforts to develop cost effective mitigation technologies for production and processing.

and from practical examples? How can a connection be made between a scientifically reliable, properly working system for continuous performance measurement (which should meet the long term sustainability goals) and practice and what will this mean for the dairy industry and more specific, for DSF and in turn for FrieslandCampina?

1.4 Knowledge gap

Research has been conducted on PMS's regarding sustainability strategies and criteria that a sufficient measurement system should include (e.g. Baumgartner & Ebner, 2010; Cagnin et al., 2005; Mayer, 2008; McCormack, 2008). However, such a system has not yet been developed specifically for the dairy industry. Many reports and researches have been published about sustainable supply chain management in general. However, there are only a minimal number of publications about the actual impacts of PMS's on specific supply chains and/or production chains (Rietbergen & Blok, 2013).

To this extent, there is a clear knowledge gap in the actual design and impacts of PMS's and, more specifically, for a certain system in the dairy industry. Therefore, this study provides a good opportunity to develop and design a process approach for continuous improvement towards corporate sustainability, which will be measured in levels of maturity and will be applied to the dairy industry. Lastly, this study will provide knowledge about the consequences the implementation would have for FrieslandCampina in particular.

1.5 Scientific relevance

First of all, because a PMS does not yet exist in the dairy industry, it is interesting to conduct research in this area and contribute to its development. Furthermore, the research is relevant to get more insights into the effects a PMS could have on a specific sector or organization.

Moreover, the research that is performed for this thesis is relevant, because it examines the process of designing a PMS for the dairy industry. By applying this to a specific company, it will help to investigate a business' ability to cope with a certain system and the strategies available for a company to improve itself and reach a higher level of maturity. So, this study will give insights in how a PMS could be implemented and how companies will react on this type of systems.

Lastly, and most importantly, the research contributes to the field of Sustainable Development, because it provides insights into how the food industry can become more sustainable by implementing corporate sustainability PMS's. In addition, it is relevant to identify if and in what ways PMS's are raising the bar in businesses' strategies towards sustainable production. The research will also contribute to the specific department of environmental governance. Due to the fact that the system will be assessed by the organizations themselves, it is a form of self-assessment and corporate governance (De Jong, De Jong, Mertens & Wasley, 2005). It is interesting to investigate how this form of governance will be applied to a specific case and how organizations react on this.

1.6 Societal relevance

The transition towards sustainable food production is important for both society and the environment. Sustainable food production is needed to maintain and preserve current and future food demand by allowing for the rapid population growth. Interest among farmers and companies towards a more sustainable production and cooperation, i.e. CS, has grown and societal concern about the environment and animal welfare is increasing. Besides this, sustainability of the food industry has also been high on the agenda of the government and agricultural organizations for years (Van Calker et al., 2005).

Because of societal concern, i.a. stated by the Food and Agriculture Organization of the United Nations (FAO, 2015), it is relevant to invest more into research regarding sustainability in the food industry. Furthermore, sustainability is a changing concept, which is time specific. Therefore, it is useful to design a model which measures long-term sustainable improvement instead of a specific moment in time, which is the case in most existing standardization systems and traditional PMS's.

A solid, clear measurement system for the dairy industry provides advantages for both society and businesses. From a societal perspective, a solid system offers increased transparency about the sustainability practices of dairy producers. For consumers the existence of a unified system provides clarification. This is an advantage in relation to the large number of existing standards and labels, which can cause confusion amongst consumers (Nadvi & Wältring, 2001). A clear PMS for the whole dairy industry will provide advantages for businesses. The use of one consistent system means that separate auditing on a farm scale is no longer needed and saves costs for the company (Watson & Emery, 2004).

1.7 Aim and research questions

The goal of this research is to make sustainability in the dairy industry measureable and to develop a PMS further to the DSF. This should provide information which allows a comparison of dairy producers regarding their continuous improvement in the area of sustainability.

As previously mentioned, there is a clear social and scientific demand for the implementation of an international PMS towards sustainability in the dairy industry. Therefore, the research questions that will be answered in this thesis are as follows:

- *What are the key characteristics of a Performance Measurement System for the continuous improvement of corporate sustainability?*
- *How can this information be used for the design of a corporate sustainability Performance Measurement System for the dairy industry?*

In order to answer these main questions, various sub-questions are developed:

1. *What are the requirements for a PMS?*
2. *What are the key characteristics of the different levels of maturity in a PMS from scientific literature?*
3. *What are the key characteristics of the different levels of maturity in a PMS from practical experience?*

Besides these questions, comparable cases will be studied and assessed. These examples will be used to gain more insights into the PMS practices and these could be applied to the dairy industry and more specifically to the DSF framework. In relation to this, the following question is:

4. *What can we learn from existing systems and how can we use this information to build a conceptual PMS for the dairy industry?*

1.8 Outline of the thesis

This thesis is structured as follows: in the next section the methods of the research will be explained. In section 3, the theoretical background underlying the research questions will be explained and the literature review will be given. The results of part one of the research will be presented as well as their implications. This section will go into more



detail about corporate sustainability, sustainability standards and more specifically about performance measurements systems. The leading articles in this study will also be briefly discussed. Furthermore, a first set of requirements for a PMS is built. In the subsequent section, section 4, the results of the second part of the research will be given. These results consist of the analysis of comparable cases as well as the second set of requirements for a PMS. A further synthesis of these two parts (part one and two) and the design of the PMS for dairy, will be discussed in section 5. In section 6 the further application to the dairy industry and the testing pilot will be described. This section is followed by the discussion (section 7) and conclusions (section 8). Thereafter, the thesis will close with section 9, in which recommendations are given to the internship organization, FrieslandCampina. Finally, at the back of the thesis, appendices can be found which provide further explanations about certain parts of the thesis, to be referred to when needed.

2. Methodology

In section 2 the methods of the research will be explained. In the following sub-sections, the research strategy, data collection, data analysis and the reliability and validity of the research are addressed.

2.1 Research strategy

In order to answer the research questions, which address the key characteristics of a performance measurement system and how this information can be used to design a corporate sustainability performance measurement system for the dairy industry, the following research strategy is used. These questions show that the research conducted in this thesis is a design process with a scientific background.

The general research strategy can be explained by following the theories described by Verschuren and Doorewaard (2010). First of all, a distinction is made between a theory-oriented and a practice-oriented research (Verschuren & Doorewaard, 2010). In this study a practice-oriented research is executed, because the study is emphasizing the more practical side of research and focuses on the context of the research, in other words the wider perspective, and this context is placed within an organization (in this case FrieslandCampina) in which the research is taken place. The research consists of three parts. In order to get this wider view of the research subject and understand the context, first desk research is executed.

Part 1

In the first part of the research, desk research and a literature review is performed in order to answer sub-questions 1 and 2 (respectively; *SQ1, what are the requirements for a PMS? And SQ2, what are the key characteristics of the different levels of maturity in a PMS from scientific literature?*) Within this part, existing reports will be used in order to analyze the existing information about sustainability standards and PMS's. In this part a first set of requirements for a good working PMS is made (Section 3.5.1) and characteristics of levels of maturity from scientific literature are derived (Section 3.5.2). According to Verschuren and Doorewaard (2010) a difference should be made between empirical- and desk research. Despite the authors presume a choice between these two, for this study the two types are combined. Thus, the literature study, or desk research, forms the first part of the thesis and is followed by a second part.

Part 2

This second part includes empirical research in terms of reviewing comparable cases, conducting interviews and observations during the internship period. How and with whom the interviews are conducted will be explained later in this thesis (Section 2.2). Information obtained from this second phase gave the opportunity to develop a second set of requirements and to be able to answer sub-question 3 (*SQ3, what are the key characteristics of the different levels of maturity in a PMS from practical experience?*) Dividing the research into separate parts gives the opportunity to execute feedback or reflection of the first phase after the second phase has been completed, in order to answer sub-question 4 (*SQ4, what can we learn from existing systems and how can we use this information to build a conceptual PMS for the dairy industry?*).

Part 3

After assessing the lessons learned from the information derived from parts one and two, a first design of the PMS for the dairy industry can be made and subsequently this first design should be tested and improved. This was executed in the third part of the research. The research framework (Figure 2) describes this strategy by showing the various steps and parts of the research towards answering the research questions.

For this thesis an in-depth, practice-oriented research was conducted by making use of empirical- as well as desk research, in which the first phase consists of desk research and the second phase emphasizes empirical research. Furthermore, this study has the purpose to design a corporate sustainability PMS for the dairy industry and this process relies on the results of part 1 and 2 and is executed in part 3. This is an applied research in which a scientific base is used to apply a design process in a practical example. Consequently, the overarching type of research is a 'scientifically supported design process' towards sustainable dairy production.

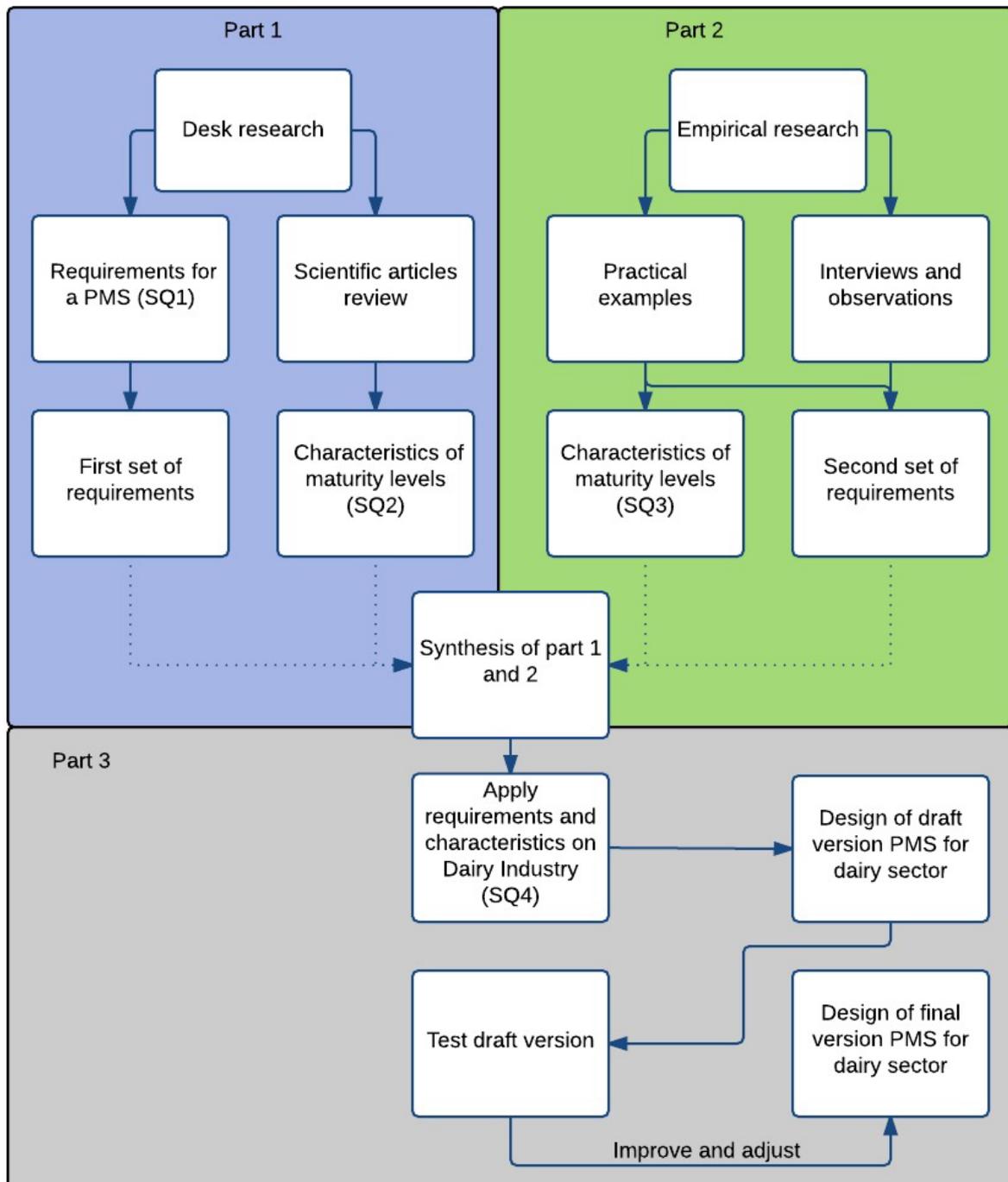


Figure 2. Research Framework

2.2 Data collection

In order to carry out the various parts of the research, different sets of data were collected. Data collection and information gathering is done through desk and/or literature research, reviewing of comparable cases, interviews and observations during the internship.

Part 1

In the first part of the study, desk research was carried out. Within this part of the research, scientific reports from others were used to analyze their findings (Verschuren & Doorewaard, 2010). Furthermore, in order to gather information, reports and scientific articles were collected from databases such as Google Scholar and Scopus. To find the right articles and reports, search terms and keywords like 'CSR', 'Sustainability Maturity', 'Sustainability Maturity Levels', 'Measuring Sustainability' were used and resulted in more than 500 articles. The articles and reports were selected by the year of publication, relevance of the subject and scope of the subject. Year of publication is important because articles that are written a long time ago could be outdated and do not provide relevant information anymore. With the scope of the subject is meant that some articles describe very general sustainable systems or even PMS in general without the focus on sustainability, whereas other articles are very specific about one case or one sector. It was useful to analyze a combination of both types, in order to be able to build a general framework. This analysis led in turn to a selection of approximately 20 articles which provided general background information about CSR, CS, performance measurement systems and maturity levels. This general information is used to build a research framework, which consists of requirements for a PMS. Seven articles from this selection were eventually used to base the maturity levels. These seven leading articles are briefly described in section 3, followed by the characteristics of the maturity levels.

Part 2

In the second part of this study, data was collected through a total of 26 semi-structured, in-depth interviews which were conducted in the period from January until March. These interviews were conducted both with representatives of FrieslandCampina (referred to as internal interviews) (10 interviews) and other relevant external organizations and sustainability experts (referred to as external interviews) (16 interviews). A list of all interviews can be found at the back of this thesis (Appendix A). These interviews were semi-structured to get detailed information from specific stakeholders and experts (Table 1) and lasted approximately one hour. The general interview structure and introducing presentation are enclosed in appendix C and D. The interviews involved specific questions as well as a critical review of the features per level from literature. This was to determine whether or not the information derived from literature is true or applicable in practical cases and business perspectives. However, the general interview structure functioned more as a guideline instead of a fixed set of questions, as the interviewees all had different backgrounds. Most of the interviews were face-to-face. However, when it was not possible to arrange an interview in person, the questions were discussed via email or phone. This was the case for only one interview. Interviews were conducted with general sustainability experts, as well as experts on specific themes. This was necessary because the DSF implements some specific themes where extra information was needed. As well as experts, organizations from comparable cases were also interviewed to get to know their specific systems better.

Table 1. Types of interviews

<i>Perspective</i>	
Internal	Internal experts
External	Sustainability experts Standardization and PMS experts Comparable cases

Additionally, information was collected from observations during the internship. These observations were from both formal and more informal meetings. The formal observations are shown and briefly explained in appendix B. Informal meetings are not included in the table and are not referred to in the thesis because they consisted of coffee break chats and short conversations. A number of formal observations are closely linked to interviews, however a clear interview structure was missing in these conversations and therefore they are interpreted as observations. Furthermore, formal observations consist of meetings with advisory and steering boards related to DSF and meetings with competitors or other important stakeholders in the dairy supply chain. Besides, meetings and consultations or agreements with sustainability colleagues during the internship are part of the formal observations. All observations led to background knowledge on the business activities at FrieslandCampina and their production process, which in turn helped to apply a maturity level system to the DSF demonstration scheme.

Part 3

In part three of the research the information from parts one and two was combined to come up with a design of the PMS for the dairy industry. The more applied results on this specific industry are described in section 5 and 6.

2.3 Data analysis

In the first part of the study, literature was analyzed and information was selected by means of relevance, confirmation by multiple reports and year of publication (Searcy, 2012). This information was selected for the first set of requirements.

The collected data from part 2 of the study is analyzed by making use of open coding. Interviews were transcribed in Microsoft Word and analyzed by making use of the analyzing program NVivo to build a clear overview of the information that was given by the interviewee. Codes were given to requirements derived from literature and new codes were added when new requirements and features were given by interviewees. By doing this, the overall transcript was broken down in smaller parts and the data has become more 'organized' (Welsh, 2002). Coding the interviews in this way helped to analyze the data in a structured manner (Yin, 2003) and provided information about which aspects mentioned by whom and how much the aspects were mentioned. The information obtained from practical examples and comparable cases was also done by conducting interviews and analyzing reports.

Thus, in order to answer the research questions, the analysis consists of various parts. First, general requirements were obtained from research reports to build the scientific background. Second, expert knowledge, practical examples and comparable cases were analyzed to learn from existing systems. This was done by means of conducting interviews and analyzing reports. Lastly, the knowledge obtained from this analysis was combined with the knowledge of the first part and was used to apply a PMS to the dairy industry.

2.4 Reliability and validity of the research

In order to guarantee reliability and validity of the research, certain measures were taken. The study is based on one specific organization and is therefore qualitative rather than quantitative, which could result in lowered reliability and validity (Golafshani, 2003). To prevent this and secure the quality of the research, various literature- and data sources were used and a large number of interviews were conducted (26 interviews in total). Moreover, interviews were held both with internal and external parties. In this way, the data sourcing is broad and contains a variety of different views and not relying on one perspective only. Furthermore, findings and observations from the internship were embedded in the companies' background, which gave it a more specific perspective and made it easier to apply the research and the final system to the company and the dairy industry. To make the research applicable to the whole dairy industry and not only to FrieslandCampina, comparable cases and practical examples from external parties were analyzed and comparisons were made during the whole process. In turn, this has helped to understand the existing systems better and learn more from them.

3. Results part one: Theoretical background & Literature review

Section 3 will explain which systems are described in literature and which information from the literature is useful for designing a sustainability PMS for the dairy industry. Criteria for PMS's will be obtained from existing scientific reports and articles (Part 1 of the research). However, before the criteria for a PMS can be described, more specific information about CS, sustainability standards, PMS's and levels of maturity will be given in the following sub-sections.

3.1 Corporate Sustainability

In this sub-section, the definition of corporate sustainability (CS) will be explained and clarified further, following from the introduction. Subsequently, the relevance of CS in the case of this research and how it can be applied to the PMS will be addressed. Once again, there are various definitions of CS (Van Marrewijk & Werre, 2003; Searcy, 2012), and much research has been conducted on this theme. Van Marrewijk & Werre (2003) define corporate sustainability as the organization's activities demonstrating the involvement of social and environmental concerns in business practices and in engagement with stakeholders. Another definition could be '*meeting the needs of the firm's direct and indirect stakeholders (such as shareholders, employees, clients, pressure groups, communities, etc.), without compromising its ability to meet future stakeholder needs as well*' (Searcy, 2012, pp. 239). However, these are the more broad definitions and it is rather obsolete to state a 'one solution fits all' definition (Van Marrewijk & Werre, 2003).

These researchers stated that each organization should choose its own relevant themes and ambitions regarding CS and these should fit to the organization's strategy and management plan (Van Marrewijk & Werre, 2003). Thus, taking into account differences amongst organizations and applying an appropriate definition of CS per organization will provide a clearer path towards sustainability and the organization's specific values. There is not one specific theme that has more importance than other themes, but it is important to include a variety of themes in the organization's strategy. In relation to this, the organizational structure will become more complex when the organization's CS is more ambitious (Van Marrewijk & Werre, 2003).

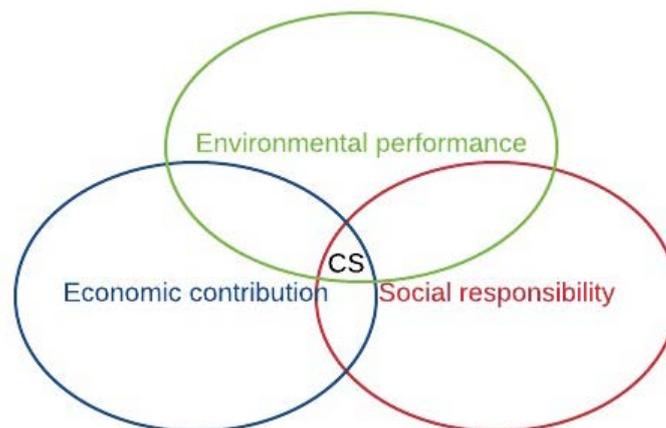


Figure 3. Corporate sustainability, the triple bottom line (Azapagic, 2003; Elkington, 2004)

Linking up to this, Azapagic (2003) describes CS as a combination of social responsibility, environmental performance and economic contribution. Additionally, this can be linked to the three P's (people, planet, profit). These elements, i.e. the triple bottom line, are strongly interrelated and in the interface of these three is CS (Figure 3) (Azapagic, 2003; Baumgartner & Ebner, 2010; Elkington, 2004). More and more industries are working towards increasing sustainability (Searcy, 2012) and to respond

to this trend and to participate in the challenge, continuous improvement should take place in this triple bottom line (Azapagic, 2003).

As stated before, many organizations are currently actively engaged in the sustainability debate (Azapagic, 2003; Searcy, 2012). By means of this they try to discover ways in which they can develop their triple bottom line and contribute to an increased sustainability. However, referring back again to Van Marrewijk & Werre (2003), the main challenge is to convert the general principles of sustainable development into actual and appropriate business principles and mainstream activities. Addressing this challenge asks for a system approach in which sustainability is integrated into the organization's business principles and is not regarded as an additional measure or task (Azapagic, 2003; Searcy, 2012; Van Marrewijk & Werre, 2003). However, many organizations face this challenge because CS is seen as a very complex issue and is considered as uncertain and ambiguous, and is organization specific. Furthermore, organizations experience difficulties in implementing CS, because they are not sufficiently educated in these issues and have a lack of authority (Searcy, 2012; Van Marrewijk & Werre, 2003).

Because of these challenges, to reach a higher level of sustainability and gain insights into the required business activities, a robust sustainability management framework is required. As indicated by Azapagic (2003), this framework should facilitate understanding of the key sustainability topics and actions needed to address them, ensure continuous improvement by measuring performance and evaluating progress, and involve good communication to relevant stakeholders about sustainability policies and development.

Stakeholder engagement is considered to be an important part of CS and stakeholder theory is a widely applied theoretical framework for research on CS (Searcy, 2012). According to Searcy (2012) *'stakeholder theory implies that corporations have obligations to individuals and groups both inside and outside the corporation, including shareholders, employees, customers, and the wider community'* (pp. 239). In turn, an increase in stakeholder engagement is an important indicator to get higher on the maturity ladder (which will be explained later on in this thesis).

3.2 Sustainability standards

Due to the increase in sustainability awareness, globalization of trade patterns and the increased acknowledgement of the governmental limitations in collaboration with the global supply chain, voluntary sustainability standards have arisen. A voluntary sustainability standard should start from the basic principle that all actors in a sector are able to aim for compliance with a given set of criteria related to the given standard or index. Because of this, the unique aspect of sustainability standards is that they are generally applicable for all types of actors in the market (Busch, 2011; Potts, Lynch, Wilkings, Huppe, Cunningham & Voora, 2014). According to ISEAL Alliance (2014), the definition of a sustainability standard in general is *'a standard that addresses the social, environmental or economic practices of a defined entity, or a combination of these'* (pp.7). In addition, sustainability standards are useful and often used in the private sector to support a more sustainable and green economy. Furthermore, voluntary sustainability standards play an important role in supporting investments in sustainable practices. Therefore, voluntary standards can be seen as tools to promote brand development as well as risk management (Potts et al., 2014).

Moreover, in the last two decades the number of voluntary sustainability standards has increased significantly and this in turn results in increased transparency, consistency and economic efficiency to address the challenge of sustainable development (Busch, 2011; Nadvi & Wältring, 2001; Potts et al., 2014). A large amount of these standards are related to specific areas, however an increasing number of globally applied standards are aimed at changing the worldwide trade and production nowadays. The majority of the

standards emphasizes on the agriculture and forestry sectors, due to the fact that these sectors together form a large amount of greenhouse gas emissions (Potts et al., 2014).

Furthermore, the development of sustainability standards will be considered. According to Potts et al. (2014), the sustainability standards are established by making use of four principles of development and implementation; subsidiarity, conformity assessment, traceability and continual improvement. The first principle of subsidiarity is related to the fact that effective indicators or tools are the ones who are able to be customized to local issues and settings. By making use of this local applicability, voluntary standards will be approachable to the needs and interests of stakeholders in multiple regions (Potts et al., 2014). The second principle is related to conformity assessment, which emphasizes the ability to externally verify or certify the standard, which in turn leads to a more transparent and credible market claim (Potts et al., 2014). Thirdly, the standard should consist of a traceability system which helps to safeguard the reliability of claims made by the market by providing responsibility between the produced products covered by the standard and the eventually sold products (Potts et al., 2014). Finally, the standard should have the possibility to improve continuously, due to the young and dynamic characteristics of sustainable markets and associated standards (Potts et al., 2014).

So, in order to develop a high value voluntary sustainability standard, it should be applicable at specific regional settings and it should have the ability to be externally verified. The product chains covered by the standards needs to be traceable and the standard should be continuously improved over time (Potts et al., 2014).

Derived from ISEAL Alliance (2014), there are ten credibility principles to build a proper standard. These principles are:

- Sustainability
- Improvement
- Relevance
- Rigor
- Engagement
- Impartiality
- Transparency
- Accessibility
- Truthfulness
- Efficiency

A few of these principles correspond to the four principles of Potts et al. (2014); these are (continuous) improvement and traceability or transparency. The other principles all together strengthen the credibility or reliability of the sustainability standard. Furthermore, ISEAL Alliance (2014) comes with additional requirements, for example; consistency with other standards. With this is meant that, when new standards will be developed, they should be consistent with existing standards, but in the same time should not be a duplication of other standards.

After assessing the relevance of sustainability indicators and standards in general, the more specific type of these systems, the corporate sustainability Performance Measurement System, will be clarified in the following sub-section.

3.3 The relevance of a Performance Measurement System

According to what was said in sections 3.1 and 3.2, to give CS some more 'body', a certain framework is required and sustainability standards are in place. This framework or standard could have the shape of a PMS with maturity levels. Due to the fact that CS only does not provide specific and/or measurable tools to improve sustainability, a certain system is needed to make CS more concrete and to provide organizations with insights in which steps should be taken in order to increase their sustainability performance. Referring to Searcy (2012), a key element of any CS initiative will be the development of a CS performance measurement system (PMS). The definition of a PMS will be given as; *'a successful performance measurement system is a set of performance measures that provides a company with useful information that helps manage, control, plan, and perform the activities undertaken by the company.'* (Searcy, 2012, pp. 240).

Moreover, CS PMS's lack certain governmental requirements and therefore these systems can be regarded as a form of self-regulation (Searcy, 2012).

An advantage of using a PMS is that it will measure performance over a long time span (Searcy, 2012). Searcy (2012) makes a distinction between a 'normal' PMS and a 'corporate sustainability PMS'. This distinction is made because he argued that a PMS only focuses on environmental and social aspects and a corporate sustainability PMS concerns of environmental, social and economic performance (triple bottom line). He gave the following definition of a corporate sustainability PMS; *'a system of indicators that provides a corporation with information needed to help in the short- and long-term management, controlling, planning, and performance of the economic, environmental, and social activities undertaken by the corporation'* (Searcy, 2012, pp. 240). A good working corporate sustainability PMS is helpful for decision makers and organizations in steering the challenges of corporate sustainability by providing them with better understanding of their current situation and their desired future position and/or final goal (Kaplan & Norton, 2000; Searcy, 2012; Stichting Milieukeur, 2016). In relation to this, a PMS can thus assist organizations in monitoring their implementation of business strategies, which in turn will contribute to their organizational success (Kaplan & Norton, 2000). Referring to Lockamy & McCormack (2004), the growth in organization's process maturity leads in turn to an increase in process capability and consistency through the entire organization. Thus, higher levels of process performance mean higher levels of maturity.

Continuing on performance, two key dimensions could be distinguished. According to Neely et al. (1995) these dimensions are considered in terms of efficiency and effectiveness. Neely et al. explain on one hand, efficiency as the extent to which customer requirements are met. On the other hand, effectiveness measures how the organization's resources are utilized in economic sense, when providing a given level of customer fulfilment (Neely et al., 1995). Consequently, stated by Neely et al. (1995), a PMS can be defined as *'the set of metrics used to quantify both the efficiency and effectiveness of actions'* (pp. 2).

Subsequently, it will be more and more important for organizations to become best in their sustainability performances, because of the increased competition due to globalization and other market factors (Ashurst & Doherty, 2003). Additionally, an increased consumer interest for sustainable products, asks for a clear independent system to prove sustainability of the products (Stichting Milieukeur, 2016). As a result of this, the implementation of PMS's becomes even more important nowadays and in the future. PMS's are seen as essential tools to improve and give evidence of the organization's sustainability performance (Sharma, Bhagwat & Dangayach, 2005). On one hand a PMS provides clarification for producers to get insights in whether or not they comply with claims of the customers and which aspects should be improved to obtain the customers claim. On the other hand, organizations that produce in a very sustainable way and are frontrunners in the industry (high level of maturity) can get rewards, because they are 'best in class' (Stichting Milieukeur, 2016).

Described by Behn (2003), measuring performance is relevant and useful for organizations. There are a variety of reasons why performance measurement is useful. However, not all mentioned reasons were relevant in the case of this research, because they were too broad or too specific, and therefore only the most relevant ones are described here. First of all, organizations could measure their performance by means of evaluating their practices. This evaluation is about making comparisons with a standard. A certain comparison is needed, due to the fact that without this it is impossible to determine whether the organization has a good or poor performance (Behn, 2003). The second reason why measuring performance is relevant is because it makes organizations able to control their output. Another very relevant reason in the case of DSF is that measuring performance motivates behavior, so employees are more motivated to

change, when they see actual results in their performance (by means of measuring). Showing results, helps to explain to society what measures are taken and how the organization performs (Behn, 2003).

Related to this, measuring performance can have a promotional advantage, in order to convince society that the organization is successful and efficient (Behn, 2003). Additionally, what was said before, a comparison with standards or with other organizations captures attention, in contrast to raw information or results, which does not. Learning can also be a positive result of measuring performance. Learning is a step further than evaluation only and benchmarking or comparisons with other organizations is seen as a learning strategy. A measurement system is considered to be a reflection of what organizations predict to see and how they supposed to react (Behn, 2003). Lastly, related to learning, performance measuring will result in improving the organization's strategies. It will improve the organization's functioning since performance measurement is not an outcome or end in itself, it must be used to make continuous improvements (Behn, 2003). Moreover, stated by Behn (2003) it is important to realize that there is not one measure that is appropriate for all circumstances. It depends on the case or issue, which measure is the most appropriate.

Furthermore, a voluntary sustainability standard is often used in the form of a self-assessment. A self-assessment questionnaire is seen as a type of formative and informative assessment and is in practice often expressed as the active monitoring and regulation of a variety of learning processes (Nicol & Macfarlane-Dick, 2006). These processes include the setting of and orientation towards goals, the strategies and management procedures used to achieve the goals, reactions on feedback and the final outcomes (Nicol & Macfarlane-Dick, 2006). Despite the fact that these authors describe this type of assessment with an example of learning at university, it could be applied to a wider learning process which could be seen in organizations. The fact that goals should be defined in an organization as well as the road towards achieving these goals is similar to learning processes at university. This is a good example to use in this case. Besides, according to Nicol and Macfarlane-Dick (2006) a self-assessment could be more effective than top-down regulations, because in this example students will have more opportunities for feedback from fellow students and in turn learn more from each other. When applying this to the dairy industry, a self-assessment will be effective, because organizations should be more motivated to work together and exchange knowledge. Participants of a self-assessment will learn more in contrast to top-down regulations. So, a self-assessment could be seen as an active process towards improvement (or in this article learning) (Nicol & Macfarlane-Dick, 2006).

Subsequent to the previous mentioned advantages, there are a number of issues which have positive effects but could also cause some negative effects for organizations. And these are important to emphasize as well. One of the possible results of a PMS with maturity levels could be that organizations will face a certain competitive pressure. This could be seen as a negative effect, however it will also encourage organizations to do things better and faster than their competitors (Gomes, Romao & Caldeira, 2013). Furthermore, organizations cannot depend on traditional strategies any longer and therefore it is crucial to change their strategies in more sustainable ones (Gomes et al., 2013). So, even if it could create a negative effect, sustainability will be something to compete on in the future, in contrast to lowering costs etc. PMS's and other tools are necessary to capture this new way of doing business (Gomes et al., 2013).

A second issue that can create positive as well as negative effects is the possible support of a 'race to the top'. The implementation of a PMS could bring about a race to the top, because consumers could set a certain target on which level they purchase the dairy products. Therefore, they will raise the bar for the producers. However, related to the first issue, this could again result in more competition, because consumers could possibly outbid one another. So, for example one customer sets the 'bar' at level 3, but

another customer could react on this and sets its 'bar' at level 4 etc. However, this will not mean that when the target is set on level 4, level 3 and lower are not relevant anymore. This will cause some difficulties and therefore it could be a disadvantage (External Interview 3). So, somewhere in the system something should be included and good leadership is needed in order to prevent this (External Interview 3; Maak, 2007).

After assessing the relevance of CS, sustainability standards and more specifically the relevance of PMS's, an overview of the main articles used for this study will be briefly explained.

3.4 Introduction of leading articles

Before the scientific literature synthesis can be performed, a short introduction of the leading articles helps to get a better understanding of the research area. For the characteristics of the levels derived from literature, several articles are used. In this introducing section, the main articles will be briefly explained. So, why are these articles used and what are the main differences and similarities between the articles? Thereafter, it is important to identify from which perspective the articles are written and what their main characteristics are.

For this literature review, seven articles are used as the base for the maturity levels. From these articles, the most corresponding characteristics are used to build the system. When only one article states a certain characteristic, it is not always been used in the final system. The table below (Table 2) shows which articles are used and provides information about their key elements.

Table 2. Brief overview of main articles for maturity levels

<i>Authors</i>	<i>Title of the article</i>	<i>Research field/ area of expertise</i>	<i>Key issues addressed in article</i>
<i>Baumgartner & Ebner (2010)</i>	Corporate sustainability strategies: sustainability profiles and maturity levels.	Emphasis on corporate sustainable development of an organization (three P's)	Distinguishes 4 levels of maturity, clear description of different levels
<i>Cagnin et al. (2005)</i>	Business sustainability maturity model.	Sustainable business perspective, but it has a management focus as well How can an organization adjust its strategy to become more sustainable?	Article fits very well to the situation researched in this thesis, so the characteristics described in this article are valued high Distinguishes 5 levels of maturity Main purpose of the system is continuous improvement of sustainability Based on Capability Maturity Model (CMM)
<i>Figge et al. (2002)</i>	The sustainability balanced scorecard—linking sustainability management to business strategy.	Focuses on business management in three P's Main emphasis on strategic business management Contains aspects of finance	Triple bottom line embedding is essential Focus on Balanced Scorecard (Appendix E) to measure the implementation of CS strategies

<i>Kirkwood et al. (2011)</i>	A maturity model for the strategic design of sustainable supply networks.	Focus on maturity levels which provide companies with information what their current state is and how they can improve in the future Review of previous scientific literature about maturity levels	Contains three dimensions (three P's) Based on CMM model and reviews a variety of existing systems Network capabilities are closely linked to network maturity Framework that is built in the article is tested by making use of a pilot Article focuses on networks/network design
<i>Lockamy & McCormack (2004)</i>	The development of a supply chain management process maturity model using the concepts of business process orientation.	Emphasis on management maturity	Clear description of 5 maturity levels Continuous improvement is important --> rather many small steps than revolutionary steps
<i>McCormack et al. (2008)</i>	Supply chain maturity and performance in Brazil.	Management perspective, describes systems that support improvement of performance in the supply chain Maturity levels in the management performance of the supply chain	The article is not explicitly focusing on sustainability, but on improvement of performance in general Article is based on SCOR model (=Supply Chain Operations Reference model, Appendix E) Distinguishes 5 levels of maturity
<i>Van Marrewijk & Werre (2003)</i>	Multiple levels of corporate sustainability.	Emphasis on corporate sustainability with an management point of view	Distinguishes 6 developmental stages, clear description of these different stages/levels These stages provide the organization a proper path towards CS (emphasis: not 'one solution fits all' principle)

However, the question is, why are these seven articles specifically chosen? First of all, the purpose was, to find articles which were written within the last years, to prevent outdated articles and reasoning that are not relevant anymore. Furthermore, these articles are used because most of them are written from a management perspective, which is relevant for this type of system. This is because its main focus is on the management and/or processes side of the organizations. This will be explained further on in this thesis.

All seven articles focus on the sustainable development of an organization in general, not on the dairy industry in particular and therefore with these articles it is possible to build a framework from a broad background and from a wider perspective. It is important to have an overview of the articles and their main reasoning before the actual synthesis is executed.

Most articles describe sustainability performance by means of effective implementation of their strategies and not the actual output itself. From these sources, it became apparent that it is difficult to measure the actual output of an organization in a performance measurement system, because output differs amongst regions and types of organizations (Kirkwood et al., 2011; Van Marrewijk & Werre, 2003; External Interview

5,8,9,11,12,13; Internal Interview 3,9). To prevent a system which is only applicable in one specific region or for one type of organization, a more general point of view, and thus a focus on the process performance, is chosen. However, this subject will come back later in this thesis.

Once again, all articles focus on the management side and sustainability strategies of an organization. Although they call it performance measurement systems, they focus on process rather than on actual output performance only. This is recognized in all articles. So, they describe which strategies organizations have toward sustainable performance, but quantitative Key Performance Indicators (KPI's) or similar are missing. From this, it can be concluded that it is difficult to measure the actual output. However, if the processes and strategies are well defined, the performance of sustainability outcomes will improve too (Baumgartner & Ebner, 2010; Cacioppe & Edwards, 2005; Lockamy & McCormack, 2004). Thus, for this research a management and strategy point of view is chosen to look at the sustainability performance of the organization.

The general information about CS, sustainability standards and the introduction of the main articles served as a theoretical background and basis for the further research in this thesis. The last step in this part of the study was to use this general information for the first set of requirements for the PMS, which is described below.

3.5 Towards an intermediate conclusion

In this first part of the research, the assessed definitions and background of CS, sustainability standards, PMS's and the introduction of the main articles, helped to set up a first set of requirements for a sustainability PMS.

First of all, before determining the characteristics per level and defining the different levels of maturity, a general overview is useful to get some insights in the design of a PMS. So, the first set of requirements (section 3.5.1) is based on a variety of reports and scientific articles. Followed on this, the characteristics of maturity levels were determined and were based on the leading articles. This is discussed in sub-section 3.5.2.

3.5.1 First set of requirements and conditions for a good working PMS

The requirements and conditions described below and showed in table 3 are relevant to take into account when building a well-defined, good working, reliable system. These requirements were derived from various articles and reports and these form the first set of requirements. Thus, from a scientific literature study and embedding of this literature, information is gathered to build a first set of requirements for a PMS. After this first set, comparable cases were assessed and interviews were conducted to build a second set of requirements. This is executed in design phase two (Section 4).

From a variety of articles it became clear that a PMS mainly consists of 4 or 5 levels of maturity. This is useful to take into consideration, before the actual design of the PMS for sustainable dairy begins. Another issue of the design that is important is that when organizational processes improve, the management of that organization changes from mainly internal towards a more external focus (Lockamy & McCormack, 2004). Derived from Lockamy and McCormack (2004), maturity levels provide information about the process capability of the organization. In other words, the higher the maturity level, the more capable the processes are in the organization. When building the levels, it is crucial to include the attributes of the previous level in each subsequent level. By doing this, maturity levels or certain steps cannot be skipped (Cagnin et al., 2005).

Obtained from Vermeulen & Witjes (2016) a company or sector should change continuously and learn in a transformative way to get success and improvements towards sustainability. An important requirement for a PMS is therefore, that it should measure changes in a long-term perspective. Furthermore, indicators should measure

the triple bottom line and should be calculated periodically, in order to point out whether the production is becoming more sustainable over time and is related to CS (Mayer, 2008). Furthermore, derived from Munda (2005) and Folan and Browne (2005) criteria should be measurable (in terms of money, emissions etc.) and the weighting or relative importance of each criterion is important to include in the process towards designing a sustainability PMS. Furthermore, Folan and Browne (2005) state that PMS's should contain possibilities for feedback. Derived from the mentioned researchers, the following first set of criteria is determined and shown in table 3.

Table 3. First set of criteria for a Performance Measurement System

First set of criteria	Source
System consists of 4 or 5 levels of maturity	Baumgartner & Ebner (2010) Cagnin et al (2005) Lockamy & McCormack (2004) Rietbergen & Blok (2013)
As processes improve, management moves from internally-focused to externally-focused	Lockamy & McCormack (2004)
Achieving a higher level of maturity means a higher level of process capability for the business	Lockamy & McCormack (2004)
Each advanced level includes the attributes of the previous maturity level	Cagnin et al (2005)
Each level includes characteristics regarding predictability, capability, control, effectiveness and efficiency	Lockamy & McCormack (2004)
A PMS should measure long-term changes	Mayer (2008) Folan & Browne (2005) Vermeulen & Witjes (2016)
Indicators should address all areas of corporate sustainability; social, economic and environmental	Mayer (2008) Cagnin et al. (2005) Vermeulen & Witjes (2016)
The indicators should be calculated periodically	Mayer (2008)
The criteria and/or indicators should be measurable (in terms of money, energy, emissions etc.)	Munda (2005) Folan & Browne (2005)
The weighting or relative importance of each criterion/indicator have to be included	Munda (2005)
A PMS should contain possibilities for feedback	Folan & Browne (2005)
A PMS should be supportive to and consistent with the business's goals, objectives and programs	Folan & Browne (2005)

3.5.2 Characteristics of maturity levels from literature

From the already introduced, leading articles, characteristics per level were defined. To get these characteristics ordered in a clear way, different sub-aspects were defined and the entire overview of these aspects could be found in appendix F. The main steps in passing the levels are described in this section and illustrated in tables. First of all, a general description per level will be given, followed by a general reasoning of maturing.

From these leading articles, 5 levels of maturity are divided. The reason why is chosen for 5 levels instead of 4 is i.a. that the first three levels should be more easily achievable, and towards level 4 and 5 bigger improvements need to be made and in these two highest levels a differentiation will be made between the good and the slightly less good (External Interview 12). By applying these 5 levels it is easier to make certain smaller steps towards the top, or in other words 'the dot on the horizon'. Each level has its own characteristics and requirements. In table 4 the various levels derived from the leading articles are presented in order to get a clear overview. After this overview, the features of the levels of maturity were determined. The choices for the 5 levels as well as their characteristics are shown subsequent to the table.

Table 4. Levels of maturity of main articles

Author	Level allocation
Baumgartner & Ebner (2010)	Level 1 Beginning Level 2 Elementary Level 3 Satisfying Level 4 Sophisticated/Outstanding
Cagnin et al. (2005)	Level 1 Ad hoc Level 2 Planned in isolation Level 3 Managed with no integration Level 4 Excellence at corporate level Level 5 High performance sustainability net
Kirkwood et al. (2011)	Level 1 Accidental/Initial Level 2 Repeatable Level 3 Defined Level 4 Managed Level 5 Mastered/Optimized
Lockamy & McCormack (2004) and McCormack et al. (2008)	Level 1 Ad hoc Level 2 Defined Level 3 Linked Level 4 Integrated Level 5 Extended
Van Marrewijk & Werre (2003)	Level 1 Red Level 2 Blue Level 3 Orange Level 4 Green Level 5 Yellow Level 6 Turquoise

For the first level, we use the term *initial* (Cagnin et al., 2005; Kirkwood et al., 2011), because this emphasizes the beginning stage of maturity and this level is defined as 'sustainability is poorly defined and practices are unstructured' (McCormack et al., 2008; Lockamy & McCormack, 2004). The entrance requirement is that the organization at least should have an overview of the local laws and regulations (Baumgartner & Ebner, 2010; Van Marrewijk & Werre, 2003).

Level 1 Initial

Entry level, overview of local laws and regulations
There is no aim for sustainability and if present, sustainability is poorly defined and practices are unstructured
Communication, reporting and stakeholder engagement is low
Approach is informal

The second level of maturity is called the *recognized* phase, in this level organizations do not have defined their actual plans yet, but have recognized the need for a more sustainable approach and had a certain awakening (Kirkwood et al., 2011). The term 'recognized' is not found in the articles. However, the descriptions given in the articles fit well with the definition of recognized and the explanation of this level is 'sustainability is recognized and/or identified' (Cagnin et al., 2005).

Level 2 Recognized

Sustainability is recognized and/or identified
In compliance with minimum requirements
Sustainability management plans starts to be developed, not in alignment with main targets
Stakeholder engagement and collaboration only with main stakeholders
Basic processes are documented
Communication and reporting is still mainly internal

Thereafter, the *defined* phase takes place (Kirkwood et al., 2011; Lockamy & McCormack, 2004; McCormack et al., 2008), which is the third level in this study. Although, the 'defined' stage is the second level in the articles, for this study is chosen to apply it at the third level, because the recognized level is placed in-between. Within this level the organization has 'an increased interest towards continuous improvement and sustainability' (McCormack et al., 2008; Lockamy & McCormack, 2004). After these three phases, in which the focus was more on the management side of the organization, the fourth and fifth level are more focusing on actual performance and goals setting (Baumgartner & Ebner, 2010).

Level 3 Defined

Sustainability and continuous improvement become more important
Approach moving from mainly internal towards more formal and proactive
Sustainability starts to integrate into business procedures
Increased (active) cooperation with stakeholders
Increased internal and external communication and reporting

The fourth level is defined as *managed* (Cagnin et al., 2005; Kirkwood et al., 2011) and is characterized by 'sustainability as the major driver in the organization, processes are measured and controlled, however sustainability is still not always in total harmony with other goals' (Cagnin et al., 2005; Kirkwood, 2011). The term managed is used in many cases and clarifies the state in which the organizations are in this level; they have managed their sustainability processes.

Level 4 Managed

Sustainability is a major driver and embedded in the organizational management, but in conflict with other goals
Approach is proactive and formal
Internal support is strong
Cooperation along supply chain is high
Communication and reporting is consistent and integrated
Room for feedback
Processes are measured and controlled
Focus on goals setting and goals achievement

The highest level of maturity is called the *optimized* phase (Cagnin et al., 2005; Kirkwood et al., 2011), which is characterized by 'sustainability is the central issue in the organization's practice' (Cagnin et al., 2005; Van Tulder et al., 2014). The term *optimized* is chosen for this level, because it clearly states that this is the highest level and organizations have optimized, or perfected, their strategies.

Level 5 Optimized

Sustainability is central in organization's practices (integrated and optimized)
Approach is proactive and formal
Continuous improvement and learning
Cooperation and information sharing along supply chain and society is high
Communication and reporting is complete
Goals are achieved

Although literature describes 'strict' features per level which is described in the boxes above, it is important to understand or realize that a PMS does not prescribe a strict path towards CS. Due to the diversity of organizations, every organization follows different steps or stages of development (Van Tulder et al., 2014; External Interview 11,12). However, a PMS does offer a framework that provides clarification about how to become more sustainable and serves actually as kind of a mirror for reviewing an organization (Van Tulder et al., 2014). Thus, with this in mind, a general reasoning will be given and an overall overview of the features per level can be found in a table in the appendix (Appendix F).

First of all, one of the main trends that was recognized from literature is the movement from an internal approach towards an external approach (Table 5) (Cagnin et al., 2005; Kirkwood et al., 2011). Also strategies and management plans are mainly non-documented in the lower levels and extensively, formal documented in the highest levels of maturity (Cagnin et al., 2005, Kirkwood et al., 2011; Lockamy & McCormack, 2008). Thus, in the first two or three levels the organizations focuses on their own organization and internal activities (Baumgartner & Ebner, 2010; Cagnin et al., 2005; Lockamy & McCormack, 2004; Van Marrewijk & Werre, 2003) and in the higher levels the organizations are active in engaging other organizations or parties as well. In other words, the leading organizations take a more active role in stakeholder engagement (Table 7) (Cagnin et al., 2005; Kirkwood et al., 2011; Maak, 2007; McCormack et al., 2008; Lockamy & McCormack, 2004). Good leadership can be seen as '*building and cultivating trustful sustainable relationships with stakeholders inside and outside the organization to achieve mutually shared objectives based on a vision of business as a force of good for the many, and not just a few*' (Maak, 2007, pp. 331).

Another important path is the transition from an inactive towards a proactive approach. According to Van Tulder et al. (2014) the difference is made between inactive, reactive, active and proactive. In an inactive organization, societal problems and sustainability are seen as irrelevant. Inactive organizations assume that the government or another

authority makes the right choices for them and they follow the rules set by these authorities. The reactive organization has interest in sustainability further than only rules and regulations and takes part in stakeholder dialogues. However, they act not in own initiative, but react on other organizations. The focus is on operational management and costs are more important than sustainability (Van Tulder et al., 2014).

An active organization is an organization which takes its responsibility for societal problems and is aware of the fact that legal requirements are not always sufficient regarding environmental or societal issues. These organizations are motivated to improve their sustainability practices. Another step further is the proactive organization, in which sustainability is completely integrated into business strategies. These organizations are aware of major trends in global society and take into account indirect societal developments as well (Van Tulder et al., 2014). A proactive organization is ahead of future laws and regulations (Internal Interview 7,8). Related to this proactive approach, the organization's important issues, i.e. the organization's materiality assessment, are also defined in compliance with societal needs or concerns and is not only based on a direct stakeholder analysis (Van Tulder et al., 2014). Additionally, according to Porter and Kramer (2011) it is the responsibility of organizations to bring business and society back together. Still in many organizations social responsibility is a minor issue and not embedded in the core business. This is the same as Van Tulder et al. (2014) refer to, when organizations really want to be sustainable, they should engage society.

Table 5. Levels of maturity in the category Approach & Vision

Category	Level 1 <i>Initial</i>	Level 2 <i>Recognized</i>	Level 3 <i>Defined</i>	Level 4 <i>Managed</i>	Level 5 <i>Optimized</i>
Approach & Vision	Reactive approach ^{1,2} Informal ² There is basically no aim for sustainability or CS ⁶	Reactive approach ² Mainly informal ² <i>Sustainability issues start to emerge</i> ⁶	Reactive approach, moving towards proactive ¹ <i>From informal towards formal</i> ²	Business functioning is mainly proactive ^{1,2,4} <i>Formal</i> ^{1,2,4} <i>Long-term commitments</i> ^{1,2,4}	<i>Proactive approach</i> ^{1,2,4,6} Formal ^{1,2,4} Company's strategy has long-term focus and commitment ^{1,2,4} <i>A shared vision for sustainability is defined in agreement among employees</i> ¹

1) Cagnin et al., 2005; 2) Kirkwood et al., 2011; 3) McCormack et al., 2008; 4) Lockamy & McCormack, 2004; 5) Baumgartner & Ebner, 2010; 6) Van Marrewijk & Werre, 2003; 7) Figge, 2002

The approach or focus of the organization is related to cooperation and engagement of stakeholders and parties in the supply chain as well as communication and reporting (Table 6) (Baumgartner & Ebner, 2010; Cagnin et al., 2005; Kirkwood et al., 2011; Lockamy & McCormack, 2008; McCormack et al., 2008, Van Marrewijk & Werre, 2003). Internal and external communication and stakeholder engagement are very important topics in the process of CS, because they lead to better embedding and support of sustainability within the organization and with its stakeholders (Cramer, 2005). In the highest level, even engagement with society should be considered, because this is an important aspect of CS. Organizations at a high sustainability level should consider the needs and concerns of society and social issues as well (Van Tulder et al., 2014).

Table 6. Levels of maturity in the categories Cooperation & Engagement and Communication

Category	Level 1 <i>Initial</i>	Level 2 <i>Recognized</i>	Level 3 <i>Defined</i>	Level 4 <i>Managed</i>	Level 5 <i>Optimized</i>
Cooperation & Engagement	Engagement and cooperation is low, the organization is not an active partner in networks ^{4,5}	<i>Collaboration starts to be felt</i> ^{1,4}	<i>Increased and active cooperation between organizations and main 'actors' in supply chain and intra organizational</i> ^{1,3,4,5}	<i>Cooperation along supply chain is in place</i> ^{1,4}	<i>Collaboration and information sharing (multi-organizational/cross-enterprise) is high</i> ^{1,2,3,4}
	If present, cooperation is hierarchical (competition amongst partners) ¹	Collaboration and stakeholder engagement is mainly internal ^{1,4}	<i>Structured collaboration with main partners starts to be noticeable</i> ¹	Cooperation between organization and stakeholders/society ⁵	
	Informal contact with stakeholders ⁵		Focus on own organization and other actors in supply chain ^{1,5}	<i>Active participation in most dialogues</i> ¹	<i>Organizations actively engage with stakeholders</i> ^{1,2,3,4,6}
	Focus on own organization only ^{4,5,6}	Contact with stakeholders become more systematic ¹		Cooperation with main stakeholders regarding all environmental, social and economic issues ¹	
Communication	Minimal communication processes and channels in place ¹	<i>Communication of sustainability is mainly internal</i> ^{1,5}	Internal as well as external communication is present and <i>external communication (with stakeholders) starts to increase</i> ⁵	From rather internal towards external and public ^{1,2}	Communication channels are fully present (internal and external) ¹
	Communication about sustainability is exceptional ¹	External communication only with most important partners and not consistent over time ⁵		<i>Communication regarding sustainability is consistent and integrated</i> ¹	<i>Learning and knowledge flows are part of communication</i> ¹
	Internal ¹	Internal ¹		Communication processes and channels are based on feedback loops <i>Shift towards external focus</i> ^{1,2}	<i>Communication is complete</i> ¹ External focus ¹

1) Cagnin et al., 2005; 2) Kirkwood et al., 2011; 3) McCormack et al., 2008; 4) Lockamy & McCormack, 2004; 5) Baumgartner & Ebner, 2010; 6) Van Marrewijk & Werre, 2003; 7) Figge, 2002

Another important indicator for measuring sustainability maturity is the governance structure of an organization. The structure of an organization is crucial for the effectiveness of organizational management. Clear structured organizations will be more effective in their management and provide leadership towards their stakeholders. Therefore, structured organizations will be more effective in taking decisions (Harris, 2000).

In the higher levels, approximately from level 3, the actual results, measurements and goals setting become more important. In the highest level, goals should have been achieved or even have been outreached (Baumgartner & Ebner, 2010; Van Marrewijk & Werre, 2003). In turn, this will mean that in the lower levels the emphasis is on the process and in the higher levels outcomes and results get more attention. In relation to this, long-term versus short-term predictions, outcomes or goals are crucial in the maturity steps (Table 7) (Cagnin et al., 2005; Kirkwood et al., 2011; Lockamy & McCormack, 2004; Van Marrewijk & Werre, 2003). Referring to Knoepfel (2001), organizations that are orientating on sustainability over a long time span provide more predictable outcomes, which results in less negative surprises. Because of that, customers will choose for leading organizations in sustainability, not only because of their outstanding performance, but of their trust and predictability (Knoepfel, 2001). The increased trust and predictability of performance and goal achievement is also confirmed by Cagnin et al. (2005) and Kirkwood et al. (2011).

Table 7. Levels of maturity in the categories Data Collection & Measuring and Goals achievement & Commitment

Category	<u>Level 1</u> <i>Initial</i>	<u>Level 2</u> <i>Recognized</i>	<u>Level 3</u> <i>Defined</i>	<u>Level 4</u> <i>Managed</i>	<u>Level 5</u> <i>Optimized</i>
Data collection & Measuring	Sustainability data is not constantly collected or analyzed ¹ Performance/process measurement is not in place ⁴	<i>Sustainability measurement is growing</i> ¹ <i>Basic processes are documented</i> ⁴ <i>Performance measurement for own organization</i> ¹	<i>Performance improvement measures take place</i> ^{3,4}	Sustainability goals are quantified and measurable ¹ Measurement of sustainability is clear and generally understood ¹	<i>Measurements and monitoring of process performance are common</i> ^{1,3,4} <i>Advanced process documentation</i> ^{1,3,4}
Goals achievement & Commitment	Highly unpredictable and uncertain ¹ Sustainability direction and goal setting are only partly known ^{2,3,4} Estimation of future performance is not in place ^{2,3,4} Short-term commitments ⁶ Compliance with law and regulations, mandatory requirements ^{6,5}	Process performance becomes more predictable ¹ Goals or/and targets are defined ¹ , but missing more often than not ^{1,4} Estimation of future performance is partly planned ^{2,3,4} Short-term commitments ⁶ Compliance with law and regulations and some <i>voluntary frameworks</i> as well ^{5,6}	Process performance is more predictable, but only in a qualitative way ^{1,2} <i>Estimations of future performance are in place</i> ¹ <i>Medium-term commitments</i> ⁶ Mandatory and voluntary frameworks towards CS governance are focused on ⁵	Process performance is more predictable and <i>targets will be achieved and will be measured in a quantitative way</i> ^{1,2} Estimations of future performance are in place ¹ <i>Long-term commitments</i> ^{1,2,4,6} Goals are quantified and measurable ¹	<i>Trust and reliability is high</i> ¹ <i>Proactive process improvement and goals are achieved</i> ^{5,6} Estimations of future performance are in place and plans to reduce future impacts are clear ¹ <i>Goal is to reach zero impact on the environment</i> ⁶ Long-term commitments ^{1,2,4,6}

1) Cagnin et al., 2005; 2) Kirkwood et al., 2011; 3) McCormack et al., 2008; 4) Lockamy & McCormack, 2004; 5) Baumgartner & Ebner, 2010; 6) Van Marrewijk & Werre, 2003; 7) Figge, 2002

When going a little bit deeper into monitoring and measuring and referring to Azapagic (2003), monitoring is an important determinant whether or not the organization is moving towards increased sustainability. By doing this, one can measure to what extent goals are met or progression is made. So, monitoring or measuring will be a clear indicator whether the organization continuously develops (see also Table 7). Besides, measuring and reporting can help in evaluating the organization's performance, which in turn provides a feedback loop for evaluation and learning (Vermeulen & Witjes, 2016).

Closely related to measuring is data collection, which is also a clear indicator of the maturity of the organization. In the beginning stages, data is not or only informally collected and collection of data will increase during the levels (Cagnin et al., 2005; Lockamy & McCormack, 2004). In the highest stages, advanced measurement and documentation take place and monitoring of data is common (Cagnin et al., 2005; Lockamy & McCormack, 2004; McCormack et al., 2008). Information about monitoring will provide insights in the internal and external communication of the organization, and this is mainly through sustainability reporting (Azapagic, 2003). Hereby, a differentiation is made between one-way communication (in the beginning levels) and two-way communication (at the highest maturity level) (Baumgartner & Ebner, 2010). With one-way communication is meant that information is given, however interaction does not take place or feedback will not be given (Lasswell, 1948). On the other hand, two-way communication is considered more like a dialogue, whereby information is shared and interaction between different actors take place (Lasswell, 1948).

In relation to this, communication is again a very important measurement of maturity (see again table 6). Communication is crucial to create internal as well as external support (Azapagic, 2003). In most literature, a trend could be recognized from internal communication towards external communication by growing in maturity (Cagnin et al., 2005; Kirkwood et al., 2011; Baumgartner & Ebner, 2010). Internal communication is the first step towards sustainable improvement, because it has a significant effect on the corporate culture of the organization (Azapagic, 2003). However, when growing in maturity, society and external stakeholders expect to get more information about the sustainability practices of the organization (Azapagic, 2003; Baumgartner & Ebner, 2010). Thereafter, external communication is mainly through sustainability reports (Table 8) (Azapagic, 2003; Baumgartner & Ebner, 2010; Vermeulen & Witjes, 2016). In turn, reporting is very important to create transparency for external stakeholders and shows which steps are made towards increased sustainability (Baumgartner & Ebner, 2010; Vermeulen & Witjes, 2016). In the first two levels, external reporting about sustainability is not considered or not very common. And from the third level, sustainability reporting becomes more common and more sustainability issues will be addressed (Baumgartner & Ebner, 2010).

Table 8. Levels of maturity in the category Reporting

Category	<u>Level 1</u> <i>Initial</i>	<u>Level 2</u> <i>Recognized</i>	<u>Level 3</u> <i>Defined</i>	<u>Level 4</u> <i>Managed</i>	<u>Level 5</u> <i>Optimized</i>
Reporting	No consideration of sustainability issues in any report or communication channel ⁵	<i>Only the most relevant sustainability issues are reported internally</i> ⁵	Sustainability issues are reported internally and the most important issues are also reported external ⁵	<i>All sustainability issues are presented in annual reports and through internal communication channels (one-way communication)</i> ⁵	All sustainability issues are presented in the annual reports and through internal communication channels (<i>two-way communication</i>) ⁵

1) Cagnin et al., 2005; 2) Kirkwood et al., 2011; 3) McCormack et al., 2008; 4) Lockamy & McCormack, 2004; 5) Baumgartner & Ebner, 2010; 6) Van Marrewijk & Werre, 2003; 7) Figge, 2002

Referring back to the stakeholder engagement, stakeholder dialogues are important and this is also a leading indicator for change (Kaptein & Van Tulder, 2003). In addition, the link between reporting and stakeholder dialogues is strong. So, with active stakeholder dialogues, environmental issues faced by the organization are better understood by its stakeholders and operating areas. This increases transparency, information sharing and the structure of reporting which in turn has a positive influence on the CS practices of the organization (Kaptein & Van Tulder, 2003). Furthermore, large multinational organizations have the power to positively contribute to society and therefore they should take the chance to do so (Maak, 2007).

In order to determine whether an organization shifts from one level to another, the existence of tipping point is crucial to understand. Tipping points are crucial highlights beyond which the entire system changes (Van Tulder et al., 2014). Tipping points also represent points at which a new balance is achieved, so in other words, when a new level is reached. These milestones are important, but it is not exactly known where these points are. However, they should determine when an organization steps from one level to another, in other words, to increase in their sustainability practices (Van Tulder et al., 2014).

Concluding, in this section a first set of requirements for a good working PMS, which should be taken into account when designing a PMS, has been described as well as the characteristics of maturity levels from scientific literature. The main requirements defined by the scientific articles are that the system should contain 4 or 5 levels of maturity. The levels should include the attributes of the previous (lower) maturity levels. The system should measure long-term development and the criteria should be measurable. For this research, 5 levels of maturity are developed and the features per level as well as the steps that should be taken are described.

After assessing the first set of criteria and characteristics of maturity levels, it is useful to compare scientific literature to actual cases and comparable systems. In the following section (part 2 of the research), the main comparable cases are described and after this a second set of requirements and characteristics of maturity levels will be defined.

4. Results part two: Comparable cases & Lessons from practice

In the following section, the results of part two of the research will be discussed. Within this part, the comparable cases used for this study are explained, a second set of requirements is built and the characteristics of maturity levels derived from empirical research are investigated.

4.1 Comparable cases

To become familiar with sustainability PMS's, comparable cases were analyzed, besides pure literature review and interviews. Most common used PMS's in the Netherlands are the CO₂ Performance ladder and the MVO Performance ladder (MVO is the Dutch translation for CSR). Besides Dutch systems, also international systems were analyzed, as the PMS for the dairy industry should be internationally applicable. To investigate how the systems are built and how they are applied, interviews were conducted and reports were analyzed. In the following sub-sections, the main comparable cases are assessed.

4.1.1 Introduction of the cases

In this sub-section, the five systems that were used for the analysis of comparable cases will be briefly discussed. Information will be provided about the key characteristics of the system, how and when they are developed or set up and how the systems are organized. For this analysis, interviews as well as reports of the systems are used. The five systems with their main characteristics are shown in table 9.

Table 9. Overview of the comparable cases used in this study

Comparable Case	Key Features	Scope	Levels	Underlying system	Year of commissioning
CO ₂ Performance ladder	To get insights in GHG reduction and energy efficiency in order to become more sustainable Independent system	the Netherlands	5	CMM	2009
ProRail Safety ladder	To create more awareness amongst employees to work safely and to support cooperation Focus on process performance Emphasis on internal cultural changes	the Netherlands	5	Hearts & Minds Model of Parker	2011
MVO Performance ladder	Measurement tool for CRS Focus on process performance	the Netherlands	5	-	2010
Dow Jones Sustainability Index	Emphasize the CS performance of an organization Global approach with regional indices	Global	No levels, results are displayed in score and organizations are benchmarked	-	1999
Carbon Disclose Project	To promote the share and management of environmental information amongst companies Independent system	Global	No levels, results are displayed in score and organizations are benchmarked	-	2000

CO₂ Performance ladder

The first system that was analyzed is the CO₂ Performance ladder (in Dutch CO₂ Prestatieladder). The CO₂ Performance ladder (CO₂PL) is examined by various researchers, in order to investigate its potential impact and to explain how the different levels are determined. In 2009 the CO₂PL was developed by ProRail to give insights in the greenhouse gas emissions and energy efficiency performance of companies in ProRail's supply chain (Rietbergen & Blok, 2013). However, the CO₂PL is now an independent system, owned by SKAO (Stichting Klimaatvriendelijk Aankopen & Ondernemen, 2015) and is not owned by ProRail anymore. Nowadays, the system has its own scheme owner and scheme administrators (External Interview 9).

In the CO₂PL, five levels of performance are constructed (Rietbergen & Blok, 2013; External Interview, 9). Among members, CO₂ emission awareness is increasing and their strategies are assessed and certified. The more conscious and consistent the members work towards CO₂ reduction, the higher they will climb the ladder (Dorée, Van der Wal & Boes, 2011).

ProRail Safety ladder

The Safety ladder is built a few years later, in 2011 and is still owned by ProRail. The set-up phase of the system is similar to the early stages of the CO₂PL (External Interview 15). The goal of the Safety ladder is to obtain more awareness in organizations of their possibilities to work in a safer way, and the goal is to do this together. So, cooperation along the supply chain and with all stakeholders, from the bottom of the supply chain up to the top, is a crucial part of the Safety ladder (External Interview, 15). Like the CO₂PL, the Safety ladder also originates from the construction industry (External Interview 15; ProRail, 2015). Although, the system is originated from one specific industry, it aims to be applicable for other industries as well, because it makes use of a generic terminology and it is not needed to use different versions for different organizations or industries (ProRail, 2015).

MVO Performance ladder

The MVO performance ladder (in Dutch MVO Prestatieladder) is a measurement tool regarding CSR and relies on the principles of the ISO 26000 standard⁸ (MVO, 2013¹). The MVO performance ladder (MVO PL) arose from the fact that there was a clear need for a measurement tool and a credible, independent certification on CSR (ISO 26000 Scan, n.d.). The requirements and certification norms of the MVO PL are founded by the Foundation Sustained Responsibility (FSR) and developed in 2010 (MVO, 2013¹). Like the CO₂PL, the MVO PL is again an independent system. The system distinguishes 5 levels of maturity in which the first two levels are entry or starting levels (MVO, 2013²; ISO 26000 Scan, n.d.). The certification body is responsible for the execution of the certification of the MVO PL and determines at what level the organization is on its CSR management performance (MVO, 2013²).

Dow Jones Sustainability Index

The Dow Jones Sustainability Index (DJSI) is launched in 1999 (Knoepfel, 2001) and is therefore the oldest sustainability standard used in this study. The index emphasizes the CS performance of an organization and a sub-division of the index in people, planet and profit is based on the principles of Elkington (1997). The index measures the CS performance of the world's leading organizations and serves as a benchmark for the integration of sustainable practices. Besides, it also provides insights in how to adopt sustainable practices in the organization (RobecoSAM, 2015¹). Within the index, organizations are ranked, like a benchmark, by their sustainability performances (RobecoSAM, 2015¹).

Carbon Disclosure Project

The Carbon Disclosure Project (CDP) is a non-profit organization that operates worldwide and was founded in 2000. The organization is set up to promote the share and management of environmental information amongst companies. It possesses one of the biggest international data collections for corporate climate change and greenhouse gas emissions, water and forest-risk (CDP, n.d.; Kim & Lyon, 2011). The organization has the goal to drive sustainable economies by reducing greenhouse gas emissions, enhance water resources and prevent deforestation by companies and cities. This goal will be achieved by promoting issues of climate change in the organizations strategies (CDP, n.d.). Consequently, the CDP provides an independent scoring system to benchmark

⁸ISO 26 000 provides guidance in social responsibility. Therefore, it helps businesses and organizations to change their principles into actual, effective actions (ISO, n.d.).

organizations performance on sustainability and environmental topics (CDP, n.d.; Kim & Lyon, 2011).

4.1.2 Scope of the cases

The scope of the system is useful to take into consideration in the assessment of the different cases. Three of the fore mentioned five systems are applicable for the Netherlands. However, they are still useful to study, because they provide information about how to apply a certain system to a specific area. The other two systems have an international scope. In this sub-section the scope of the five comparable cases are briefly discussed.

First of all, the CO₂PL is built on the Dutch situation and 'culture' and therefore it is more difficult to apply it on an international scale (External Interview, 9). Similarly, the Safety ladder developed by ProRail operates within the national boundaries of the Netherlands as well. The third system that functions at a national level is the MVO PL. Like the CO₂PL and the ProRail Safety ladder, the MVO PL is applied to the Dutch situation only. Thus, these three systems are built on the country specific characteristics of the Netherlands. However, the MVO PL is also applied in Belgium (Reed, 2012). Furthermore, all three systems have a 'chain approach' and look further than at company level only. Especially in the Safety ladder, the 'chain approach' is strong, they stated that *'improvements are done together, if you want to make further steps on the ladder, your chain partners have to come with you'* (External Interview 15).

In contrast to the previously discussed systems, the DJSI and the CDP are developed with an international perspective (CDP, 2015¹; Knoepfel, 2001). The DJSI is thus a global applied and also well-known, international system and is used for a variety of products and industries (Knoepfel, 2001). Furthermore, the DJSI is structured in one global index and a variety of regional and country specific applicable indexes. These regional indexes are derived from the global index and therefore they are sub-divisions of the world index (Knoepfel, 2001). The CDP has a regional application too, even though it is a global system, certain issues are addressed on a regional scale. An example of a regional issue is water management. The water issue strongly depends on regional or local differences and is therefore difficult to measure on a large, even global scale (CDP, 2015¹). These regional differences are important to take into account for the dairy industry as well. In addition, the DJSI has a 'company level approach' as well as a 'sector broad or chain approach'. This means that it scores the sustainability performance of the organization as well as a sector average score (SustainAbility, 2013). This index thus takes into account the individual performance as well as the performance of the entire sector. In contrast, the CDP primarily focuses on the own organization score. According to Kim and Lyon (2011), organizations which are transparent about their disclosure, could positively affect their shareholder value and thus higher share prices. They state that the use of the CDP, enables organizations to show their good environmental performance. Investors appreciate this and it is more likely that they invest in the organization (Kim & Lyon, 2011).

4.1.3 Reliability of the cases

For the reliability of the comparable cases, various criteria are considered. These criteria comprise the year in which the systems were commissioned and the amount of organizations that work with or participate in the case. By knowing the reliability of the case, the information derived from the case can be taken into consideration in terms of usefulness.

The year of commissioning is relevant because, it says something about how experienced the system is. When a system is in use for a long time, one can assume that mistakes made in the beginning as well as the limitations of the system have been

solved during the long development period. Derived from an interview (External Interview 15), when a system has just been in use, it is still in development. Thus, when a system is in place for a long time, there could be stated that it is more reliable to use as a case than systems that are built just a few years ago (5 years or less). Notwithstanding, a performance measurement system or benchmarking system is always in development, because it supports continuous improvement (Bourne et al., 2000). Therefore 'older' systems are still in development too.

Furthermore, when a comparable case has a lot of participants or members, it could be stated that it is a reliable and/or good working system, because, apparently, a large number of organizations trust the system to work with.

For the comparable cases used in this study, the above mentioned criteria were applied. First of all, the systems are all in use for more than a few years, which tells something about the reliability. The ProRail Safety ladder, however, is the system that is in use for the shortest period. One could thus assume its reliability is less, compared to the other systems out of the five cases. Nevertheless, the system is in place for 5 years now, so the most beginning ambiguities should have been solved through the years (External Interview 15).

When looking at the criteria related to the number of members of the system, the DJSI is by far the most used system. It has over 3300 invited organizations and at least 1800 questionnaires analyzed (RobecoSAM, 2014). Due to this, one could state, that it is a very reliable system to use for this study. Additionally, the other cases are also well-known systems and especially the international systems are widely used. So, one issue that makes the Dutch systems less reliable for international research is because they lack the global or international focus.

After determining the reliability of the cases, the methodology will be analyzed in the next sub section.

4.1.4 Methodology of the cases

A number of comparable cases are based on other already known systems. It is beneficial to get to know the underlying systems, because they could provide new insights in a broad variety of measurement systems. If the systems are not based on an existing system, the methodology of the case will be addressed.

First of all, the CO₂PL is mainly based on the Capability Maturity Model (CMM) (External Interview 9), which is a widely used system in measuring maturity of organizations and is introduced by the software engineering branche (Cagnin et al., 2005; Kirkwood et al., 2011; Rietbergen & Blok, 2013; Paulk et al., 1993). Within the CMM a distinction is made between 5 levels of maturity, which describe *'the capability of an organization to perform important processes to deliver a certain product or a process'* (Rietbergen & Blok, 2013, pp. 35). Thus, the CMM is a framework that describes the main components of an effective process and provides organizations with useful information about the performance of their processes (Paulk et al., 1993). In relation to the performance of processes, cultural changes in the organization are crucial elements towards higher levels of maturity as well (External Interview 15). The 5 levels of maturity are clearly described in many studies (Cagnin et al., 2005; Kirkwood et al., 2011). Although the system is derived from a specific sector, it is a widely used system in other sectors and applicable for a wide variety of cases (Paulk et al., 1993). The model is applicable for other sectors and processes, because of the general requirements and conditions that are established per level.

The second case, the ProRail Safety ladder, is mainly based on two other systems or frameworks. Firstly, the Safety ladder is based on Hearts & Minds of Shell. This

framework includes useful elements, however it is more like a self-measurement system and this is not the case for the ProRail Safety ladder (External Interview 15). Thus, Hearts & Minds is a useful tool in the sense that it provides organizations with techniques and tools to involve their employees in dealing with health and safety in their business practices. Additionally, the Hearts & Minds toolkit is relevant to both Shell as well as 'non-Shell' companies and is worldwide applicable (Hearts & Minds, 2016). The international scale could be seen as an advantage, however for the Safety ladder this caused some difficulties, because this system had the intention to be more specific for the Dutch situation and was too general beforehand (External Interview 15).

Besides Hearts & Minds, the Safety ladder is also based on the model of Parker, which emphasizes the cultural perspective of safety (ProRail, 2015). The model developed by Parker, Lawrie and Hudson (2006), is a framework for progress and maturing of the organizational safety culture and which consists of 5 levels as well. The format of the framework was based on the oil and gas industry (Parker et al., 2006).

The third case, the MVO PL, is not built on specific example systems, but is inspired by various standards including ISO 26000, AA1000 Stakeholders, Global Reporting Initiative (GRI), the triple bottom line (People, Planet, Profit) and ISO 17021 (MVO, 2013¹). Due to the fact that the before mentioned standards are not intended for certification purposes, the MVO PL was established (MVO, 2013¹).

The fourth case, the DJSI, is not specifically based on already existing systems. The DJSI is not a system which includes levels of maturity, but functions as an index or benchmarking system to compare businesses in their sustainability performance (Knoepfel, 2001). The index is aligned with the standards of the GRI. The companies who join the benchmark are all stock market related organizations and CS is increasingly being used to convince investors in buying a stock (Knoepfel, 2001). The participating organizations fill in an extensive questionnaire, to determine their score on the DJSI (External Interview 2).

The fifth and last case, the CDP, is also not based on an existing system. The CDP has similarities with the DJSI, because they both are benchmarking systems to compare organizations by means of a questionnaire and scoring. The questionnaire of the CDP will be repeated every year in order to get the organization's awareness of greenhouse gas emissions, risks, opportunities and management strategies towards a cleaner future (CDP, n.d.; Kim & Lyon, 2011). The results of the questionnaire are public and online available on the website of CDP. Participating in the CDP, and thus disclosing the organization's information and data, is voluntary (Kim & Lyon, 2011).

4.1.5 Process vs performance

When assessing different systems and comparable cases, an important remark is the difference between process and performance (or outcome). Despite the fact that most of the systems are called 'performance ladder' or 'performance measurement system', the emphasis is still on the process side of the organizations. The focus is thus on the management and strategies instead of actual results. In this sub-section, the comparable cases will be analyzed in order to investigate whether they are focusing on processes or performance.

Starting again with the CO₂PL, one could argue that the system is mainly focusing on the process side of an organization, rather than actual results or sustainability performance. The higher the level of the organization, the more mature the organization is. For example, this will mean that targets are improved, measurement systems and data collection have been increased and communication plans are getting better (External Interview 9). Due to this, it becomes clear that the actual results are not the main objective of the performance ladder.

The second case, the ProRail Safety ladder, is also fully orientated on process performance and measures improvements in the processes towards a safer organization (External Interview, 15; ProRail, 2015). From the interview for this comparable case (External Interview 15) it became clear that ProRail deliberately has decided to exclude the word 'performance' out of the system, because the organization has the purpose to keep the emphasis on the processes towards increased safety in the workplace. Related to this, the safety ladder values the internal culture of an organization instead of its results.

The third case, the MVO PL, gives evidence whether an organization has sufficient management policies towards CSR. Like the CO₂PL, the MVO PL gives insights in sustainable development in an objective and certifiable way. In turn, the MVO PL also emphasizes the process performance of an organization (MVO, 2013²).

The fourth case, the DJSI, executes the benchmark both from a process perspective as well as from an actual (quantitative) performance perspective. In their questionnaire they ask for actual numbers, percentages, thresholds or KPI's for example related to greenhouse gas emissions or water use. This represents the actual performance in sustainability practices. However, the index also contains of process and/or management related parts, for example how environmental reporting and communication is arranged or what measures are in place to mitigate risks (RobecoSAM, 2015²).

The fifth case, the CDP, is divided in two types of scores: (1) the disclosure score in which the comprehensiveness of disclose is measured and (2) the performance score which scores the degree of actions taken on climate change (CDP, 2015²). This means that the total score consists of two separate scores: one for the disclosure and one for the performance. The system suggests that performance as well as process will be measured. This is true to a certain degree, because the system measures performance data (CDP, 2015²). However, in order to be eligible for the performance score a minimum in the disclosure score should be achieved and so the focus is on disclosure. Furthermore, the performance score measures emission data but also actions undertaken in order to have a positive contribution on climate change. Despite the fact that it belongs to the performance score, it is still partially 'process' related.

In conclusion, the majority of the above analyzed examples has a process oriented perspective. It can be stated that, although they call it a performance ladder or system, the actual outcome performance is not or partially taken into consideration. From these systems it can be derived that it is difficult to include the actual performance in a performance measurement system. Due to the difficulty of measuring actual performance, some of the comparable cases chose explicitly not to include the performance part and focus only on the process and management side of the organization. However, this will not mean that performance is not measured at all, but the emphasis is more on the strategies and management plans (Neely et al., 1995).

4.1.6 Main advantages and disadvantages

After assessing the five mentioned comparable cases, the main advantages and disadvantages of using these examples can be determined.

The main disadvantage of the first three systems (the CO₂PL, ProRail Safety ladder and MVO PL) is the limitation due to the Dutch focus. These systems do not or only partially include international perspectives, although the PMS for the dairy industry should have an international scope. Due to the national emphasis of the systems, the information derived from these is less applicable for a worldwide system. However, the national focus of the CO₂PL and ProRail Safety ladder is related to the fact that it is designed for the construction industry, which is a less internationally orientated industry than for example, the dairy industry (External Interview, 9). In contrast with the CO₂PL and the

Safety ladder, the MVO PL does not focus on one specific theme or industry only and is applicable for organizations in all sectors who have management in place regarding CSR or CS (MVO, 2013¹) and which can be seen as an advantage. Additionally, the general reasoning of the maturity levels in these 3 systems was in turn very useful to take into account for the design of the PMS for sustainable dairy. The main disadvantage was the focus on the Dutch situation and the main advantage was that the levels were very clearly described and therefore useful as a base or general framework for the design of the sustainable dairy PMS.

The other two systems (DJSI and CDP) were a bit different than the first three. These two systems have the international scope, which is an advantage. Furthermore, both systems have, besides the global view, a regional subdivision which takes into account regional differences (CDP, 2015¹; Knoepfel, 2001) and in turn provides a fair benchmark. However, a disadvantage of these systems is that they are missing maturity levels and containing only a scoring and benchmarking system. Moreover, besides the CO₂PL, Safety ladder and CDP, the DJSI is used for a variety of products and sectors worldwide (Knoepfel, 2001), and because its features are widely applicable it is a useful system to consider.

By combining the information derived from these five systems, the limitations of the first three and the last two systems could be overcome. In that way, the systems formed a useful contribution to design a PMS for the dairy industry.

4.2 Towards an intermediate conclusion

In this section, the comparable cases used for this study were addressed. These comparable cases are the CO₂ Performance ladder, the ProRail Safety ladder, MVO Performance ladder, Dow Jones Sustainability Index and lastly the Carbon Disclosure project. The most useful information derived from these cases was, how to differentiate levels of maturity and how to manage a certain system, thus the underlying management issues of a PMS. Furthermore, the DJSI and the CDP are international systems. These were helpful in investigating how the design of such a system on a global scale can be performed, taking into account regional differences. Additionally, all five systems were useful to be taken into consideration for this study. They were reliable due to the number of members or participants and period of commissioning. However, for most of the comparable cases the emphasis is on the process side of performance and therefore it is difficult to discover how actual (outcome) performance can be measured in a performance measurement system.

The definition of the first set of requirements (section 3), the analysis of practical cases and examples of existing PMS's and interviews, has led to more insights in how to build a proper, credible system. This in turn, has led to a second set of requirements to which the system should comply. These requirements are described in the following subsection and shown in table 11.

4.2.1 Second set of requirements and conditions for a good working PMS

First of all, from a number of interviews (13 out of 26) it became clear that although the PMS will be a global system, regional differences should be taken into account. Thus, the system should be usable on a worldwide scale and should be applicable for different regions (External Interview 3,4,8,9,11,12,13,14;15 Internal Interview 7,8,9,10). In relation to this, the PMS should also fit to a range of organizations and it should be recognized that every organization is different, like size or type. Besides regional differences, also organizational differences should be taken into account (External Interview 5,8,9,11,12,13,14; Internal Interview 3,7,8). So, the system should contain a general, worldwide framework with a regional applicability.

Besides these differences, one of the most important characteristics of a PMS is that it should support continuous improvement (Bourne et al., 2000; Cagnin et al., 2005; Vermeulen & Witjes, 2016), so the system should be updated regularly. Therefore, the system requests for flexibility and feedback possibilities. A good working PMS should include an effective mechanism for the review and revision of targets and standards (Bourne et al., 2000). This is also confirmed by 11 out of 26 interviews (External Interview 3,6,7,9,13,14,15; Internal Interview 2,4,7,8) and corresponds with comparable cases. The example systems emphasize the importance of continuous improvement and feedback possibilities as well and have similar procedures in place to improve the system over time. In turn, it is crucial that the levels and criteria per level are changeable over the years. Thus, the question arose whether the criteria will be changed over time or new levels will be added? From the interviews became clear that there was conformity about the best way to change the criteria.

Table 10. Continuous improvement of the system

Changing criteria over time	Adding levels over time
Internal Interview 2, 7, 8	Internal Interview -
External Interview 3, 4, 9, 10, 13, 14, 15	External Interview 12

The majority of the interviewees prefer to change the criteria over time instead of adding new levels after level 5 (Table 10). This is also mainly due to the fact that the system will grow intensively if levels will be added throughout the years and complexity will rise.

Furthermore, the system should have the right balance between challenging and discouragement (Bourne et al., 2000). In other words, the PMS should challenge frontrunner organizations to become even better and should also support organizations who start to orientate themselves on sustainability (External Interview 8,11,12). Thus, the entry level should be achievable for a wide range of organizations and at the same time the highest levels should be challenging enough for frontrunners. According to this, the system should have a positive approach, to motivate organizations in their sustainability journey and should not discourage the beginning organizations (External Interview 6,7).

Moreover, there is a possibility that a PMS could bring up a certain 'race to the top' (Observation 10; External Interview 11,12,16). The purpose of a PMS is to support and help organizations to improve their sustainability practices and this should in turn not lead to a rapid value loss of the performance of organizations in the lower maturity levels. To prevent a certain 'race', that organizations go to the highest levels as fast as possible, it is important to build the levels in such a way that each level includes and transcends the previous level (Van Marrewijk & Werre, 2003; External Interview 9). With this, it is not possible to skip levels or go through the levels too fast. Furthermore the criteria should be updated regularly. By doing this, competition and progress is in place, though the 'top' will not be reached too easily (External Interview 16).

Despite that a PMS should support organizations in their strategies, it could discourage or 'scare' organizations because they will not easily provide their internal information or they are afraid for more administrative burdens (Bourne et al., 2000; External Interview 9). Despite these possible negative effects, External Interviews 6 and 7 emphasize the positive, supporting role of a PMS or benchmarking system in general, because it provides insight in how to improve the organization.

From literature was derived that internal support is important to create certain sustainable awareness (Cramer, 2005). One of the first things an organization should undertake, derived from the interviews (External Interview 9; Internal Interview 1,5) was that, in order to create support amongst employees, the internal culture of the

organization has to be changed towards a more sustainable one. To create this cultural change, internal communication is important and should be carried out before external communication takes place (Cramer, 2005). Thereafter, and related to this requirement, transparency is one of the most important issues within the organization as well as in the entire supply chain (Cramer, 2005; External Interview 4,5,12; Internal Interview 5,6).

Table 11. Second set of criteria for a Performance Measurement System

Second set of criteria	Source
A PMS should be applicable for different regions and therefore regional differences should be taken into account	External Interview 3,4,8,9,11,12,13,14 Internal Interview 7,8,9,10 Observation 5,6
A PMS should be applicable for different types/sizes of organizations and therefore organizational differences should be taken into account	External Interview 5,8,9,11,12,13,14 Internal Interview 3,7,8 Observation 5
A PMS should support continuous improvement and should be updated regularly	Bourne et al. (2000) Cagnin et al. (2005) Vermeulen & Witjes (2016) External Interview 3,6,7,9,13,14,15 Internal Interview 2,4,7,8
Changing the criteria per level is preferred instead of adding new levels over time	External Interview 3,9,13,14,15 Internal Interview 2,4,7,8
A PMS should have the right balance between challenging frontrunner organizations and not discouraging starting organizations	Bourne et al. (2000) External Interview 8,11,12
A PMS should support organizations to improve their practices rather than supporting 'a race to the top' only	External Interview 8,11,12,16
Prevent a race to the top, by making sure that each level includes and transcends previous levels	Van Marrewijk & Werre (2003) External Interview 9
A negative effect of a PMS could be that organizations are scared or discouraged, because they will not provide their internal information or are afraid of administrative burdens	Bourne et al. (2000) External Interview 9
One of the basic things of the system is the creation of support amongst employees in the internal base of the organization	Cramer (2005) External Interview 9 Internal Interview 1,5
Transparency within the system and within the organization is one of the most important issues	Cramer (2005) External Interview 4,5,12 Internal Interview 5,6

Thus, the criteria shown in table 11 provide information about the general system and not yet about sustainability maturity levels. Keeping the above mentioned criteria in mind by designing the PMS, helps to build a more reliable and credible system. Additionally, in the next sub section, the features of maturity levels derived from empirical research will be given.

4.2.2 Characteristics of maturity levels from empirical research

After assessing the characteristics of maturity levels derived from literature, in this section the main features per level derived from empirical research will be discussed. In addition, a general table that shows the overall overview of the characteristics per level and per sub-category can be found in the back of the thesis (Appendix G).

First of all, derived from the analysis in NVivo, there could be determined which subjects or issues were the most prevalent in the conducted interviews. Derived from this analysis, the setup of the different levels was by far the most discussed subject (17 out of 26 interviews). Furthermore, how to apply continuous improvement mechanisms in the PMS was often discussed (12 out of 26) and the process versus output 'dilemma' (13 out of 26) came across in many interviews as well as examples of comparable cases which were useful for this research.

When looking at the different categories discussed in the desk research results, the path described in the approach and vision of the maturity levels is quite similar to the results derived from desk research. Again, the transitions from reactive towards proactive and from informal towards formal defined programs were recognized (Table 12) (External Interview 3,16; Internal Interview 3,7,8; Observation 9,11).

Table 12. Levels of maturity in the category Approach & Vision

Category	Level 1 <i>Initial</i>	Level 2 <i>Recognized</i>	Level 3 <i>Defined</i>	Level 4 <i>Managed</i>	Level 5 <i>Optimized</i>
Approach & Vision	<p>The organization does not have a plan or program^{O 11}</p> <p>Starting orientation/awareness of sustainability/CS should be in place^{EI 4,6,7,8}</p>	<p>The organization is developing a plan or program (informal)^{O 9c,11}</p> <p>Emphasis is on reactive approach rather than proactive^{EI 3}</p> <p>A clear plan is lacking^{EI 3}</p>	<p>A plan or program is formally defined^{O 9c,11}</p>	<p>A program is defined and communicated^{O 9c,11}</p> <p>Proactive^{II 3,7,8}</p>	<p>A program is defined, communicated and integrated^{O 9c,11}</p> <p>Proactive approach^{EI 16; II 7,8}</p> <p>Continuous improvement and emphasis is on 'to be better than the competition'^{II 1}</p>

CC) Comparable Case; EI) External Interview; II) Internal Interview; O) Observation

From interviews and comparable cases became clear that communication (Table 13) is a very important driver and/or indicator for continuous improvement of sustainability and to create internal support amongst employees (External Interview 5,9,16; Internal Interview, 1,5,7,8). One of the interviewees said: 'You can't buy sustainability, it is an experience' (Internal Interview 2). This will also mean that, at the highest maturity level, sustainability strategies and management plans should be totally embedded in the organization (External Interview 4). Only one interviewee explicitly suggested that communication is not essential for measuring levels of CS (External Interview 3).

Table 13. Levels of maturity in the category Communication

Category	Level 1 <i>Initial</i>	Level 2 <i>Recognized</i>	Level 3 <i>Defined</i>	Level 4 <i>Managed</i>	Level 5 <i>Optimized</i>
Communication	<p>First step: create internal support and awareness^{EI 9; II 5,6}</p> <p>Internal^{II 5,6}</p> <p>Internal towards external is better, because the first step to go forward is to create the right culture in the organization^{II 1,5,6}</p>	<p>Internal support and awareness^{EI 9; II 5,6}</p> <p>Internal^{II 5,6}</p>	<p>Internal support creation amongst employees and external embedding^{EI 5; II 5,6}</p> <p>Internal towards external^{II 5,6}</p>	<p>From rather internal towards external and public (between company and NGO's and governmental parties)^{CC}</p> <p>Emphasis on external, but also internal^{II 5,6}</p>	<p>Communication is very important in all levels, through the entire process of maturity^{EI 16; II 5,6,7,8; O 13}</p> <p>Importance of transparent external communication is high^{EI 4; II 5,6}</p>

CC) Comparable Case; EI) External Interview; II) Internal Interview; O) Observation

Furthermore, from desk research it became clear that stakeholder engagement was also very important (Kaptein & Van Tulder, 2003). This issue was addressed in the interviews and one can state that the interviewees confirmed its importance. This issue gets more conviction by adding a few statements: one of the interviewees explicitly suggested that *'Stakeholder engagement and commitment are very important, they are the ones who must agree with the goals the organization imposes.'* (External Interview 6). Another quote related to stakeholder engagement is of External Interviewee (15) *'You'll have to climb the ladder together'*. However, another interviewee suggested that the emphasis is more on 'how' stakeholders are engaged rather than 'how often' they are engaged in stakeholder dialogues or other types of meetings or feedback sessions (External Interview 2). Thus, this is important to take into account when designing a PMS.

In relation to this, another interviewee emphasized the importance of societal engagement rather than direct stakeholder engagement only (External Interview 16). Again, this can be illustrated by a quote from the interview *'When you engage your stakeholders and these stakeholders go with you in the right direction, that's fine. However, what actually counts is whether the organization is concerned with societal issues'*. (External Interview 16). This could also be confirmed by a comparable case example. In one of the comparable cases an organization could only reach a high level of maturity (mainly level 5), if it engages the entire society. Societal issues and concerns should be taken into account as well as the opinions of the most important or direct stakeholders. In other words, the perspective should be changed. Instead of determining important issues with the main stakeholders, societal concerns should be at the base of the materiality assessment (External Interview 16). Again, the focus or approach of the organization has to move from internal towards external by going through the levels. So, this is both from internal towards external communication and reporting, as well as the focus on own organization versus cooperation with stakeholders in the supply chain (External Interview 6,7,9,13,14,16; Internal Interview 5,6).

Regarding communication, a sustainability expert suggested that in some cases it could be that communication starts from an external perspective and changes towards a more internal focus, due to the fact that organizations would like to show what progress they make towards increased sustainability and does not have appropriate internal support (External Interview 5). This is true in some cases. However, two communication experts stressed that the 'internal towards external approach' is the most optimal one. This is again related to the internal embedding and support of sustainability, which is one of the first steps that should be taken in order to increase sustainability practices (Internal Interview 5,6).

Consequently, and related to stakeholder commitment, transparency is seen as an important factor; *'accountability and measurability is very important in order to gain trust among your stakeholders'* (Internal Interview 6). Moreover, to gain trust, the organization should be transparent about its processes on sustainability and this feature is seen as a crucial one to score high on the maturity system (External Interviews 4,5,12; Internal Interviews 5,6). This means that organizations should be able to show which activities they undertake and this is what transparency is supposed to be (External Interview 12; Internal Interview 5,6).

Table 14. Levels of maturity in the category Materiality Assessment

Category	Level 1 <i>Initial</i>	Level 2 <i>Recognized</i>	Level 3 <i>Defined</i>	Level 4 <i>Managed</i>	Level 5 <i>Optimized</i>
Materiality assessment	No assessment ^{EI 2; O 9a} Organization should start with determining which (environmental) impacts it has ^{EI 8}	<i>Internal assessment</i> ^{EI 2,7} Assessment, but not verified ^{O 9a}	<i>Starting with an external assessment, with selection of stakeholders</i> ^{EI 2,7} Assessment is communicated, not verified yet ^{O 9a}	<i>Extensive stakeholder survey to assess materiality</i> ^{EI 2,16} Assessment is communicated and verified with external parties ^{EI 8; O 9a}	Stakeholder engagement as well as societal engagement ^{EI 2,16} Assessment is communicated, verified and maintained ^{EI 8; O 9a}

CC) Comparable Case; EI) External Interview; II) Internal Interview; O) Observation

The path of maturing in materiality assessment and determining key issues is very similar to the path described in literature. In the initial level there is just an informal assessment, which only addresses the main impacts of the organization, or there is no assessment at all. This is followed by a more formal assessment and finally by a comprehensive stakeholder dialogue with the possibility to give feedback and to engagement society. This, in order to assess materiality of the key issues for the specific organization (Table 14) (External Interview 2,6,7,8,16; Internal Interview 5; Observation 9A).

Besides transparency and communication, is the importance of reporting. Reporting is common from approximately level 3. In the higher levels, besides reporting about plans and processes, results and goals should be taken into account as well (Table 15) (External Interview 2,13,14). This is in compliance with what was described in literature.

Table 15. Levels of maturity in the categories Reporting, Output & Results

Category	Level 1 <i>Initial</i>	Level 2 <i>Recognized</i>	Level 3 <i>Defined</i>	Level 4 <i>Managed</i>	Level 5 <i>Optimized</i>
Reporting, Output & Results	Output and results will not weight very much ^{EI 5,6,7} Reporting is not in place ^{II 5}	Output and results will not weight very much ^{EI 5,6,7} <i>Reporting starts to be in place</i> ^{II 5}	Output and results become important to include in the scoring ^{EI 3} <i>Focus will be more on outcome</i> ^{EI 6,7; O 13} <i>Reporting about most important sustainability issues</i> ^{II 5}	Output and results should be clear and targets should be met ^{EI 1,3} Focus on outcome ^{EI 6,7; O 13} <i>Besides results, also goals and measures are well reported</i> ^{II 5} <i>Reporting as well as communication have to be transparent</i> ^{II 5,6}	Organization should prove its results and own targets as well as supply chain goals are met ^{EI 1,3} In the end, the focus should be on outcome ^{EI 6,7; II 5; O 13} Besides results, also goals, measures and progress towards the goal are well reported ^{EI 2; II 5} Reporting as well as communication have to be transparent ^{II 5,6}

CC) Comparable Case; EI) External Interview; II) Internal Interview; O) Observation

On the contrary, a new issue, not yet frequently seen in literature, came across in a few interviews and an observation (External Interview 3,8,12; Observation 9A). Verification is an indicator that is not frequently mentioned, although it is seen as an important one. The trend in verification is as follows; in the lower levels, verification is not in place or is internally verified only. From approximately level three, external verification starts to occur and in the highest levels verification is in place and should be certified (External Interview 3). Additionally, verification of criteria is a useful indicator to determine whether or not the criteria in the maturity level system are measurable and in turn relevant (External Interview 3). This can be linked to transparency again, because organizations should be able to give evidence for their practices.

Additionally, one of the issues that came across in several interviews (13 out of 26) is the division of process and output, which was also addressed in the comparable case section (sub-section 4.1.5). These 13 interviewees missed the actual output or results side of the performance measurement system. However, from literature and comparable cases became clear that it is difficult to measure actual outcomes because of regional differences and organizational differences. So, the solution to tackle this is to keep the guidelines general in order to be applicable on a larger scale and specify some regional dependent guidelines to keep the regional applicability (External Interview 4,9,12,15; Internal Interview 7,8,9). Furthermore, goals or targets could be determined, however these have to be relative improvements, rather than actual, absolute targets, because absolute targets are difficult to compare between different organizations (External Interview 2). So, again there could be concluded that actual performance is difficult to measure in a PMS and therefore the focus will be more on the process towards sustainability.

Subsequently, another issue that was not derived from literature beforehand and was not predominant in literature was the difference between leading and lagging indicators. This was discussed during Internal Interviews (7,8). They stressed that the difference between these two is important in measuring maturity, because this indicates whether the organization is at a high level or is still in the beginning phase of the system. After these interviews, the scientific articles were re-examined to investigate whether these indicators were researched. One of the articles, Neely et al. (2000), described on one hand lagging indicators as measures that relate to results, for example cost related performance. On the other hand, leading indicators are related to causes of the results, so these emphasize on quality, flexibility, innovation etc. (Neely et al., 2000). Thus, lagging indicators emphasize on past performance and do not take into account future prospects (Neely et al., 2000). Referring to Internal Interviews (7,8) the more leading indicators an organization uses, the more mature the organization is, because *'lagging indicators are already passé for these organizations'* (Internal Interview 7,8).

Thus, the main findings from empirical research were the importance of communication, to create internal support of sustainability as well as to be transparent towards stakeholders in order to create trust amongst your stakeholders and society. Furthermore, materiality assessment is leading in order to determine the priority of different sustainability issues. Here it is interesting to see that the path towards maturity is quite similar as described in literature. Additionally, the main differences and similarities between the literature review findings and the results presented in this section are discussed in the following section (section 5).

5. Results part one and two: Synthesis & PMS Design for dairy

After assessing the general sets of requirements, a synthesis of both parts of the research (one and two) is executed. With this synthesis a comparison is made between data derived from scientific literature and data derived from interviews and comparable cases. In other words, in this sub-section there will be determined whether the criteria from literature correspond with data from interviews and practical examples and what can be learned from practice. Furthermore, in this section the structure of the design of the PMS for the dairy industry will be discussed.

5.1 Main differences and similarities

The findings from literature and interviews were very similar, most of the interviewees agreed with the features mentioned in literature. However, there are some differences between what is described in literature and what is seen in actual cases. After describing the results from desk research and empirical research separately, a more detailed comparison could be made and will be given in this sub-section.

One of the main differences between interviews and desk research was that in scientific reports, clear characteristics per level were defined, whereas in the interviews a more general growth in maturity was given. Consequently, in the interviews mainly the trends from low towards high maturity in sustainability practices were described rather than strict features per level. However, after all interviews were conducted, these trends were translated in features per level as well. Another remarkable finding from empirical research was regarding communication. In certain cases the transformation from external towards internal is recognized, instead of from internal towards external, which is seen as the most desired path. In literature the focus was on the development from internal towards external only. Additionally, from interviews (8 out of 26) became clear that communication in general should have the emphasis, because this is seen as an essential part for internal support creation and in turn for reaching higher levels of corporate sustainability. In desk research, this focus was less prominent. Furthermore, transparency within the organization as well as towards the stakeholders and society was highlighted as an important point of attention in the interviews, but was not frequently found in the first part of this research (desk research).

On the contrary, from both perspectives the results also had many similarities in the characteristics of a level based measurement system and the similarities emphasized the importance of the specific characteristics. Main similarities found in this study were related to stakeholder engagement and continuous improvement of the system. The importance of stakeholder engagement and/or stakeholder dialogues was stressed in both parts of the research. Because of this, one can state that the involvement of stakeholders is crucial for the sustainable development of an organization. Furthermore, continuous improvement of the system is recognized as an important aspect. Regarding continuous improvement, sufficient feedback mechanisms and communication (internal as well as external) is essential. The preference is given to changing criteria over time, to keep the system up to date.

The difference between process versus performance and the 'lack' of performance is an essential element that came across in many interviews and was studied in scientific reports and example systems too. However, from this research it became apparent that actual performance is difficult to measure and well-defined process performance will in turn lead to better results and actual performance (Baumgartner & Ebner, 2010). How this balance between processes and performance is applied to the dairy industry will be explained in more detail in sub-section 5.2.6.

5.2 Towards an intermediate conclusion

One can state that the results from both parts of the research were quite similar. In both parts of the research it became apparent that stakeholder engagement is very important to improve the sustainability practices of an organization. However, in the empirical research, the emphasis was on communication, which was not predominant in the findings from desk research. In order to become higher on the sustainability ladder, communication is essential and internal support amongst the employees is seen as a starting point for improving sustainability practices. The importance of the creation of a transparent and reliable supply chain is emphasized in the interviews as well.

After assessing these different research techniques, a broad framework is developed and the subject of PMS's is studied in-depth. However, thus far most findings were described very generally and were not yet applied to the dairy industry. By knowing the characteristics derived from different kinds of sources, this general information can be applied to the dairy industry. Thus, at this point, the general framework and knowledge will be applied to the dairy industry and in particular to the Dairy Sustainability Framework (DSF).

5.2.1 The structure of the PMS for the dairy industry

In this sub-section, the structure of the PMS for the dairy industry will be explained in more detail. The system will contain 5 levels, which was previously discussed in sub-section 3.5.2, and these levels are based on the information provided in the previous parts. Thus, level 1 will be the *initial* level, level 2 is *recognized*, level 3 is *defined*, level 4 is called *managed* and level 5 is the *optimized* level. With the knowledge of these levels in mind, specific features for the levels are applied to the various parts of the DSF Demonstration Scheme (Figure 4). In line with the five maturity levels, each statement that belongs to a certain part of the DSF demonstration scheme can be tested on a 5-point scale. A statement belongs to each point on the scale with which organizations can identify themselves. So, each column is divided in different sub issues and each sub issue has a few statements which are subdivided in 5 levels.

The DSF Demonstration Scheme (Figure 4) shows different steps an organization should take in order to function in a sustainable way. Besides these steps, business principles and core themes regarding sustainability are included. Furthermore, enablers and already existing standards, that influence the sustainability management and performance, are included in the scheme as well. The scope of the DSF is the entire dairy value chain (DSF, 2015). Although this demonstration system is developed, levels of maturity are not yet included into the scheme. In this study these levels were designed and applied to the existing scheme in order to develop a PMS towards continuous improvement of sustainability in the dairy industry.

The PMS for the dairy industry, based on the DSF, will be an independent, online self-assessment tool. This self-assessment will provide organizations with insights in their sustainability performance. By making use of an independent self-assessment, the system is easily accessible for all types of organizations and is unbiased by one organization and these are also two of the principles of ISEAL Alliance (2014) in order to build a proper system. Once more, before this research, there were no levels of maturity, only the one dimensional framework or scheme was present.

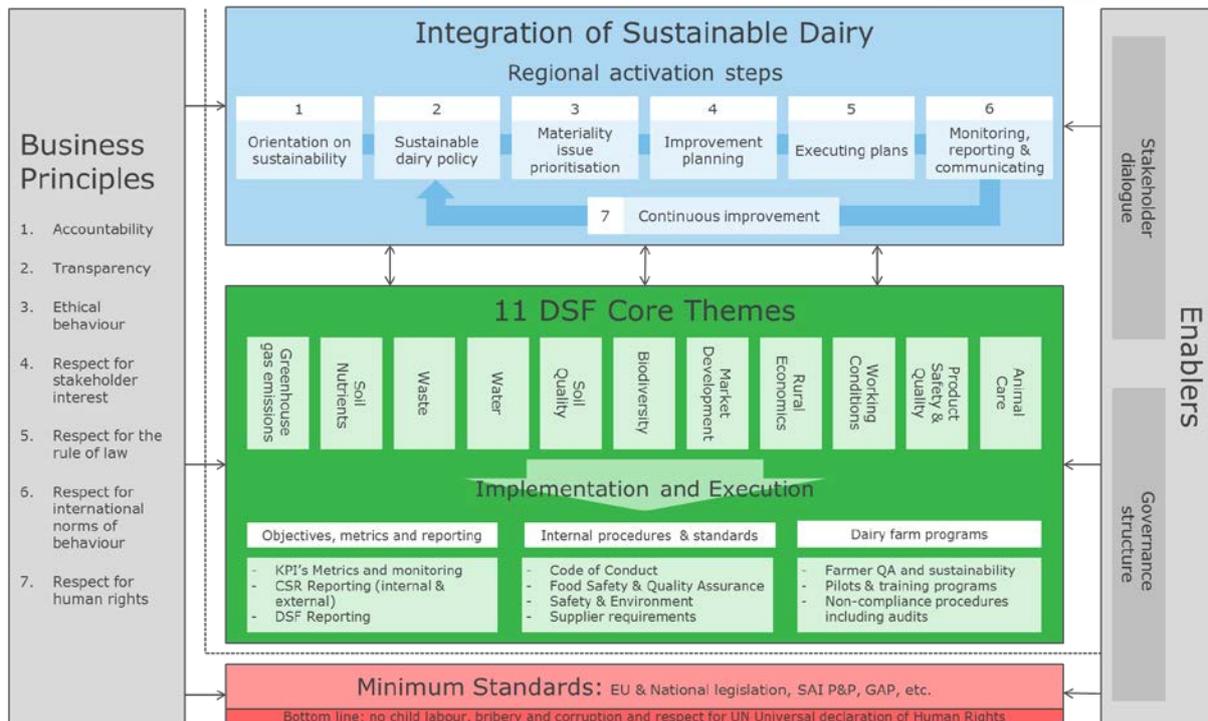


Figure 4. DSF Demonstration Scheme (FrieslandCampina, 2015)

The adding of maturity levels has been started at the left column, *Business Principles*. These principles are mainly based on the ISO 26000 standard and therefore, the statements representing the levels underneath the 7 principles were based on ISO 26000 too. ISO 26000 is a standard regarding social responsibility and the business principles in the left outer column represent the items that are needed to operate in a socially responsible, transparent way (Castka & Balzarova, 2008; ISO, n.d.). Therefore, they are important determining factors for the CS maturity of an organization. Furthermore, GRI, ISO1000, 9001 and 14001 were also used to base the statements. Additionally, it is important to connect the system to standards like ISO, to give it more credibility and it makes the system 'unquestionable' in comparison with systems that are not based on a standard (External Interview 1; Observation 13). Moreover, the levels built in these principles were based on the knowledge obtained from parts one and two of this research.

The *Business Principles* are followed by the *Enablers* part (right column), in which stakeholder dialogues and the governance structure of the organization are introduced. Statements related to these issues are mainly derived from Philipson et al. (2011), Baumgartner and Ebner (2010), Maak (2007) and Harris (2000), who describe the relevance of stakeholder engagement, dialogues and knowledge exchange. Furthermore, for stakeholder dialogue and engagement the standard AA1000 is used. The AA1000 is a standard regarding the improved quality of social relations and stakeholder engagement of organizations. With this standard, organizations will be able to better base their decisions regarding their stakeholders and in turn increasing trust (AccountAbility, 1999).

The levels for these parts were again grounded on the main articles and empirical data. An appropriate governance structure is an important enabler, because it determines if the organization has a leadership role, is capable to share knowledge and takes part in information systems (Harris, 2000), which in turn leads to a higher maturity level. Besides governance structure, stakeholder dialogues are crucial to develop the organization's CS (External Interview 6,7,16). With stakeholder dialogues, and more general stakeholder engagement, the organization increases its embedment in society

and in turn its sustainability (Kaptein & Van Tulder, 2003). So, the better the organization is embedded in its broader context, the higher its CS will be.

The next step was to build levels for the *Integration of Sustainable Dairy* (blue column). These 7 steps follow the implementation of the DSF 11 core themes (green column) and are developed according to the plan-do-check-act principle. The statements underneath these steps were developed in consultation with interviewees and are also linked to ISO 14001 and 9001. Again the division of levels was based on the leading articles. The plan-do-check-act cycle, in short PDCA cycle (Figure 5), needs some clarification related to the integration of sustainable dairy.

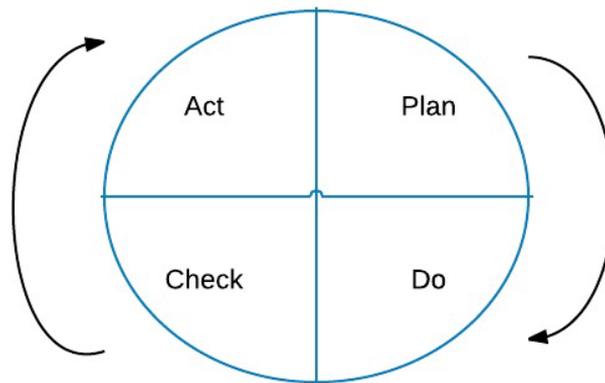


Figure 5. Plan-do-check-act cycle (Moen & Norman, 2006)

The PDCA cycle is a widely used approach to measure continuous improvement in environmental management systems since the late 1980's (Moen & Norman, 2006; Vermeulen & Witjes, 2016). Many organizations worldwide use this approach as a single or repeated loop to identify improvement steps (Vermeulen & Witjes, 2016). When applying this cycle to the DSF, the 'plan' phase is linked to the setting of plans and policies towards sustainability, i.e. implementation steps 1 till 4. When this planning phase is completed, the plans can be implemented in the 'do' phase, i.e. step 5, where after they should be checked for effectiveness and goals achievement, which is executed in the 'check' part of the cycle and linked to monitoring and measuring, i.e. step 6. Eventually, in the 'act' phase organizations will adjust their plans and goals when needed and are able to improve their activities towards achieving their goals, i.e. step 7 (DSF, 2015). By making use of a repeated cycle of activities, continuous improvement will take place and this is also one of the requirements of a sustainability standard (ISEAL Alliance, 2014; Potts et al., 2014), which states that a system should be able to adapt to the dynamics related to sustainable markets.

Furthermore, following the content of DSF, organizations should be in compliance with local rules and regulations to join the framework and therefore this will be the entry level of the PMS as well. Due to this, *Minimum Standards* (red box at the bottom of the scheme) are important to include in the system. To apply the minimum level at a large variety of cases, a distinction is made between global minimum standards and regional minimum standards. The features of the entry level and the minimum standards will be further explained later in this thesis (sub-section 5.2.2).

After assessing the previously explained, which are more general outer 'columns', the core of the system is elaborated by adding levels to the implementation of the *11 Core Themes*. For these themes, interviews were conducted with experts of the specific themes (External Interview 13,14; Internal interview 7,8,9,10) and a variety of statements is related to GRI-G4 indicators. The GRI-G4 is focusing on the sustainability reporting guidelines and principles and standard disclosure of specific aspects (GRI, 2015). These aspects could be linked to the DSF themes. These *Core Themes* were the most difficult to divide into levels, because organizations face a variety of themes due to

regional differences, for example water use or animal care. Because of this, there is chosen for a worldwide approach of the outer columns (*Business Principles, Integration of Sustainable Dairy and Enablers*), a division of regional and global *Minimum levels* and a more regional applicability for the *Core Themes* (External Interview 3,8,11,12,13,14; Internal Interview 9,10; Observation 5). However, these subdivisions and regional differences will be explained in more detail in sub-section (5.2.3).

Subsequently, it is important to determine whether the system is designed in such a way that the members of the DSF are willing to join or participate in the system and will use the system to measure their maturity in implementing DSF. To determine if a system is reliable and what the requirements for such a system are, a reference to section 3.2 will be made, in which the requirements for sustainability standards were given. In accordance with this, a few of the principles state that the system should be transparent, truthful and relevant (ISEAL Alliance, 2014) and that the chance that DSF members trust and support the idea of a maturity level system will grow if the system is linked to or based on well-known standards, which makes the system more transparent and relevant (External Interview 1). Moreover, truthfulness is also achieved by relying on existing standards, because they are seen as the ones who set the right claims. In turn, relevance is confirmed by linking the system to existing systems and standards because they are held up-to-date and are approved by several parties. Therefore, a large number of statements for the PMS are based on standards like ISO and GRI (Global Reporting Initiative) and already existing systems.

Lastly, although this is not mentioned by ISEAL Alliance or other associations for sustainability standards, many comparable cases offer a certain reward (in Dutch Gunningsfactor) for joining the system. This motivates organizations to participate in the system (External Interview 15). Thus, in this system something like a reward should be included. In the dairy industry one can think of a sustainability certification when organizations become higher on the maturity ladder, or certain subsidies for sustainable investments could be given.

5.2.2 Entry level

The requirements of the entry level are crucial for the successfulness of the PMS for dairy (Observation 5,13). The entry level serves as the lower boundary of the system and ensures that entering the system is attainable for all kinds of organizations, even if they just started, as not to exclude too many organizations from joining DSF (External Interview 4,6,7,8,9,12). However, even if the entry level is low, the organization should have a certain orientation towards sustainable development and certain cutoff criteria should be included (External Interview 5,6,7,8).

One of the main features of the entry level is, amongst others, compliance with local laws and regulations and this can be seen as the base level for joining DSF (External Interview 2,4,11; Internal Interview 4,7,8,10; Observation 11). This means that, with the implementation of this entry level a certain threshold is built to prevent organizations which do not comply with laws and regulations, to score high on the ladder. So, even if organizations have the right programs and management processes in place, but do not comply with a minimum standard, they will not score higher than level 1. So, from level two or higher, at least members should comply with the minimum standards.

By applying local laws and regulations as the baseline, the system is applicable for all regions and takes into account regional variation. Besides the minimal requirements, customers will make agreements together, to determine which level should be their level for the 'license to produce' (External Interview 3). When certain agreements are not made, a logical consequence could be that the highest level is the only relevant one (External Interview 3) and this should be avoided, because for some organizations the

highest level is yet impossible or too far away to achieve. This balance between challenging and discouragement was also previously discussed in section 4.2.1.

Besides the requirements, transitional periods are also important to take into account. Transitional periods refer to the time allowed for passing the levels. From various comparable cases it became clear that the first levels should be achieved in one year and if this is not met, the organization should leave the system. The higher levels, which are more difficult to achieve, should have a longer transitional period and in the higher levels it is the choice of the organization whether or not they would like to become best in class (External Interview 5; Internal Interview 1). So, they will not be placed out of the system. From comparable cases is also derived that there is a second type of transitional periods, which apply when the criteria of the levels are changed by updating the system. These transitional periods are in place to give the organization a certain transitional period of one year to meet the new requirements and otherwise it will drop to a lower level. However, this is not a subject that is key for the first design of the PMS, but it is good to take into account by further developing and improving the system throughout the years.

5.2.3 Differentiation in regions

From interviews it became clear that some parts of the system cannot be compared on a global scale and that for some parts regional differentiation is needed. It has been decided that the general 'outer columns' (*Business Principles, Enablers and Integration of Sustainability*) of the system are comparable on a global scale, because in these parts the main focus is on process performance (Observation 5,6,13). The *Minimum Standards* are divided into global standards and regional standards. However, the specific *Core Themes* should not only emphasize process performance but also outcomes or goals achievement. Besides, cultural and local environmental differences make it complex to execute a global comparison in the themes. Therefore, these themes will be compared on a regional scale (Cagnin et al., 2005; CDP, 2015¹; External Interview 8,9,13; Internal Interview 9; Observation 6; Matson et al., 1997). When referring back to the principles for reliable standards, the first principle of Potts et al. (2014), the principle of subsidiarity, is related to the regional differentiation of a system. This principle stated that effective tools are the ones that are able to be modified to local issues and settings. By making use of this local applicability, voluntary standards can be approachable to the needs and interests of stakeholders in multiple regions (Potts et al., 2014). So, this principle proves that a successful system needs a certain regional classification and applicability.

When looking for a possible classification, the existing division of DSF by SustainAbility (see left column of table 16) is considered. However, this division is insufficient, because a number of regions is very broad and is defined too generally and there are still many differences within these regions. Therefore, a new, more specific regional classification is made (Observation 6) and is shown in the right column of table 16 and figure 6.

Table 16. Regional division

Regional division from SustainAbility (DSF)	Regional division tightened
Asia Pacific	1. North America (Canada & VS)
Australia	2. Central America (Mexico, Panama, Costa Rica etc.)
Canada	3. South America 1 (Brazil, Peru, Colombia etc.)
Eastern Asia	4. South America 2 (Chile, Argentina, Bolivia etc.)
Eastern Europa	5. Europe
Latin America	6. North Africa
New Zealand	7. Mid Africa
Northern & Western Europe	8. South Africa
Southern Asia	9. Russia
Sub-Saharan Africa	10. Middle East
United States of America	11. India, Pakistan (includes Sri Lanka, Nepal, Bangladesh)
	12. China & Japan
	13. South-East Asia
	14. Oceania (Australia & New Zealand)

The regions in the right column are based on climatological differences and continents as well as existing regional classifications (IFCN, 2015; Kottek et al., 2006). Climatological differences are important when considering themes like water use and water availability, but also animal care (possibilities to behave naturally) and soil related issues (External Interview 8; Internal Interview 9,10). Especially water is a theme which is not easily measured worldwide, because it faces challenges on both local and global scale and therefore faces various regional variations. Water is a very region specific issue, due to i.a. precipitation patterns, differences in natural use etc. (CDP, 2015¹). So, goals or results related to for example water could not be compared globally and should be determined on a regional scale to be able to compare similar organizations and measure their own challenges and performance.

However, in certain regions or countries, goals and targets are not explicitly set, nevertheless this does not necessarily mean that their final performance is not as good as organizations that do have defined goals (Observation 14). So, this is important to take into account as well. When specific goals are not defined in a specific region, this could be adjusted for that region, because that is where the regional categorization or division is made for. In other words, the DSF has to develop or set guidelines per region and so the organization specific goals or progress could be compared to the regional guideline to see whether the organization's progress is ambitious and in compliance with the standards of that specific region (Observation 15).

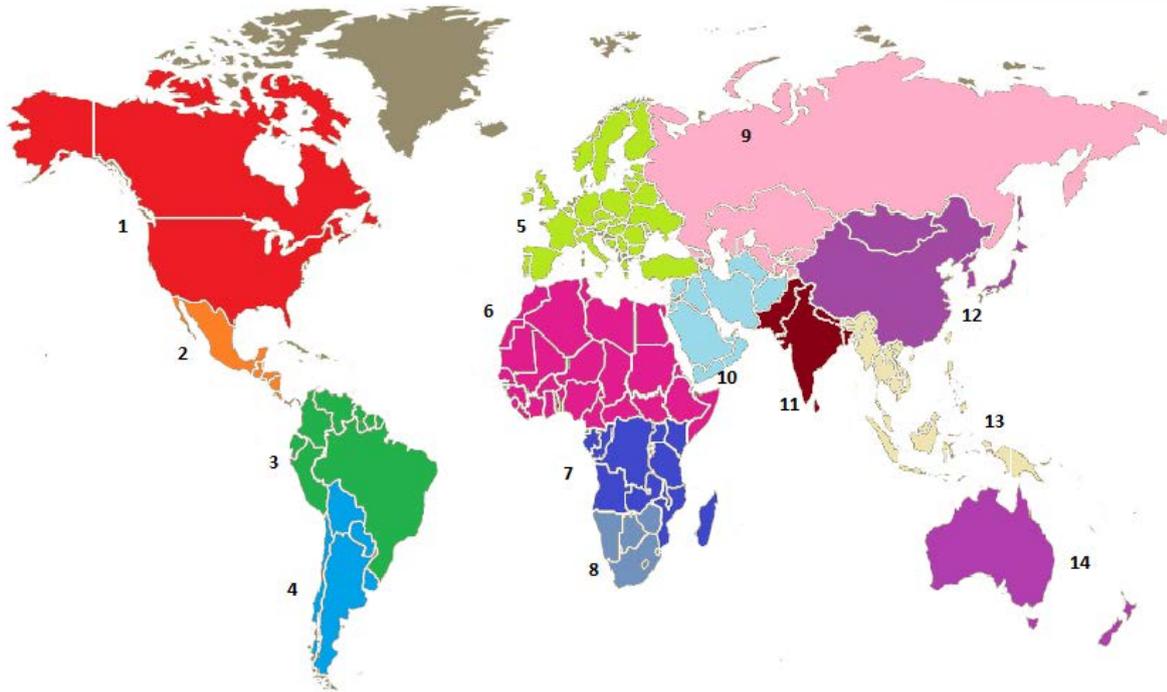


Figure 6. Regional Classification

In conclusion, it is important to take into account regional differences and in order to make a fair and useful comparison between organizations a clustering of organizations in comparable regions is required.

5.2.4 Differentiation in types of organizations

Besides differences in regions, differences in types of organizations should be taken into consideration as well. Again, this came along in a number of interviews (6 out of 26) (External Interview 5,8,11,12,13; Internal Interview 3). According to External Interviewee 15, the structure of the questions in the self-assessment should be in a certain extent generally described, in order to make the system applicable for different types of organizations. This is also seen in comparable cases. However, in some situations, if organizations differ too much, a clear comparison cannot be made. The same as for regional differences, a comparison between organizations of different sizes is unfair. In other words, it is unfair if very small and very large organizations compete against each other (External Interviews 5,8,13,14; Internal Interview 3). In order to prevent this, a certain clustering should be made (External Interview 5,11; Internal Interview 3). From the interviews it became clear that, another difference related to organizational size is communication. In large organizations communication is organized differently in comparison with smaller organizations, due to hierarchical differences and the size of the organization affects the complexity of communication processes (Ghobadian & Gallear, 1997; Internal Interview 3). So, the larger the organization, the more complex the communication process will be (Ghobadian & Gallear, 1997; Internal Interview 3).

Thus, the main difference that should be taken into consideration is the size of the organization and therefore a certain clustering should be made. However, another possible distinction on supplier level is between dairy originated from primarily grazing cows or indoor held cows (External Interview 8). The difference between these two types of organizations can have effect on the specific theme performance, mainly on greenhouse gas emissions and biodiversity (Natuur & Milieu, 2015).

Concluding the concrete steps that should be made are; clustering different types of organizations by differences in size, because this has been found as the leading indicator. There are more differences, such as type of the organization. This should be taken into account as well, but this can also be linked to size. So, to provide a fair comparison between organizations, at least a clustering of organizations with the same size has to be made.

5.2.5 Continuous improvement of the PMS

Subsequently, after assessing regional and organizational differences, another important part of the system is how to manage continuous improvement or, in other words, how should the system be updated through the years? Even though the system will encourage organizations to improve their sustainability practices continuously by making use of the PDCA cycle, the system itself should be continuously improved as well. Continuous improvement of the system is important, due to the ever *'changing world around us'* in which sustainability trends and improvements are changing rapidly (ISEAL Alliance, 2014; External Interview 15; Internal Interview 7,8). So, the system must remain challenging, relevant and should reflect current stakeholders understanding (ISEAL Alliance, 2014). Therefore, continuous improvement of the system is needed in order to *'change together with the developments'* and keep the system up to date (Internal Interview 2).

In order to keep the system up to date, the question arose; whether the criteria per level will be adjusted or new levels will be added (see section 4.2.1 and revisiting Table 10 at page 37). From this previously discussed section (4.2.1) it became clear that preference is given to adjustment of the criteria. To delve deeper in the issue of improvement, comparable cases have been used to get insights in how to apply continuous improvement strategies to the system. From a comparable case and interviews (External Interview 15) is derived that a maturity level system needs to develop over the years. According to ISEAL Alliance (2014) review and revision of the system is desired to keep the system relevant over time. When a system comes into practice for the first time, some shortcomings and flaws are inevitable. By means of this, having a clear feedback system is important to keep improving the system (ISEAL Alliance, 2014; External Interview 9,10,15). Furthermore, tightening the criteria of the system does not mean making large changes, since the criteria will not change entirely, but making the system more consistent and/or univocal over time. However, the criteria of the system should not be adjusted too fast, because organizations need time to get used to and adapt their strategies to the new requirements (External Interview 9,10,15; Internal Interview 2). Therefore, certain transition periods are important to determine (External Interview 9; Internal Interview 4,7,8). Additionally, it is very important to engage society and stakeholders in the improvement process of the system (External Interview 15) and to have a clear leading committee or scheme owner who is responsible for the changes made in the scheme and communicates about it (External Interview 9,10,16; Internal Interview 7,8).

Thus, for the improvement of the system a clear feedback loop which provides insights in possible adjustments is needed, as well as a scheme owning organization or committee to enter the desired adjustments and to communicate these adjustments to its members (ISEAL Alliance, 2014). Furthermore, transitional periods of approximately one year are necessary in order to give members the opportunity to meet the new criteria of the continuous improved system.

5.2.6 Process vs performance

This sub-section will discuss the *'dilemma'* of process versus performance applied to the PMS for the dairy industry. Within the PMS, a difference is made between on one hand processes, i.e. the management side of the organization, and on the other hand, the

performances side, i.e. the actual results or outcome of the organization. This 'dilemma' is previously discussed in the comparable case section (Section 4.1). However, here it will be applied to the PMS for dairy specifically.

During the research it became apparent that it is difficult to measure the actual performance of an organization, if you want to keep the system generally applicable in different regions and on different types/sizes of organizations (External Interview 5,8,9,11,12,13,14; Internal Interview 3,7,8). Due to this, for the PMS for dairy the choice is made to focus on the process side of sustainability in the 'outer columns' and emphasize more on the actual performance in the *Core Themes*. (External Interview 3,9). By dividing these two parts, the system has a general, worldwide applicable part, as well as a regional, more outcomes oriented one.

Additionally, derived from observation (14), other variables or indicators that measure sustainable development of an organization, besides goals only, need to be considered. Stated by Ness, Urbel-Piirsalu, Anderberg and Olsson (2007), the assessment of goals is needed for the transition to sustainability. Though, Ness et al. (2007) argued that material flow analysis could also be an indicator for measuring sustainability performance. This could be used in and could be a good alternative for areas where goals assessment is not performed. This analysis is also divided in scale, namely in national and wider spatial levels. With material flow analysis (in)efficiency of the organization can be determined and this analysis regulates or measures the resource flows and resource use in an organization (Ness et al., 2007). Furthermore, Ness et al. (2007) state that there is a difference between project related assessment tools and policy related tools, whereas project related tools emphasize on local scale assessments and policy related assessment tools focus on a wider, global scale. So, this confirms the division made in the PMS for dairy that the global 'outer core' is focusing on policy and management and the inner '*Core Themes*' focusing on project and outcome on a more regional scale.

It is of high importance to mention that the measurements that are done through the PMS for the dairy industry are measuring relative progress, this means that the 'results' are given per year and the progress will be determined in terms of progress per year. So, an increase in a certain outcome provides insights in the progress of organizational performance towards higher sustainability (Azapagic & Perdan, 2000). Referring back to what was said in section 3.3 about the usefulness of performance measurement, Behn (2003) stated that measuring performance leads to better possibilities for evaluation. And with regard to evaluation, a comparison with a standard or other organization is desired. In the case of the PMS for dairy a comparison should be made with the organizational performance of other organizations or on beforehand determined guidelines (from DSF).

In conclusion, the balance or difference between processes versus performance is important to keep in mind by designing a PMS. However, from this study became clear that actual performance is difficult to measure. Therefore, in this study and in the designed PMS for dairy the emphasis will be on organizational performance and the process towards sustainability. However, in the *Core Themes*, there is used a more results and goal achievement related approach.

5.2.7 Questionnaire for self-assessment

After discussing the underlying principles of the system, the content of the system will be explained in this sub-section. The PMS for dairy will consist of a self-assessment questionnaire (the so called certification scheme).

The questionnaire is based on all themes and categories of the DSF, and consists of statements (and a variety of questions) with multiple answer possibilities which

represent the levels of maturity. This questionnaire provides the organization with a self-assessment through an online tool. Besides the certification scheme, a certification protocol is written, in which the outline of the scheme is explained. There is chosen for a self-assessment questionnaire due to the voluntary characteristics of DSF and referring back to Nicol and Macfarlane-Dick (2006), self-assessments have positive effects on learning and improvements within organizations.

In the questionnaire, it has been determined to use statements rather than questions only, because this makes it clearer and stronger (External Interview 10) and in most comparable cases also statements with multiple choice answers are used instead of direct questions. Generally, these statements represent the optimal situation and the answers are ascending from low (level 1) to high (level 5).

Besides the entire self-assessment questionnaire, which is comprehensive, a quickscan is developed (Appendix H), which measures in a few questions in which level the organization will be. However, this is a prediction, and the final outcome could differ from the quickscan, this should be taken into account. This quickscan will provide basic insights in what the leading parts and biggest points of improvement are. For this quickscan a variety of examples were used, to get insights in how a certain quickscan could be developed (B Corporation, 2016; Sustainable Golf Project, 2012). This quickscan contains of about 20 questions which addresses the most important issues of the PMS for dairy. However, this only provides information at a very general scale, more in-depth information about the maturity of the organization could only be given after completing the entire self-assessment. A difference is made between these two 'questionnaires', because the quickscan provides organizations with general information about their degree of sustainability which can be used internal or as a starting point for further sustainable development. However, specific steps that should be taken are not given and external verification is not possible. External verification or external promotion of the organization's sustainable practices is possible after filling in the self-assessment and having checked the given answers by external parties. Another difference between these two is that the quickscan will be a free tool and for the self-assessment organizations should pay, in ratio of organizational size, in order to get access to the questionnaire.

5.2.8 Display of results

The results of the self-assessment will be presented in an overall score, a score per category and a clear overview presented in spider diagrams, to compare the achieved results with the results of other organizations. When getting insights in the results of others and in comparison to the own results, like a benchmark, learning will be facilitated (Behn, 2003). An impression of the display of the results will be showed in the figure below (Figure 7).

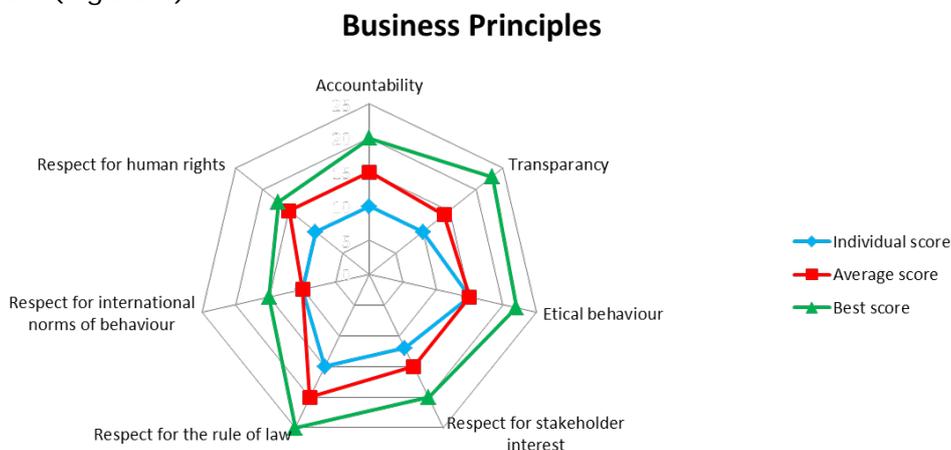


Figure 7. Artificial impression of the results of one of the columns of DSF

Coming back to the learning strategy, when the results of the self-assessment are a clear comparison between other organizations, learning takes place. According to Behn (2003), a comparison between organizations, in other words benchmarking organizations, is a traditional type of performance measurement that is designed to enable learning. Additionally, Behn (2003) describes various ways in which benchmarking will support learning; first of all, it shows the organization in what parts of the assessment they score high, and thus what they are doing well. Secondly, it provides information about in which parts the organization scores lower and thus in which parts they are not doing well. And lastly, the question rises; what does the organization needs to do to improve the parts that are not scored high enough yet? (Behn, 2003). The scores of the individual parts of the DSF within the own organization are shown in the bar chart (Figure 8).

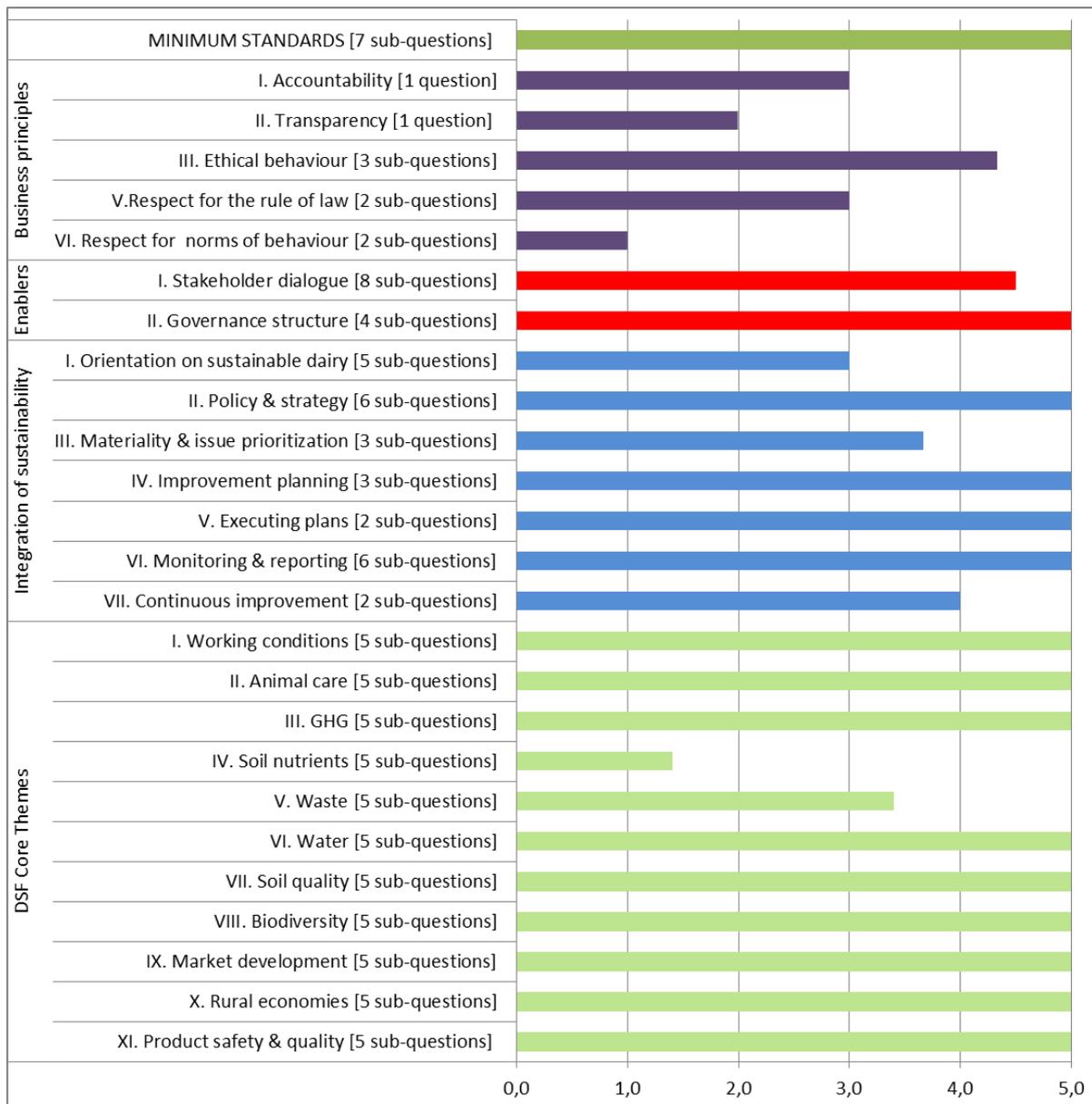


Figure 8. Bar chart that shows artificial results. The bar chart is an example of the individual score of an organization after completing the self-assessment questionnaire

In Conclusion, by applying the general knowledge and framework to the dairy industry, a few issues were important to take into account. First of all, the actual design of the system was based on knowledge derived from the first two parts of the research. This was further based on observations during the internship and in turn bringing forward the ideas and set-up designs of the PMS system to experts and advisory boards. By means of these meetings and discussions, the design of the PMS was improved. Furthermore, the focus was more on organizational performance and process performance than actual output. When going deeper into the actual design, the requirements of the entry level were important, because these should form the base level of the system, but also a certain threshold to participate in the system. So, the entry level should not be a too large barrier for stepping in the system, but should also contain of a strict starting point whether or not the organization is allowed to become a member. Due to this it is decided that from level 2, organizations have to meet the minimum standards anyway. Additionally, differences in regions and types of organizations should be taken into account, to make the system generally applicable. Thereafter, continuous improvement of the system will be done by tightening and adapting the criteria per level over time instead of adding levels further than level 5. Finally, the system will be used as a self-assessment to make it possible to test the organization's own performance and get insights in how to improve its sustainability practices.

6. Results part three: Further application to the dairy industry

After assessing the application to the dairy industry in general in the previously discussed section (Section 5.2), the system will be tested by a pilot of FrieslandCampina. This pilot, i.e. feasibility assessment, is useful to get insights in the usability of the system and to discover weaknesses or unclearness in the statements (ISEAL Alliance, 2014). Before this pilot will be discussed, the situation of FrieslandCampina will be briefly explained.

6.1 The situation of FrieslandCampina

For this research one organization is used to investigate whether the information provided from reviewing literature is applicable to organizations in the dairy industry. FrieslandCampina is explored in detail, to test and refine the results derived from literature research and to make it possible to apply the results to the dairy industry and the DSF. FrieslandCampina is chosen in this case, because it is one of the largest dairy producers and the first organization that is interested in a sustainability framework that consists of levels of maturity.

Royal FrieslandCampina NV is one of the six largest dairy companies in the world and provides consumers worldwide with dairy products containing essential nutrients. The company has offices spread over 32 countries and has more than 22,000 employees. The Central Office is situated in Amersfoort, the Netherlands. FrieslandCampina is owned by Zuivelcoöperatie FrieslandCampina U.A. which consists of over 19,000 member dairy farmers in the Netherlands, Germany and Belgium (FrieslandCampina, n.d.²).

FrieslandCampina's goal is to deliver better nutrition for the world and a fair income for its farmers now and in the future (FrieslandCampina, n.d.³). To achieve this goal, FrieslandCampina set up the Route2020 strategy⁹. This sustainability strategy also includes the DSF (section 1.3), in which stakeholders are brought together and collaborate (FrieslandCampina, n.d.³).

In 2014 FrieslandCampina, together with two key customers, started an implementation pilot project. By making use of the Dairy Sustainability Framework (DSF), FrieslandCampina wants to demonstrate the increasing sustainability of dairy farming to customers and consumers (FrieslandCampina, n.d.¹) with use of the DSF Demonstration Scheme (revisiting Figure 4 at page 45). Once more, this scheme demonstrates different steps an organization should take in order to function in a sustainable way. Besides these steps, business principles and core themes regarding sustainability are included. Furthermore, enablers and already existing standards, that influence the sustainability management and performance, are included in the scheme as well. The scope of the DSF is the entire dairy value chain (DSF, 2015).

6.2 Pilot findings

After completing the PMS for the dairy industry, the system was tested by using the self-assessment questionnaire, which was completed by a sustainability expert in the context of FrieslandCampina. In the so called first pilot, i.e. feasibility assessment, the entire system, with all statements and questions, was included in order to discover whether it was possible to answer the questions and statements. In this way, the organization can evaluate the system and can test its feasibility (ISEAL Alliance, 2014). As a result, errors and ambiguities in the system can easily be detected and could be solved and improved before the system is actually put in place or used in a bigger pilot.

⁹ The main principles within the route2020 strategy are sustainable growth of the firm and value creation of the produced milk.

Testing the system was useful, because it determined whether the system worked as intended, gave the possibility to identify errors (Bertolino, 2007) and indicated whether the system was user friendly for the organization FrieslandCampina. This is important to test before the actual system is in place, because the main errors and unclear questions and statements should be eliminated by then. Besides testing the statements, an extensive comparison with ISO 26000 was carried out. This was to determine whether the system could be linked to, overlap or discourage parts of the ISO standard. This can be seen as a gap analysis, to investigate whether the PMS for dairy is comparable to the ISO standard.

After performing the pilot and the gap analysis, the main strengths and weaknesses were discovered and weaknesses were diminished. These strengths and weaknesses were discussed during feedback sessions of the pilot, which are referred to as observations 15 and 17.

The main strengths demonstrated by the pilot were that the PMS for dairy is an extensive system and contains all relevant elements for measuring sustainability. Another important finding from the gap analysis was that the PMS is more detailed in comparison to ISO 26000 and therefore the PMS goes a step beyond the ISO standard (Observation 15). All statements of the PMS can be linked to one of the statements of ISO. The *Business Principles* statements within the PMS are literally based on the ISO principles. But from the ISO 26000 point of view, not all questions and statements can be linked to the PMS for dairy, because these statements are not relevant for the dairy industry and therefore not included in the PMS. Questions that are not relevant for the dairy industry are questions that are too general. An example is *'In which social responsibility initiatives and possible matching instruments does your organization partake?'* This question is too broadly defined and therefore not applicable for DSF and the PMS for dairy (Observation 15).

Another advantage of the PMS for dairy, and a shortcoming of ISO 26000, is the continuous improvement of the PMS and the support of continuous improvement of sustainability instead of the more static approach of the ISO standard. In other words, the PMS for dairy related to DSF has a plan-do-check-act (PDCA) cycle to support the organizations to become higher on their sustainability practices and keeps the system up to date. For the ISO 26000 this approach is not applicable (Observation 15). Thus, the PDCA cycle of the PMS is seen as a big advantage. This PDCA cycle is i.a. associated with the statements covered by the *Integration of Sustainable Dairy* part (Blue column at the top of the scheme, Figure 4). An example of ISO 26000 questions which have a more static approach are *'Which subjects are relevant?'*, *'Which criteria did you use in determining the issues of high priority?'* and *'Which issues have priority?'* (Observation 15). These questions do not address the process or development over time, but only consider one moment in time. On the contrary, an example of a statement in the PMS which addresses the process towards determining issue prioritization is *'The organization has established and implemented an objective process to determine the materiality of the 11 themes that reflects the significant economic, environmental and social issues or aspects that substantively influence the decisions of stakeholders.'*

Besides these strengths, some weaknesses were discovered. The main weaknesses of the system detected during the pilot were ambiguities or overlaps in the statements (Observation 15,17). The statements must be clearly understandable for all members and participants of the DSF and therefore an accurate review of the statements was useful. This was executed during the pilot and a few statements were adjusted, in consultation with sustainability experts, to make them clearer and more easily understandable. A few statements still had too much overlap with other statements, and therefore it was decided to delete and/or combine some of them. This was done in order to keep the system efficient (Observation 15,17). A few examples of these statements

can be given to clarify this. Firstly, a statement covered by the *Integration of Sustainable Dairy* part was that the organization should have determined its resources needed to come to an effective execution of the sustainability policy. However, the term 'resources' was too broad and had to be clarified. This was done by adding a number of examples. The statement is adapted by: *The organization has determined resources (in terms of money, time, people) needed to come to an effective execution of the sustainability policy.* Another example of a statement which has to be adjusted was within the *Business Principles* part and sub element 'Respect for human rights'. This statement is related to violation and is as follows: *The organization has established and implemented a process to deal with violation of human rights.* However, in the answer possibilities, violation was too narrow formulated as 'discrimination'. Discrimination does not cover all elements of violation and in order to make the answers more generally applicable, this is changed in violation.

Additionally, from the feasibility assessment it became clear which departments of FrieslandCampina are related to which parts of the PMS for sustainable dairy and in turn which departments were able to answer specific parts of the self-assessment. This is especially important for the specific *Core Themes* of DSF, because this part contains issue specific statements. After the pilot, a few examples can be given. The 'Rural Economies' theme can be strongly connected to the Dairy Development Programme of FrieslandCampina, in which they support the development of dairy farms and raising the standards of living of dairy farmers in Asia, Africa and Eastern Europe (FrieslandCampina, n.d.⁴). The specific statements which are covered by the 'Product Safety & Quality' theme were linked to the Foqus Food Safety & Quality department of FrieslandCampina, so these statements could be answered by them (FrieslandCampina, n.d.⁵). Another example of a theme that strongly related to one of the programs that FrieslandCampina has in place was 'Animal Care'. Within FrieslandCampina this theme is covered by the Foqus Planet program, and therefore these statements could be answered by this department (FrieslandCampina, n.d.⁶).

6.3 Towards an intermediate conclusion

From the pilot can be concluded that the overall usability of the system is seen as positive and the system contains all relevant elements for measuring sustainability in the dairy sector. Furthermore, the system goes beyond the ISO standard, because the PMS is more specifically related to the dairy sector and therefore more useful. On the contrary, a number of statements still had too much overlap and have therefore been combined, deleted or adapted. Additionally, the pilot gave insights in which parts of the PMS can be answered by whom.

Further steps that have to be taken in order to implement the system and increase the credibility will be addressed in the discussion and recommendations section of this thesis.

7. Discussion

This research was executed in order to explore what a PMS is, how to characterize such a system and how to design a PMS with levels of maturity to a specific industry, in this case the dairy industry. This research was executed during an eight-month internship period and this gave extra insights, besides desk research and empirical research. The methods used in this research were not regarded as a problem. Literature research gave enough opportunities to develop a general framework, due to the large variety of reports written about PMS and/or levels of maturity, and with the knowledge derived from the empirical research, the framework could be strengthened and complemented. Also, data collection within the empirical research part of the study was not seen as a problem. The 26 interviewees were willing to provide me with information and were easily accessible.

However, there are some limitations of the research. First of all, the comparable cases that were used in this study provided insights and useful knowledge about how to set up a maturity system to measure sustainability. However, a few of these systems were designed for a specific industry, the construction industry (CO₂PL, ProRail Veiligheidsladder), which differs a lot from the dairy industry. The most important difference between these industries was the non-international character of the construction industry versus the international character of the dairy industry. Furthermore, some of the example systems did not make use of maturity levels (CDP, DJSI), but were only a benchmark or scoring system. Nevertheless, they gave general, useful information about how sustainability related systems work and how they are applied to international markets. In order to solve the limitations in further research, comparable cases which are international, include levels of maturity and are related to the food industry could be used. However, it is difficult to find these kinds of systems, because there are not that many with these specific characteristics. Additionally, since there are not yet many systems that measure sustainability in levels of maturity in one specific industry, the designed PMS for dairy is valuable. It is a contribution to science and society in the field of sustainable development and more specifically in sustainable food production and consumption.

Another limitation is, that the actual design of the PMS for dairy is not yet tested in a wider perspective. Interested organizations and members of DSF have made a contribution to the design process of the PMS. They have provided me with extra notifications. However, for further research and implementation of the system, a number of pilots, or feasibility assessments, are needed to test the design of the PMS for dairy. The design is internally tested in a pilot at FrieslandCampina. This gave insights in the usability and applicability of the system, but this has to be expanded to other organizations in the dairy industry as well. Besides this pilot, public consultation and extensive stakeholder engagement is required, to give stakeholders enough time and opportunities to provide input on the PMS (ISEAL Alliance, 2014).

It appeared from the research that concrete results and outcomes are difficult to measure in a PMS if it has to be generally applicable on a global scale. Actually, this was not expected in advance, but it became clear during the research. However, process performance is eventually an indicator of outcome, because when the processes towards the goals are in development, the outcome will be improving as well. Measuring outcomes is possible if the system is designed for one specific region or type of organization. This is not the case for the dairy industry, which asks for a global applicable system in order to measure sustainability in the entire supply chain.

In conclusion, for further research the use of more specific comparable cases is recommended. In that way, more specific information about how these systems work will be obtained. In addition, to further implement the PMS for dairy, a second pilot study, i.e. feasibility assessment, is needed to test the applicability and usability of the system.

An international, diverse test panel of organizations is recommended as well as enough opportunities for stakeholders to give their input on the system.

8. Conclusions

The main goal of this thesis is to investigate how to build a Corporate Sustainability Performance Measurement System (PMS) and with this knowledge, how to apply this type of system to the dairy industry. For this thesis the research was subdivided in three parts, first a general framework from literature research was set up. This general framework gave insights in the requirements a certain system should comply with and what the characteristics per maturity level should be. After executing this first, general part, a second study was executed which consists of empirical research, in which comparable cases were analyzed in more detail and interviews were conducted. These analyzes gave insights in how a PMS is working in existing cases and what the differences and/or similarities are between literature and 'reality'. In order to get detailed information, 26 in-depth interviews were conducted. Thirdly, the knowledge obtained from the first two parts was used to apply on the dairy industry. For this application one specific organization is used, namely FrieslandCampina, one of the largest dairy producers in the world. This organization is used in order to build a PMS with levels of maturity on the specific framework that was already in place in the dairy industry, the Dairy Sustainability Framework (DSF). Although this system already exists, levels of maturity were missing and therefore this research was executed.

Once again, this thesis is executed to answer the following questions:

- *What are the key characteristics of a Performance Measurement System for the continuous improvement of Corporate Sustainability?*
- *How can this information be used for the design of a Corporate Sustainability Performance Measurement System for the dairy industry?*

The requirements for a PMS were derived from desk research and expanded with information derived from interviews and comparable cases. A basic requirement that is needed for a PMS to measure corporate sustainability is that the system is related to sustainability (and covers the three dimensions of sustainability; people, planet, profit) and that it supports continuous improvement of sustainable practices on the long term (Cagnin et al., 2005; ISEAL Alliance, 2014; Kirkwood et al., 2011; Lockamy & McCormack, 2004; Mayer, 2008; Folan & Browne, 2005). Additionally, the criteria within the system should be relevant and measurable in order to determine sustainability performance (ISEAL Alliance, 2014; Munda, 2005; Folan & Browne, 2005). The system should also be applicable to a variety of different regions and therefore the regional differences should be taken into account when designing the PMS. So, in order to get a well-working system the principle of subsidiarity of Potts et al. (2014) should be considered. By making it local applicable, the system is approachable to the needs and interests of stakeholders in multiple regions (Potts et al., 2014).

Besides, the system should continuously be kept up to date due to the fast changes made in sustainability. In order to go along with these changes, the criteria per level need to be tightened over time (Bourne et al., 2000; Cagnin et al., 2005; Potts et al., 2014; Vermeulen & Witjes, 2016; External Interview 3,4,6,7,9,10,13,14,15; Internal Interview 2,4,7,8). Related to this, a clear feedback mechanism will help to improve the system (Folan & Browne, 2005; External Interview 9,10,15; Vermeulen & Witjes, 2016). Also, for the design of the PMS, it is important that each advanced level includes the aspects of the previous maturity levels, in order to prevent that organizations can skip certain levels or criteria (Cagnin et al., 2005). Besides, the system should challenge frontrunners and at the same time should not discourage organizations who are just at the beginning of sustainability (Bourne et al., 2000).

Secondly, desk research showed that a PMS with levels of maturity usually consists of 4 or 5 levels and in this study there is chosen to use 5 levels. These 5 levels are called respectively; *Initial*, *Recognized*, *Defined*, *Managed* and *Optimized*. Additionally, a few trends were recognized; first of all, the movement from an internal, reactive approach towards an external, proactive approach (Cagnin et al., 2005; Kirkwood et al., 2011). So, in the first two or three levels the organizations focus on their own organization and internal activities (Baumgartner & Ebner, 2010; Cagnin et al., 2005; Lockamy & McCormack, 2004; Van Marrewijk & Werre, 2003) and in the higher levels organizations actively engaging with other organizations or parties, i.e. the leading organizations take a proactive role in stakeholder engagement (Cagnin et al., 2005; Kirkwood et al., 2011; Maak, 2007; McCormack et al., 2008; Lockamy & McCormack, 2004). This is also the case for communication processes in and outside the organization (Cagnin et al., 2005). The importance of communication is also confirmed by the findings from empirical research (External Interview 5,9,16; Internal Interview, 1,5,7,8). In relation to this, the transition from inactive, reactive towards active and proactive is often seen in the sustainability transition of large organizations (Van Tulder et al., 2014).

According to empirical findings, the creation of internal support is very important to reach higher levels of sustainability. Internal support could be linked to communication, because communication is an important driver to create internal support amongst employees and in turn continuous improvement of sustainability (External Interview 5,9,16; Internal Interview 1,5,7,8). In relation to this, not only internal support is important, but also external support and thus stakeholder engagement is crucial in order to become higher on the sustainability ladder (External Interview 2,6,15). Once more, the perspective that is seen in the levels is that organizations move from an internal towards an external approach. So, this is both from internal towards external communication and reporting, as well as the focus on own organization versus cooperation with stakeholders in the supply chain (External Interview 6,7,9,13,14,16; Internal Interview 5,6). Additionally, engagement with society, besides engagement with direct stakeholders only, is important if the organization wishes to achieve the highest levels of maturity (External Interview 16).

Consequently, and related to stakeholder commitment, is transparency, which is seen as an important factor. To gain trust of the stakeholders, the organization should be transparent about its processes towards sustainability and this feature is seen as a crucial one to score high on the maturity system (External Interviews 4,5,12; Internal Interviews 5,6). When passing the levels, data collection, data measuring and goals achievement increase and improve (Kirkwood et al., 2011; McCormack et al., 2008; Lockamy & McCormack, 2004). In the lower levels data collection and measuring is nihil and the focus is more on the processes. In the higher levels, advanced measurement and documentation are used, as well as the achievement of goals (Cagnin et al., 2005; Lockamy & McCormack, 2004; McCormack et al., 2008).

Finally, the knowledge derived from the first parts of the research was applied to the dairy industry, and more specifically to the Dairy Sustainability Framework (DSF). During the design of the PMS for dairy, a few important things came across. First of all, the emphasis of the system is on the processes and organizational performance of organizations. This is because actual output is difficult to measure on a global scale when taking into account regional- and organizational differences. However, in order to be able to include results into the system to a certain degree, the PMS for dairy is divided in two parts. On one hand, the outer columns of the DSF are globally applicable and focus on processes, and on the other hand, the *Core Themes* of the DSF are emphasizing more on output and have a regional applicability. Furthermore, the requirements for the entry level of the PMS for dairy were important to keep in mind, because these form the base level of the system as well as a certain threshold to participate in the PMS. Lastly, the system will be used as a self-assessment to make it possible for organizations to test their own performance and get insights in how to improve their sustainability practices.

9. Recommendations to business

Attached to this research was the eight-month internship at the Central Office of FrieslandCampina in Amersfoort. After completing the research a few advices could be given to the business.

The research conducted here is a scientific based design process for a PMS for the dairy industry and therefore, the recommendations to the business are related to next steps that have to be taken in order to successfully implement the system in the entire DSF network.

- The first step that has to be taken in order to implement the PMS for dairy is setting up a broader pilot study to test the usability of the system. So, before the actual implementation in the DSF network can take place, an international, wider scaled pilot is needed to see whether the system is applicable for more types of organizations than FrieslandCampina only. This could also be linked to ISEAL Alliance (2014). First of all, the testing pilot, by ISEAL Alliance referred to as feasibility assessment, is part of the system development through testing the feasibility and auditability of the requirements in the system (ISEAL Alliance, 2014). Secondly, ISEAL Alliance prescribes that for the successful implementation of a new system, public consultation is important. By means of this step, stakeholders have the opportunity to provide input on the system and can get insights in which manner their input is taken into account (ISEAL Alliance, 2014). The first step I suggest is to start a pilot, i.e. feasibility assessment, with two or three other large dairy companies in the North Western part of Europe. The pilot should be taking place in North West Europe, because the system is built with this situation in mind. When these organizations are able to make proper use of the system, it can be extended to other parts of the world. By starting a testing pilot, the largest weaknesses of the system will be discovered and the system can be improved. Thus, the pilot will provide participants with information about the usability and functioning of the system. After this pilot, the content of the system should be available for other stakeholders as well, in order to start the public consultation and give other stakeholders and organizations the opportunity to provide input (ISEAL Alliance, 2014).
- From interviews it became clear that sufficient communication is required in order to create more internal support. Therefore, the next step that has to be taken is, to create more internal support amongst DSF members and its secretariat. During the process it became apparent that the internal support of a maturity level system should be increased in order to successfully implement the PMS. In order to create this support, it is recommended to set up a strong communication network about the advantages of the PMS for dairy. This could be done during or after the pilot with the participating organizations of the pilot. When they share their experiences of filling in the self-assessment and the useful insights they received from the system, the other members and the DSF secretariat could be convinced more easily.
- Another recommendation that will be given is more related to the content of the system. I recommend to expand the regional classification or to make it more specific. Although, the new classification (sub-section 5.2.3 and Table 16) is more specific compared to the classification of SustainAbility, the regions still have to be sub-divided into smaller areas within the 14 regions in order to be able to make a more detailed and fair comparison between organizations (External Interview 4; Internal Interview 9, Observation 6). However, this more specific sub-division could be linked to a pilot, to see whether some regions need more specification, whether certain part and/or themes are not applicable or relevant for a specific region and how or if the scoring should be adjusted for these variations. Thus, by making use of

a pilot, insights can be obtained how to implement the *Core Themes* on regional specific situations and how the scoring can be adjusted to this.

- Furthermore, the continuous improvement of the system could be done in several ways to keep the system up-to-date and relevant to the current sustainability norms and progression. I will recommend to appoint a commission who reviews the system once a year and participants of the system should have the opportunity to give feedback. Therefore, a few feedback questions regarding the main strengths and weaknesses could be included in the self-assessment tool. By doing this, members are able to improve the system by themselves and together with the other participants. Related to continuous improvement of the system is the continuous improvement of the participating organizations. The system has to support organizations to improve their sustainability practices and therefore, certain transitional periods are necessary (derived from comparable cases). Due to the more easily reachable criteria of the first two levels, the transition period should be no longer than one year. So, the organizations have one year to improve themselves and reach a higher level of maturity. Moreover, the higher levels are more difficult to reach and therefore the transitional period has to be longer. So, from level three on the transitional period needs to be two or three years. Eventually, the achievement of the fifth level is the choice of the organization, whether they have the ambition to become a frontrunner (Internal Interview 1), thus for level four to five a transitional period is not necessary.
- Lastly, Minimal levels and/or minimal requirements should be formulated more clearly. The threshold of the minimum level is strict and clear, however the actual requirements ask for more explanation. Currently, it is not clear what the actual requirement, law or regulation contains of. In order to provide more clarity, the DSF could come up with a detailed list of the required laws to which an organization should comply. This is important because the entry level and minimum requirements are seen as an important part of the DSF and the maturity levels. So, currently one of the requirements is 'compliance with local laws and regulations'. In order to make this more concrete, DSF could come up with a list of the local laws and regulations, clustered per region.

In conclusion, the set-up of an independent, international testing pilot, i.e. feasibility assessment, on North West European scale is the first step in order to implement the system successfully. In order to improve the objectivity of the system, this second pilot should be carried out by an independent organization. Thereafter, the internal support amongst the DSF secretary should be increased and this could be achieved by setting up a more extensive communication network and public consultation opportunities, after or during the international pilot is tested. Additionally, the regional variation regarding the *Core Themes* as well as the local laws and regulations, need some more attention and a general agreement for the continuous improvement of the system should be arranged.

References

- AccountAbility (1999). AA1000 Framework overview, standards and guidance notes. Retrieved May 30, 2016 from <http://www.accountability.org/images/content/0/7/076/AA1000%20Overview.pdf>
- Azapagic, A. (2003). Systems approach to corporate sustainability: a general management framework. *Process Safety and Environmental Protection*, 81(5), 303-316.
- Azapagic, A., & Perdan, S. (2000). Indicators of sustainable development for industry: a general framework. *Process Safety and Environmental Protection*, 78(4), 243-261.
- Ashurst, C., & Doherty, N. F. (2003). Towards the formulation of a 'best practice' framework for benefits realisation in IT projects. *Electronic Journal of Information Systems Evaluation*, 6(2), 1-10.
- Baumgartner, R. J., & Ebner, D. (2010). Corporate sustainability strategies: sustainability profiles and maturity levels. *Sustainable Development*, 18(2), 76-89.
- B Corporation (2016). Performance requirements, take the B impact assessment. Retrieved March 21, 2016 from <https://www.bcorporation.net/become-a-b-corp/how-to-become-a-b-corp/performance-requirements>
- Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public administration review*, 63(5), 586-606.
- Bertolino, A. (2007). Software testing research: achievements, challenges, dreams. *Future of Software Engineering*, 85-103
- Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. *International journal of operations & production management*, 20(7), 754-771.
- Busch, L. (2011). *Standards: Recipes for reality*. Michigan, United states: MIT Press.
- Cacioppe, R., & Edwards, M. (2005). Seeking the Holy Grail of organisational development: A synthesis of integral theory, spiral dynamics, corporate transformation and action inquiry. *Leadership & Organization Development Journal*, 26(2), 86-105.
- Cagnin, C. H., Loveridge, D., & Butler, J. (2005). Business sustainability maturity model. In *Business Strategy and the Environment Conference, Devonshire Hall, University of Leeds*.
- Castka, P., & Balzarova, M. A. (2008). ISO 26000 and supply chains—On the diffusion of the social responsibility standard. *International journal of production economics*, 111(2), 274-286.
- CDP (2015¹). Guidance for companies reporting on water on behalf of investors & supply chain members 2015, Water questionnaire 2015.
- CDP (2015²). CDP 2015 climate change scoring methodology. Retrieved March 20, 2016 from <https://www.cdp.net/Documents/Guidance/2015/CDP-climate-change-scoring-methodology.pdf>
- CDP (n.d.). The facts. Retrieved April 13, 2016 from <https://www.cdp.net/Documents/CDP-the-facts.pdf>
- Charles, H., Godfray, H. C., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*, 327, 812-818.
- Cramer, J. (2005). Experiences with structuring corporate social responsibility in Dutch industry. *Journal of cleaner production*, 13(6), 583-592.

- De Jong, A., DeJong, D. V., Mertens, G., & Wasley, C. E. (2005). The role of self-regulation in corporate governance: Evidence and implications from the Netherlands. *Journal of corporate finance*, 11(3), 473-503.
- Dorée, A, van der Wal, G and Boes, H (2011) Client leadership in sustainability: How the Dutch railway agency created CO₂ awareness in the industry *In: Egbu, C. and Lou, E.C.W. (Eds.) Procs 27th Annual ARCOM Conference*, 5-7 September 2011, Bristol, UK, Association of Researchers in Construction Management, 685-694.
- DSF (2014). The Dairy Sustainability Framework. Retrieved November 18, 2015, from <http://dairysustainabilityframework.org/the-gdaa/programmes-of-the-gdaa/the-dairy-sustainability-framework/>
- DSF (2015). The Dairy Sustainability Framework, implementation guide 2015. Retrieved December 15, 2015 from <http://dairysustainabilityframework.org/wp-content/uploads/2015/10/Dairy-Sustainability-Framework-Implementation-Guide.pdf>
- Elkington, J. (1997). *Cannibals with forks: the triple bottom line of twenty first century business*. Capstone, Mankato, MN.
- Elkington, J. (2004). Enter the triple bottom line. *The triple bottom line: Does it all add up*, 11(12), 1-16.
- Epstein, M. J., & Roy, M. J. (2001). Sustainability in action: Identifying and measuring the key performance drivers. *Long range planning*, 34(5), 585-604.
- FAO (2015). The state of food and agriculture 2015, social protection and agriculture: breaking the cycle of rural poverty. Retrieved December 12, 2015 from <http://www.fao.org/3/a-i4910e.pdf>
- 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.
- Folan, P., & Browne, J. (2005). A review of performance measurement: Towards performance management. *Computers in industry*, 56(7), 663-680.
- FrieslandCampina (n.d.¹). Sustainable dairy farming. Retrieved November 10, 2015 from <https://www.frieslandcampina.com/en/sustainability/sustainable-dairy-farming/>
- FrieslandCampina (n.d.²). Organization; our products find their way to more than 100 countries. Retrieved November 23, 2015 from <https://www.frieslandcampina.com/en/66rganization/who-we-are/>
- FrieslandCampina (n.d.³). Organization; strategy, route2020. Retrieved November 23, 2015 from <https://www.frieslandcampina.com/en/66rganization/strategy-route2020/>
- FrieslandCampina (n.d.⁴). Dairy Development Programme. Retrieved June 28, 2016 from <https://www.frieslandcampina.com/en/sustainability/dairy-development-programme/>
- FrieslandCampina (n.d.⁵). Foqus Food Safety & Quality: quality & safety during production. Retrieved June 28, 2016 from <https://www.frieslandcampina.com/en/quality-and-safety/foqus/quality-safety-during-production-and-distribution/>
- FrieslandCampina (n.d.⁶). Kwaliteit begint op de boerderij, met 'Foqus Planet'. Retrieved June 28, 2016 from <https://www.frieslandcampina.com/nl/kwaliteit-en-veiligheid/foqus/foqus-planet/>
- GDA (2011). The Global Dairy Agenda for Action on Climate Change, progress report on dairy sector commitments 2009-2011. Retrieved November 20, 2015 from <http://dairysustainabilityframework.org/wp-content/uploads/2014/08/GDAaprogress2011report.pdf>
- Ghobadian, A., & Galleary, D. (1997). TQM and organization size. *International journal of operations & production management*, 17(2), 121-163.

- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The qualitative report*, 8(4), 597-606.
- Gomes, J., Romão, M., & Caldeira, M. (2013). The Benefits Management and Balanced Scorecard Strategy Map: How They Match. *International Journal of IT/Business Alignment and Governance (IJITBAG)*, 4(1), 44-54.
- GRI (2015). G4 Sustainability reporting guidelines, reporting principles and standard disclosure. Retrieved March 20, 2016 from <https://www.globalreporting.org/resource/library/GRI-G4-Part1-Reporting-Principles-and-Standard-Disclosures.pdf>
- Harris, K.J. (2000). *Integrated justice information systems, governance structures, roles and responsibilities. A background report*. Retrieved February 14, 2016 from www.search.org/files/pdf/Governance.pdf
- Hearts & Minds (2016). About Hearts & Minds. Retrieved April 12, 2016 from <http://heartsandminds.energyinst.org/About>
- Huan, S. H., Sheoran, S. K., & Wang, G. (2004). A review and analysis of supply chain operations reference (SCOR) model. *Supply Chain Management: An International Journal*, 9(1), 23-29.
- Huang, S. H., Sheoran, S. K., & Keskar, H. (2005). Computer-assisted supply chain configuration based on supply chain operations reference (SCOR) model. *Computers & Industrial Engineering*, 48(2), 377-394.
- IFCN (2015). World Dairy Map 2015, results of the IFCN Dairy Report 2014. Retrieved April 25, 2016 from https://www.verantwoordeveehouderij.nl/upload_mm/7/e/e/5453701c-d838-46fc-9bfa-278ce11af84a_Dairy-Map-2015_final.pdf
- ISEAL Alliance (2014). Setting Social and Environmental Standards, ISEAL Code of Good Practice. Version 6.0, December 2014. Retrieved May 20, 2015 from <http://www.isealalliance.org/sites/default/files/ISEAL%20Standard%20Setting%20Code%20v6%20Dec%202014.pdf>
- ISO (n.d.). ISO 26000, social responsibility. Retrieved November 30, 2015 from <http://www.iso.org/iso/home/standards/iso26000.htm>
- ISO 26000 Scan (n.d.). MVO is nu certificeerbaar. Retrieved April 12, 2016 from <http://www.iso26000scan.nl/home/mvo-prestatieladder>
- Kaplan, R. S., & Norton, D. P. (1995). Putting the balanced scorecard to work. *Performance measurement, management, and appraisal sourcebook*, 66, 17511.
- Kaplan, R. S., & Norton, D. P. (2000). Having trouble with your strategy? Then map it. *Harvard Business Review*, 78(5), 167-276.
- Kaptein, M., & Van Tulder, R. (2003). Toward effective stakeholder dialogue. *Business and society review*, 108(2), 203-224.
- Kim, E. H., & Lyon, T. (2011). When does institutional investor activism increase shareholder value? The carbon disclosure project. *The BE Journal of Economic Analysis & Policy*, 11(1), 1-27.
- Kirkwood, D. A., Alinaghian, L. S., & Srari, J. S. (2011). A maturity model for the strategic design of sustainable supply networks. *Management*, 28(5).
- Kotler, P., & Lee, N. (2005). Corporate social responsibility. *Doing the Most Good for Your Company and Your Cause*, New Jersey.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3), 259-263.

- Knoepfel, I. (2001). Dow Jones Sustainability Group Index: a global benchmark for corporate sustainability. *Corporate Environmental Strategy*, 8(1), 6-15.
- Lasswell, H. D. (1948). The structure and function of communication in society. *The communication of ideas*, 37, 215-228.
- Lockamy III, A., & McCormack, K. (2004). The development of a supply chain management process maturity model using the concepts of business process orientation. *Supply Chain Management: An International Journal*, 9(4), 272-278.
- Maak, T. (2007). Responsible leadership, stakeholder engagement, and the emergence of social capital. *Journal of Business Ethics*, 74(4), 329-343.
- Matson, P.A., Parton, W.J., Power, A.G., & Swift, M.J. (1997). Agricultural intensification and ecosystem properties. *Science*, 277(5325), 504-509.
- Mayer, A. L. (2008). Strengths and weaknesses of common sustainability indices for multidimensional systems. *Environment international*, 34(2), 277-291.
- McCormack, K., Bronzo Ladeira, M., & Paulo Valadares de Oliveira, M. (2008). Supply chain maturity and performance in Brazil. *Supply Chain Management: An International Journal*, 13(4), 272-282.
- Moen, R., & Norman, C. (2006). Evolution of the PDCA cycle (Detroit). Retrieved May 25, 2016 from http://pkpinc.com/files/NA01_Moen_Norman_fullpaper.pdf
- Munda, G. (2005). "Measuring sustainability": a multi-criterion framework. *Environment, Development and Sustainability*, 7(1), 117-134.
- MVO (2013)¹. De MVO Prestatieladder, eisen MVO managementsysteem voor maatschappelijk verantwoord ondernemen. Deel A, Versie 1 december 2013/1. Retrieved November 12, 2015 from http://mvoprestatieladder.nl/wp-content/uploads/2016/04/2013MVOPrestatieladderdeelAManagementsysteem-versie1-12-2013_1FinalvermeldinginBijlage7-websiteFSR.pdf
- MVO (2013)². De MVO Prestatieladder, certificatieregeling voor het MVO managementsysteem. Deel B, Versie 1 december 2013. Retrieved November 12, 2015 from <http://mvoprestatieladder.nl/wp-content/uploads/2016/05/2013MVOPrestatieladderdeelB.pdf>
- Nadvi, K., & Wältring, F. (2001). Making sense of global standards. University of Sussex UK, University of Duisburg, Germany.
- Natuur & Milieu (2015). Situatie in de melkveehouderij, december 2015. Retrieved April 25, 2016 from <https://www.natuurenmilieu.nl/wp-content/uploads/2016/01/20160107-Situatie-in-de-melkveehouderij-December-2015-1.pdf>
- 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., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M., & Kennerley, M. (2000). Performance measurement system design: developing and testing a process-based approach. *International Journal of Operations & Production Management*, 20(10), 1119-1145.
- Ness, B., Urbel-Piirsalu, E., Anderberg, S., & Olsson, L. (2007). Categorising tools for sustainability assessment. *Ecological economics*, 60(3), 498-508.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, 31(2), 199-218.
- Parker, D., Lawrie, M., & Hudson, P. (2006). A framework for understanding the development of organisational safety culture. *Safety Science*, 44(6), 551-562.

- Paulk, M. C., Weber, C. V., Garcia, S. M., Chrissis, M. B. C., & Bush, M. (1993). Key practices of the Capability Maturity Model version 1.1. *Tech. Report 93(25)*.
- Porter, M. E., & Kramer, M. R. (2011). The big idea: Creating shared value. *Harvard Business Review*, 89(1), 2.
- Potts, J., Lynch, M., Wilkings, A., Huppe, G., Cunningham, M., & Voora, V. (2014). The state of sustainability initiatives review 2014: Standards and the green economy. *International Institute for Sustainable Development (IISD) and the International Institute for Environment and Development (IIED)*, 332.
- ProRail (2015). Handboek Veiligheidsladder, Samen zorgen voor bewust veilig werken. Versie 2.0 definitief. ProRail Procurement / Kwaliteit en Onderzoek. Retrieved November 20, 2015 from https://www.prorail.nl/sites/default/files/edms-3777264-v2-handboek_veiligheidsladder_2_0.pdf
- Reed, R. (2012). The CSR Performance Ladder Certification Standard. *The Executive Times*. 13-15.
- Rietbergen, M. G., & Blok, K. (2013). Assessing the potential impact of the CO₂ Performance Ladder on the reduction of carbon dioxide emissions in the Netherlands. *Journal of Cleaner Production*, 52, 33-45.
- RobecoSAM (2014). DJSI 2014, review results, september 2014. Retrieved May 24, 2016 from http://www.sustainability-indices.com/images/DJSI_Review_Presentation_09_2014_final.pdf
- RobecoSAM (2015)¹. Dow Jones Sustainability Index, DJSI family. Retrieved April 11, 2015 from <http://www.sustainability-indices.com/index-family-overview/djsi-family-overview/index.jsp>
- RobecoSAM (2015)². DJSI 2015, test companies SA, corporate sustainability assessment, questionnaire. Retrieved April 3, 2016 from <http://www.robecosam.com/images/sample-questionnaire-1.pdf>
- Roland Berger Strategy Consultants, Nederlandse Zuivel Organisatie, & ZuivelNL. (2015). *De Witte Motor*. Den Haag.
- Searcy, C. (2012). Corporate sustainability performance measurement systems: a review and research agenda. *Journal of business ethics*, 107(3), 239-253.
- Sharma, M. K., Bhagwat, R., & Dangayach, G. S. (2005). Practice of performance measurement: experience from Indian SMEs. *International Journal of Globalisation and Small Business*, 1(2), 183-213.
- Stichting Klimaatvriendelijk Aanbesteden & Ondernemen (2015). Handboek CO₂-Prestatieladder 3.0. Retrieved November 15, 2015 from http://cms2009.digitnet.nl/Uploads/CO/20150610_Handboek_CO_2_Prestatieladder_3_0.pdf
- Stichting Milieukeur (2016). Kompas voor duurzamer ondernemen, certificering, voordelen. Retrieved April 11, 2016 from <http://www.smk.nl/7/over-smk/certificering--voordelen.html>
- Steering Committee of the State-of-Knowledge Assessment of Standards and Certification. (2012). Executive Summary. In: *Toward sustainability: The roles and limitations of certification*. 1-18.
- SustainAbility (2013). Rate the raters phase five, questionnaire for raters, Dow Jones Sustainability Indices. February 2013. Retrieved May 12, 2016 from www.sustainability.com/library/attachment/426
- Sustainable Golf Project (2012). Do the quickscan! How sustainable is your golf course? Retrieved March 22, 2016 from <http://www.sustainablegolfproject.com/61/do-the-quickscan.htm>
- Van Calker, K. J., Berentsen, P. B., Giesen, G. W., & Huirne, R. B. (2005). Identifying and ranking attributes that determine sustainability in Dutch dairy farming. *Agriculture and Human Values*, 22(1), 53-63.

- Van Marrewijk, M., & Werre, M. (2003). Multiple levels of corporate sustainability. *Journal of Business Ethics*, 44(2-3), 107-119.
- Van Tulder, R., Tilburg, R. van., Francken, M. & Da Rosa, A. (2014). *Managing the transition to a sustainable enterprise. Lessons from frontrunner companies*. Oxfordshire, UK: Routledge
- Welsh, E. (2002). Dealing with data: using NVivo in the qualitative data analysis process. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 3(2), Art. 26. Retrieved from <http://nbn-resolving.de/urn:nbn:de:0114-fqs0202260>.
- Vermeulen, W. J., & Witjes, S. (2016). On addressing the dual and embedded nature of business and the route towards corporate sustainability. *Journal of Cleaner Production*, 112, 2822-2832.
- Verschuren, P. & Doorewaard, H. (2010). *Designing a research project. Second edition*. The Hague, the Netherlands: Eleven International Publishing.
- Watanatada, P., & Mak, H. (2011). *Signed, sealed... delivered? Behind certifications and beyond labels, working paper*. London, UK: SustainAbility.
- Watson, M., & Emery, A. R. (2004). Environmental management and auditing systems: The reality of environmental self-regulation. *Managerial Auditing Journal*, 19(7), 916-928.
- Yin, R. K. (2003). *Case study research: Design and Methods. Applied Social Research Methods Series* (3rd ed.). London, UK: SAGE.

Figures

- Azapagic, A. (2003). Systems approach to corporate sustainability: a general management framework. *Process Safety and Environmental Protection*, 81(5), 303-316.
- DSF (2014). The Dairy Sustainability Framework. Retrieved November 18, 2015, from <http://dairysustainabilityframework.org/the-gdaa/programmes-of-the-gdaa/the-dairy-sustainability-framework/>
- FrieslandCampina (2015). Presentation Dairy Sustainability Framework, introduction and implementation at FrieslandCampina.
- Moen, R., & Norman, C. (2006). Evolution of the PDCA cycle (Detroit). Retrieved May 25, 2016 from http://pkpinc.com/files/NA01_Moen_Norman_fullpaper.pdf

Appendices

Appendix A. List of interviews

<i>Stakeholder type</i>	<i>Function</i>	<i>Number</i>	<i>Date</i>
Internal	Supplier Quality Management	1	12.01.2016
	Corporative Affairs	2	13.01.2016
	Corporate Sustainability	3	27.01.2016
	Supplier Quality Management	4	01.02.2016
	Corporate Communication	5	03.02.2016
	Corporate Media Relations	6	03.02.2016
	Safety, Ingredients	7	22.02.2016
	Safety, Ingredients	8	22.02.2016
	Corporative Affairs	9	25.02.2016
	Corporative Affairs	10	02.03.2016
External	Consultant	1	12.01.2016
	Consultant	2	15.01.2016
	Consultant	3	18.01.2016
	Project Manager, Researcher/University	4	25.01.2016
	Sustainability expert/ University	5	26.01.2016
	Senior Researcher	6	28.01.2016
	Senior Researcher	7	28.01.2016
	Consultant/ Senior Adviser	8	02.02.2016
	Comparable Case	9	08.02.2016
	Lead Assessor	10	12.02.2016
	Accountant/ Sustainability Expert	11	18.02.2016
	Accountant	12	18.02.2016
	Consultant	13	03.03.2016
	Consultant	14	03.03.2016
	Comparable Case	15	04.03.2016
	Sustainability expert/ University	16	09.03.2016

Appendix B. List of observations

Type of observation	Number	Date and place	Explanation
Bilateral meeting	1	20.11.2015 Amersfoort	Bilateral discussion about auditing on farm level. How is auditing currently performed and what will be the changes in the future when a general system is in place? What are the consequences from a cost perspective? These things were discussed in this meeting. (Duration: 15 minutes)
Internship day for FrieslandCampina interns	2	07.01.2016 Amersfoort and surrounding areas	Presentations about FrieslandCampina, job application tips and introduction to the dairy farm (including farm visit) (Duration: one day)
Launching of new strategy	3	18.01.2016 Amersfoort	Assisting at the launch of the new strategy towards 2020, in which sustainability plays a large role. At this launching, I attended a presentation and guided the employees through the new strategy. (Duration: one day)
Meeting with advisor and experts	4	12.02.2016 Amersfoort	Discussion meeting about how to set up the self-assessment questionnaire, certification scheme and certification protocol. Furthermore, a review and discussion about the PMS was executed (Duration: one hour)
Meeting Advisory Board	5	16.02.2016 Amersfoort	In this meeting the members of the advisory board of the DSF Pilot discussed the next steps that should be taken after the pilot, which includes adding levels of maturity. Presentations (including a presentation by myself about findings of the maturity levels in this study so far) and feedback/discussion were given. (Duration: 4 hours)
Bilateral meeting	6	17.02.2016 Amersfoort	Discussion about how to include regional differences in the PMS for dairy. Examples of regional clustering are discussed and a draft regional division is made for the PMS. Conclusion from this meeting: because of the difference between processes and outcome performance, regional variation is very important to take into account. Therefore, outcome orientated parts of the DSF scheme should be determined per region. Regions are based on climatological differences. The basic regions will be used for general comparisons between organizations, for a more detailed comparison, sub-regions are necessary (Duration: 45 minutes)
Meeting with dairy competitor	7	10.03.2016 Wageningen	Discussions and presentations (including one presentation by myself about the relevance of a PMS for the dairy industry) about how to reach higher sustainability performances, by cooperating and sharing knowledge (Duration: one day)
Meeting for Advisory board and Steering	8	21.03.2016 Amersfoort	Discussion about the relevance of a PMS for the dairy industry (further to DSF) and about the agenda for the steering board and advisory



Board with advisor			board meeting at April 7 th (Duration: one hour)
Meeting with consultant	9a	15.03.2016 Utrecht	Sharing ideas and discussion about how to apply levels and statements to the specific DSF themes and regional variation. Conclusion: a uniform approach for all Core Themes provides clarity and simplicity for the self-assessment questionnaire. Materiality analysis should be linked to score on Core Themes. So, what is answered for the materiality assessment questions is important for determining the score on the themes. For the Core Themes the focus should be on the questions, rather than on the determination of the five levels. (Duration: one hour)
Meeting with consultant	9b	22.03.2016 Utrecht	Follow-up appointment for which was discussed in 9a (Duration: one hour)
Meeting with consultant	9c	30.03.2016 Utrecht	Follow-up appointment for which was discussed in 9a and 9b (Duration: one hour)
Bilateral meeting	10	23.03.2016 Amersfoort	Discussion about materiality assessment application in the PMS for dairy and how to prevent a race to the top. What is the importance of materiality assessment and how should it be used for the system? Conclusion: stakeholder engagement is very important in determining materiality and to reach higher levels of maturity. By DSF is established that all 11 themes should be endorsed by its members. However, the materiality assessment of the organization is leading in which themes get higher priority. Thus, how comprehensive this assessment is, is linked to stakeholder engagement (Duration: one hour)
Meeting with advisor Corporate Sustainability	11	24.03.2016	Discussion about the structure of the PMS, some feedback and/or tips are given and how the <i>Core Theme</i> should be assessed by the PMS. Conclusion: keep the structure of the statements for the Core Themes univocal. Clear structure is required. Keep the themes separated and do not combine them. Additionally, there was determined that organizations start with having no plan, followed by the development of a plan or program, the formal definition of the program and having a comprehensive program in place. This transition is important to take into account for the Themes. Furthermore, minimum standards were discussed; when organizations do not comply with minimum standards, they are not able to join the system further than level 2 (Duration: 45 minutes)
Meeting about DSF with sustainability experts	12	30.03.2016	Meeting about DSF and the PMS. Discussion about how to apply scoring to the PMS and feedback about the PMS questions/statements were given. Furthermore, tips about a pilot were given (Duration: 45 minutes)
Meeting Steering Board & Advisory Board	13	07.04.2016 Amersfoort	In this international meeting members of the advisory and steering board convened to discuss the progress of DSF, a review of its



			<p>pilots and presentations were given about the next steps (including a presentation by myself about the scientific relevance and background of maturity levels for DSF) Conclusion: in order to convince the entire DSF secretary, a letter is written by the members of the pilots about how the DSF was implemented in the past year, because there was determined that communication within DSF is one of the most important things to come to improvement of the system as well as to come to sustainable improvement in general. Furthermore, the emphasis was on the fact that DSF should encourage continuous improvement of sustainability. Furthermore, there was stated that a performance measurement is a good tool for credibility of the DSF. (Duration: 3 hours)</p>
Consultation by phone with sustainability expert and professor from University	14	03.05.2016 Amersfoort	<p>Discussion about the 'dilemma' of processes vs performance. In this phone call we talked about which strategies can be used to overcome this dilemma. Conclusion: To tackle the 'dilemma' of processes vs performance, separate targets and requirements should be established by DSF regarding the standards of the specific region. These targets do not have to be concrete numbers, but more like a voluntary guideline to which organizations can be measured. So, differentiation in regions is very important. (Duration 30 minutes)</p>
Meeting with project officer Sustainability	15	03.05.2016 Amersfoort	<p>Sharing experience of filling in the self-assessment, testing the first design of the PMS for dairy (pilot testing) and the comparison between ISO 26000 and the statements of the PMS, the so called gap analysis. Conclusion: A number of statements was too vague or has too much overlap with other statements and are therefore deleted or adjusted. Furthermore, from the comparison with ISO was derived that the PMS is more specific to the dairy sector and is therefore more relevant. Additionally, the support of continuous improvement of sustainability is seen as a big advantage in relation to the more 'static' approach of the ISO 26000. (Duration: one hour)</p>
Farm visit	16	16.05.2016 Winkel (NH)	<p>Visit of a dairy farm, including a guided tour and volunteering work during the day (Duration: one day)</p>
Meeting with project officer Sustainability	17	18.05.2016 Amersfoort	<p>Discussion about tightening and improving the statements of the PMS/self-assessment, what are the strengths and weaknesses and is there overlap between statements? Conclusion: certain statements or answer possibilities need to be more concrete, this is a matter of rephrasing as well as changing some terms. A few of the statements had overlap with other parts and due to that, these are rephrased or deleted. Furthermore, the system was clear and understandable. (Duration: 2.5 hours)</p>

Appendix C. Format of the interview

The interviews were conducted in Dutch. Therefore, in this appendix the interview structure is translated from Dutch to English.

Interview

Short introduction

- Short introduction about my project (see Appendix D)

At the beginning of each interview I asked if the interviewee agreed with recording the interview. The information was used anonymously for my thesis and for designing the PMS. Thereafter, I also asked whether the interviewee wishes to receive the transcript after the interview.

A few questions to know more about person's background

1. What is your exact function within the organization?
2. How long are you working in this organization and how long are you working in your current position/function?
3. What is your definition of sustainability and/or a sustainable organization? And how is your daily work related to sustainability?

Sustainability in the dairy industry

4. Are you familiar with the DSF? (see presentation)
5. Are you familiar with benchmarking systems?
6. How should a sustainability benchmarking system look like for the dairy industry according to you? So, in other words, how could dairy companies be benchmarked compared to each other/other dairy companies?
7. How would you describe/define different levels maturity for a performance measurement system?

Maturity levels in particular

8. What are, according to you, the characteristics of an organization that starts to orientate itself on sustainability (level 1)?
9. What should be the minimum level of sustainability?
10. What are, according to you, the characteristics of an organization that performs optimally on sustainability? So, in other words, what are the criteria to which an organization should comply to be 'optimized' in its sustainability performance?
11. What are the most important steps that should be taken by an organization to go through the different levels?

Scheme with criteria derived from literature

According to literature, I have defined a set of categories and derived various characteristics per theme.

12. What are, according to you, the most important criteria in sustainability performance measuring?
13. So, in other words, which criteria should weight more, and which criteria are less important/should weighted less?



14. Which criteria are still missing in this scheme, according to you?

Concluding

15. Could you please repeat which criteria are the most important for building a proper PMS (or benchmarking system)? And what are the most important characteristics of an organization operating at level 1 and at level 5?
16. Do you have additional questions or comments? Things I did not think about?
17. Would you mind if I email you for further questions during my thesis process?
18. Would you like to receive a transcript of this conversation?
19. Finally, do you have any further recommendations, for example people I can consult for questions or documents/reports which could be useful for my project and provide some additional information?

Appendix D. Interview presentation



A collaborative approach to sustainability in the global dairy value chain

Towards an international Performance Measurement System for the dairy industry

Applied to the Dairy Sustainability Framework



Willingness within the international dairy sector to become more sustainable

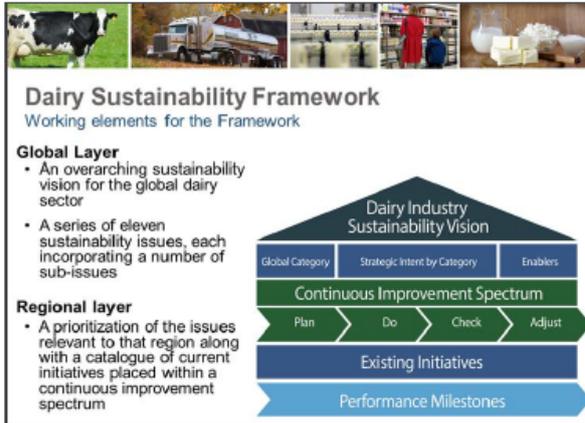


- ❖ Global trend to apply sustainability strategies in business procedures
- ❖ Moving beyond standards and towards a sector-wide cooperation
- ❖ Many initiatives from dairy industries, food companies, retailers and NGO's

- ❖ A common language for sustainable dairy → Dairy Sustainability Framework (DSF)
- ❖ Why the DSF?
- ❖ International dairy industry wants to take leadership



Dairy Sustainability Framework



- ❖ A process based approach (Plan do check adjust)
- ❖ Within 11 core sustainability themes



2



From an industry-wide agenda to the implementation of a global framework



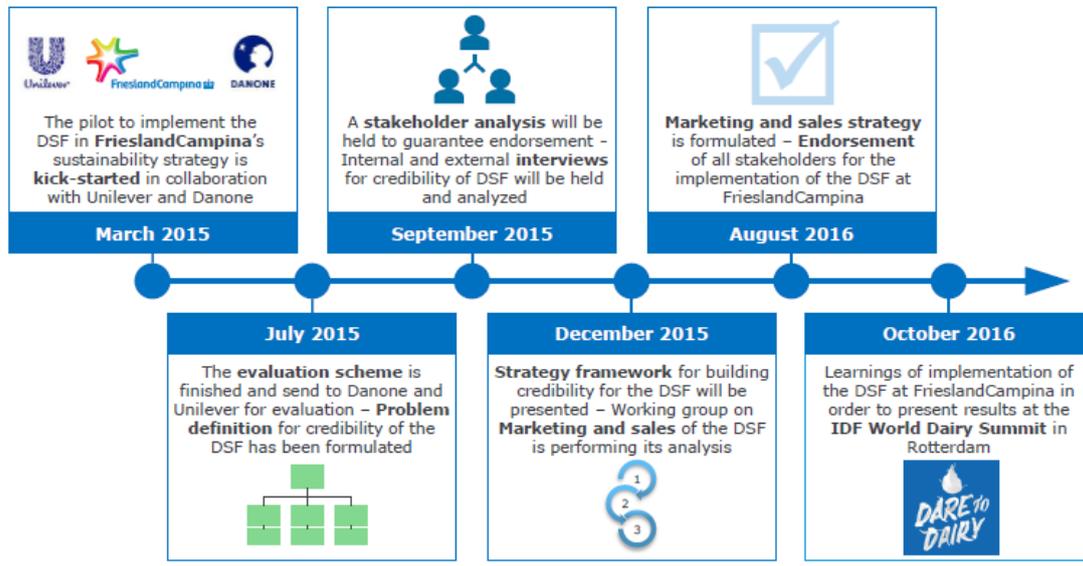
Timeline of the Dairy Sustainability Framework

3

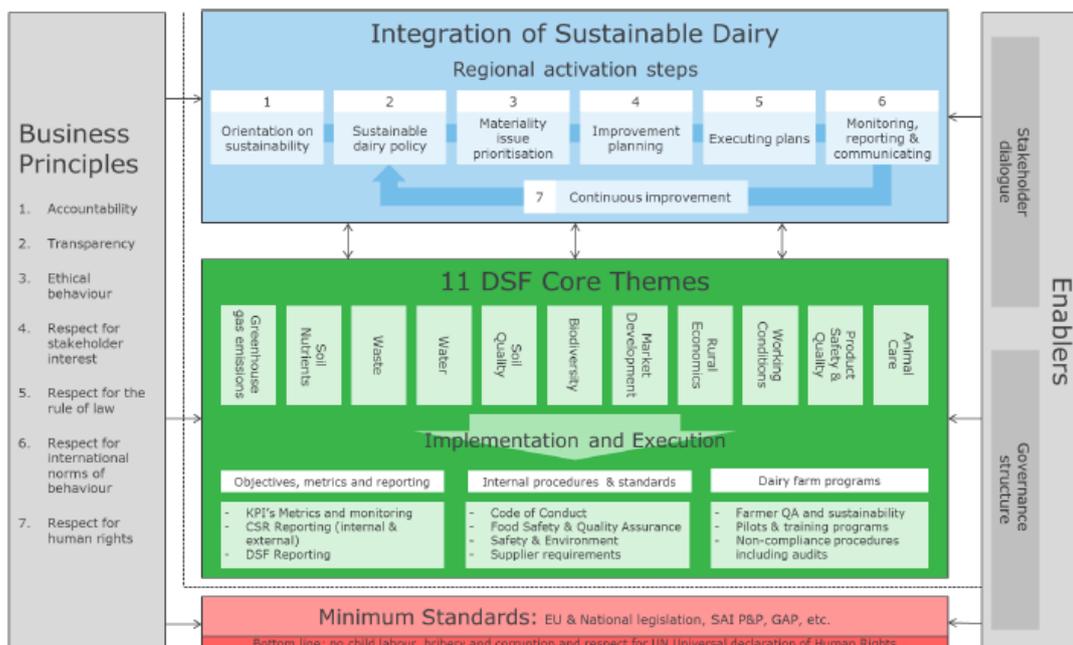


Implementation pilot with FrieslandCampina, Danone and Unilever

Timeline of the Pilot



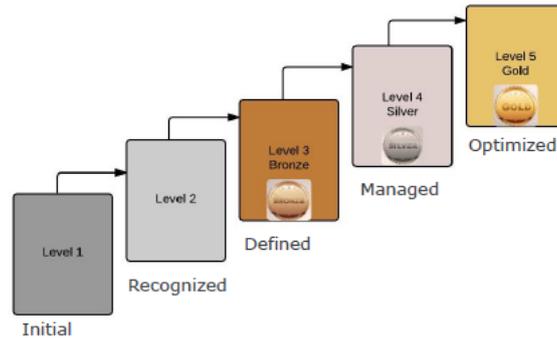
Demonstration Scheme





Towards a Performance Measurement System (PMS)

- ❖ Important in the DSF is continuous improvement of sustainable performance
- ❖ In the support towards further continuous improvement, a Performance Measurement System (PMS) provides added value. The PMS is a system which measures performance by making use of maturity levels which will give insights to the company to determine their relative performance in relation to others
- ❖ Besides, the PMS shows the company in which categories they should improve to become 'higher on the ladder'
- ❖ The PMS will be developed in cooperation with Utrecht University



6



Criteria	Levels of Maturity				
	Level 1 Initial	Level 2 Recognized	Level 3 Defined	Level 4 Managed	Level 5 Optimized
Management and strategy	<ul style="list-style-type: none"> Floors defined and practices are unstructured Most agreements and arrangements are informal Company's insight in (energy) use is partial 	<ul style="list-style-type: none"> Sustainability will be recognized and/or identified Sustainability direction and goal setting are identified, but not aligned with business main targets 	<ul style="list-style-type: none"> Increased interest towards continuous improvement and sustainability Sustainability starts to be integrated into business procedures Sustainability management is present, but still no alignment with other processes 	<ul style="list-style-type: none"> Sustainability is major driver for business, but still not in total harmony with other goals Sustainability is integrated and mainly embedded in the business functioning 	<ul style="list-style-type: none"> Sustainability is central in company's practice Sustainability is completely integrated and optimized Management is based on continuous performance improvement and learning Knowledge management is integrated
Cooperation with stakeholders	<ul style="list-style-type: none"> Cooperation is low Relevant cooperation is transactional (competition amongst partners) 	<ul style="list-style-type: none"> Proactive collaboration starts to be built Collaboration mainly internal 	<ul style="list-style-type: none"> Increased cooperation between organizations and 'actors' in supply chain and intra-organizational Structured collaboration with partners starts to be noticeable 	<ul style="list-style-type: none"> Cooperation along supply chain is in place 	<ul style="list-style-type: none"> Collaboration and information sharing (multi-organizational/cross-enterprise) is high The objective is to actively engage with stakeholders
Predictability and goals achievement	<ul style="list-style-type: none"> Highly unpredictable and uncertain Sustainability direction and goal setting are only partly known 	<ul style="list-style-type: none"> Process performance becomes more predictable Goals (and) targets are defined, but missing more often than not 	<ul style="list-style-type: none"> Process performance becomes more predictable 	<ul style="list-style-type: none"> Process performance is more predictable and targets will be achieved Process improvement goals are achieved 	<ul style="list-style-type: none"> Trust and reliability is high
Data collection and measuring	<ul style="list-style-type: none"> Sustainability data is not consistently collected or analyzed 	<ul style="list-style-type: none"> Sustainability measurement is growing 	<ul style="list-style-type: none"> Performance improvement measures take place 	<ul style="list-style-type: none"> Sustainability goals are quantified and measurable Measurement of sustainability is clear and generally understood 	<ul style="list-style-type: none"> Measurements and monitoring of process performance are common Advanced process documentation
Approach and vision	<ul style="list-style-type: none"> Reactive Informal 	<ul style="list-style-type: none"> Reactive, but moving towards proactive Mainly informal 	<ul style="list-style-type: none"> From informal towards formal 	<ul style="list-style-type: none"> Business functioning is proactive and formal Long-term commitments A shared vision for sustainability is defined in agreement among employees 	<ul style="list-style-type: none"> Proactive Formal Company's strategy has long-term focus
Communication	<ul style="list-style-type: none"> Minimal communication processes and channels in place Communication about sustainability is exceptional 	<ul style="list-style-type: none"> Communication of sustainability is still only internal External communication only with most important partners and not consistent over time 	<ul style="list-style-type: none"> Internal as well as external communication is present and external communication (with stakeholders) starts to increase 	<ul style="list-style-type: none"> From (rather) internal towards external and public (between company and NGOs and governmental parties) Communication regarding sustainability is consistent and integrated Communication processes and channels are based on feedback loops 	<ul style="list-style-type: none"> Communication channels are fully presented (internal and external) Learning and knowledge flows are part of communication Communication is complete
Reporting	<ul style="list-style-type: none"> No consideration of sustainability issues in any report or communication channel 	<ul style="list-style-type: none"> Only the most relevant sustainability issues are reported internally 	<ul style="list-style-type: none"> Sustainability issues are reported internally and the most important issues are also reported external 	<ul style="list-style-type: none"> Sustainability issues are presented in annual reports and through internal communication channels 	<ul style="list-style-type: none"> Sustainability issues are presented in the annual reports and through internal communication channels. Besides results, also goals and measures are well reported

Levels of Maturity

7



Main questions

- ❖ To which key criteria should a company comply to become 'optimized' in their sustainability performance (level 5)?
- ❖ How will a company normally develop? What are the most important steps taken?
- ❖ What are the characteristics of a start-up company (level 1)?

Appendix E. Brief explanation BSC and SCOR

The Balanced Scorecard

The Balanced Scorecard (BSC) is designed by Kaplan and Norton in the 1990s to measure performance of organizations.

The traditional performance measures for financial aspects were appropriate for the industrial era, however they were out of step with the skills and capabilities organizations faced in the 90's. According to Kaplan and Norton (1995), there was a clear demand for a system that does not make a distinction between the financial and the operational measures, but combines these two perspectives in one system. From this the balanced scorecard is derived, which is defined as *'a set of measures that gives top managers a fast but comprehensive view of the business'*. (pp. 71). The system is based on the assumption that efficiency in investments of capital is no longer the only factor for business success, but soft factors like knowledge exchange, customer orientation etc. have become important as well (Figge et al., 2002). In relation to this, the balanced scorecard assesses an organization from four perspectives; firstly, the customer perspective, which emphasizes the customer's view, how does the customer see the organization? The second perspective is from the internal organization, what must the organization excel at? Thirdly, the organization is assessed from an innovation and learning perspective, in which continuous improvement and creation of value are important issues. And lastly, the financial perspective is taken into account, which looks at the organization in relation to its shareholders (Kaplan & Norton, 1995). Furthermore, the balanced scorecard is characterized by simplicity, it aims to force organizations to focus on the most important criteria only (Kaplan & Norton, 1995).

The balanced scorecard is useful as a system for sustainability management, because the issues of corporate contribution to sustainability will be addressed in an integrative way. For this, organizations who desire to become more sustainable have to improve their strategies in all three dimensions of sustainability (triple bottom line). And these dimensions should not be considered separately, like in some approaches (Figge et al., 2002).

The Supply Chain Operations Reference (SCOR) model

The Supply Chain Operations Reference (SCOR) model is a process reference model, which aims to be an industrial standard that enables improved supply chain management (Huang, Sheoran & Keskar, 2005). The SCOR model is a frequently used, integrated system of business process re-engineering, benchmarking and process measurement. The framework or model consists of; *'standard descriptions of management processes, a framework of interactions amongst average processes, standard metrics to measure process performance, management practices that produce best in class performance, and lastly standard alignment to software features and functionality'*. (Huang, Sheoran & Wang, 2004, pp. 24).

Furthermore, the SCOR model distinguishes four processes; source, make, deliver, and plan. Additionally, these processes themselves are described in detail whereas they are further divided into process elements, tasks and activities (Huang et al., 2004). The SCOR model consists of five stages or levels of maturity which show the development of activities and process towards effective supply chain management (Lockamy & McCormack, 2004). So, via the model the above mentioned processes are measured and then there can be determined in what level the organizations are related to supply chain management. The SCOR model is useful to consider in the case of performance measurement systems, because it makes use of maturity levels as well.



Appendix F. Characteristics derived from literature

Levels of Maturity					
Category	Level 1 <i>Initial</i>	Level 2 <i>Recognized</i>	Level 3 <i>Defined</i>	Level 4 <i>Managed</i>	Level 5 <i>Optimized</i>
Approach & Vision	<p>Reactive approach^{1,2}</p> <p>Informal²</p> <p>There is basically no aim for sustainability or CS⁶</p>	<p>Reactive approach²</p> <p>Mainly informal²</p> <p>Sustainability issues start to emerge⁶</p>	<p>Reactive approach, moving towards proactive¹</p> <p>From informal towards formal²</p>	<p>Business functioning is proactive^{1,2,4}</p> <p>Formal^{1,2,4}</p> <p>Long-term commitments^{1,2,4}</p>	<p>Proactive approach^{1,2,4,6}</p> <p>Formal^{1,2,4}</p> <p>Company's strategy has long-term focus and commitment^{1,2,4}</p> <p>A shared vision for sustainability is defined in agreement among employees¹</p>
Management & Strategy	<p>Management is poorly defined and practices are unstructured^{3,4}</p> <p>Most agreements and arrangements are informal²</p> <p>Company's insight of their impacts is partial⁵</p> <p>Driving force for sustainability is governmental legislation⁶</p>	<p>Sustainability is recognized and/or identified¹</p> <p>Sustainability direction and goal setting are identified, but not aligned with business main targets^{1,2}</p>	<p>Increased interest towards continuous improvement and sustainability¹</p> <p>Sustainability starts to be integrated into business procedures^{1,6}</p> <p>Sustainability management is present, but still no alignment with other processes¹</p>	<p>Sustainability is major driver for business, but still not in total harmony with other goals^{1,2,6}</p> <p>Sustainability is integrated and mainly embedded in the business functioning^{1,2}</p>	<p>Sustainability is central in company's practice^{1,2,6}</p> <p>Sustainability is completely integrated and optimized^{1,2,6,7}</p> <p>Management is based on continuous performance improvement and learning¹</p> <p>Knowledge management is integrated¹</p>
Cooperation & Engagement	<p>Engagement and cooperation is low, the organization is not an active partner in networks^{4,5}</p> <p>If present, cooperation is hierarchical (competition amongst partners)¹</p> <p>Informal contact with stakeholders⁵</p> <p>Focus on own organization only^{4,5,6}</p>	<p>Collaboration starts to be felt^{1,4}</p> <p>Collaboration and stakeholder engagement is mainly internal^{1,4}</p> <p>Contact with stakeholders become more systematic¹</p>	<p>Increased and active cooperation between organizations and 'actors' in supply chain and intra organizational^{1,3,4,5}</p> <p>Structured collaboration with partners starts to be noticeable¹</p> <p>Focus on own organization and other actors in supply chain^{1,5}</p>	<p>Cooperation along supply chain is in place^{1,4}</p> <p>Cooperation between organization and stakeholders/ society⁶</p> <p>Cooperation with main stakeholders regarding all environmental, social and economic issues¹</p> <p>Active participation in most dialogues¹</p>	<p>Collaboration and information sharing (multi-organizational/ cross-enterprise) is high^{1,2,3,4}</p> <p>Organizations actively engage with stakeholders^{1,2,3,4,6}</p>



<p>Goals achievement & Commitment</p>	<p>Highly unpredictable and uncertain¹</p> <p>Sustainability direction and goal setting are only partly known^{2,3,4}</p> <p>Estimation of future performance is not in place^{2,3,4}</p> <p>Short-term commitments.⁶</p> <p>Compliance with law and regulations, mandatory requirements^{6,5}</p>	<p>Process performance becomes more predictable¹</p> <p>Goals or/and targets are defined¹, but missing more often than not^{1,4}</p> <p>Estimation of future performance is partly planned^{2,3,4}</p> <p>Short-term commitments⁶</p> <p>Compliance with law and regulations and some voluntary frameworks as well^{5,6}</p>	<p>Process performance is more predictable, but only in a qualitative way^{1,2}</p> <p>Estimations of future performance are in place¹</p> <p>Medium-term commitments⁶</p> <p>Mandatory and voluntary frameworks towards CS governance are focused on⁵</p>	<p>Process performance is more predictable and targets will be achieved and will be measured in a quantitative way^{1,2}</p> <p>Estimations of future performance are in place¹</p> <p>Long-term commitments^{1,2,4,6}</p> <p>Goals are quantified and measurable¹</p>	<p>Trust and reliability is high</p> <p>Proactive process improvement and goals are achieved⁵</p> <p>Estimations of future performance are in place and plans to reduce future impacts are clear</p> <p>Goal is to reach zero impact on the environment⁶</p> <p>Long-term commitments^{1,2,4,6}</p>
<p>Data collection & Measuring</p>	<p>Sustainability data is not constantly collected or analyzed¹</p> <p>Performance/process measurement is not in place⁴</p>	<p>Sustainability measurement is growing¹</p> <p>Basic processes are documented⁴</p> <p>Performance measurement for own organization¹</p>	<p>Performance improvement measures take place^{3,4}</p>	<p>Sustainability goals are quantified and measurable¹</p> <p>Measurement of sustainability is clear and generally understood¹</p>	<p>Measurements and monitoring of process performance are common^{1,3,4}</p> <p>Advanced process documentation^{1,3,4}</p>
<p>Communication</p>	<p>Minimal communication processes and channels in place¹</p> <p>Communication about sustainability is exceptional¹</p> <p>Internal¹</p>	<p>Communication of sustainability is mainly internal^{1,5}</p> <p>External communication only with most important partners and not consistent over time⁵</p> <p>Internal¹</p>	<p>Internal as well as external communication is present and external communication (with stakeholders) starts to increase⁵</p>	<p>From rather internal towards external and public^{1,2}</p> <p>Communication regarding sustainability is consistent and integrated¹</p> <p>Communication processes and channels are based on feedback loops¹</p> <p>Shift towards external focus^{1,2}</p>	<p>Communication channels are fully present (internal and external)¹</p> <p>Learning and knowledge flows are part of communication¹</p> <p>Communication is complete¹</p> <p>External focus¹</p>
<p>Reporting</p>	<p>No consideration of sustainability issues in any report or communication channel⁵</p>	<p>Only the most relevant sustainability issues are reported internally⁵</p>	<p>Sustainability issues are reported internally and the most important issues are also reported external⁵</p>	<p>Sustainability issues are presented in annual reports and through internal communication channels (one-way communication⁵)</p>	<p>Sustainability issues are presented in the annual reports and through internal communication channels (two-way communication⁵)</p>

1) Cagnin et al., 2005; 2) Kirkwood et al., 2011; 3) McCormack et al., 2008; 4) Lockamy & McCormack, 2004; 5) Baumgartner & Ebner, 2010; 6) Van Marrewijk & Werre, 2003; 7) Figge, 2002



Appendix G. Characteristics derived from interviews and comparable cases

Levels of Maturity					
Category	Level 1 <i>Initial</i>	Level 2 <i>Recognized</i>	Level 3 <i>Defined</i>	Level 4 <i>Managed</i>	Level 5 <i>Optimized</i>
Approach & Vision	The organization does not have a plan or program ^{O 11} Starting orientation/awareness of sustainability/CS should be in place ^{EI 4,6,7,8}	The organization is developing a plan or program (informal) ^{O 9c,11} Emphasis is on reactive approach rather than proactive ^{EI 3} A clear plan is lacking ^{EI 3}	A plan or program is formally defined ^{O 9c,11}	A program is defined and communicated ^{O 9c,11} Proactive approach ^{II 3,7,8}	A program is defined, communicated and integrated ^{O 9c,11} Proactive approach ^{EI 16; II 7,8} Continuous improvement and emphasis is on 'to be better than the competition' ^{II 1}
Management & Strategy		Sustainability is identified, but is not yet defined ^{EI 4}			Training within the organization is important to become higher in CS ^{II 7,8} Stakeholder dialogue outcomes are incorporated in strategy ^{EI 2}
Stakeholder engagement	No contact or only informal contact with stakeholders ^{EI 2,9; II 5; O 11} 'How' is more important than 'how often' ^{EI 2}	Contact with stakeholders is starting, but not yet publically, mainly one-to-one conversations ^{II 5} Contact with stakeholders becomes more systematic ^{EI 2} 'How' is more important than 'how often' ^{EI 2}	Contact with stakeholders is present ^{O 11} External stakeholders identified ^{EI 9} Commitment with NGO's and external stakeholder engagement starts to occur ^{II 5,7,8} Communication plan is present ^{EI 9}	Contact with stakeholders is comprehensive ^{O 11} Outcomes of stakeholder dialogues are applied in business strategies ^{EI 2} 'How' is more important than 'how often' ^{EI 2}	The organization engages the entire society, not only the direct stakeholders and supply chain ^{CC, EI 15} Commitment with stakeholders and societal engagement is high ^{EI 3,6,7,16; II 5} 'How' is more important than 'how often' ^{EI 2}
Predictability & Goals achievement	The organization does not have defined goals ^{CC, EI 9}	The organization defined short term goals or goals only for own organization ^{CC, EI 9}	The organization defined its goals ^{EI 9; II 5} Long term goals starts to be in place ^{O 9c}	The organization achieve its goals or is on its way to achieve it ^{EI 3; O 9c,11} Organization	The organization achieves or even outreach its goals ^{EI 3,9; O 11} Goals are defined and communicated ^{EI 2}



	Short term goals ^{O 9c}	Progress towards goal is in place ^{O 9c}		defines goals for own organization and supply chain ^{CC, EI 9}	The goal is to have a net positive influence on the planet ^{EI 8} Long term goals are defined and engage entire supply chain ^{EI 9; O 9c,11}
Data collection & Measuring	Performance measures are not in place ^{CC, EI 13,14; II 3}	Performance is measured for own organization ^{CC, EI 13,14; II 3}			Performance is measured for entire supply chain ^{CC, EI 13,14; II 3}
Communication	First step: create internal support and awareness ^{EI 9; II 5,6} Internal ^{II 5,6} Intern towards extern is better, because the first step to go forward is to create the right culture in the organization ^{II 1,5,6}	Internal support and awareness ^{EI 9; II 5,6} Internal ^{II 5,6}	Internal support creation amongst employees and external embedding ^{EI 5; II 5,6} Internal towards external ^{II 5,6}	From rather internal towards external and public (between company and NGO's and governmental parties) ^{CC} Emphasis on external, but also internal ^{II 5,6}	Communication is very important in all levels, through the entire process of maturity ^{EI 16; II 5,6,7,8; O 13} Importance of transparent external communication is high ^{EI 4; II 5,6}
Reporting, Output & Results	Output and results will not weight very much ^{EI 5,6,7}	Output and results will not weight very much ^{EI 5,6,7}	Output and results become important to include in the scoring ^{EI 3} Focus will be more on outcome ^{EI 6,7; O 13}	Output and results should be clear and targets should be met ^{EI 1,3} Focus on outcome ^{EI 6,7; O 13} Besides results, also goals and measures are well reported ^{II 5} Reporting as well as communication have to be transparent ^{II 5,6}	Organization should prove its results and own targets are met ^{EI 1,3} In the end, the focus should be on outcome ^{EI 6,7; II 5; O 13} Besides results, also goals, measures and progress towards the goal are well reported ^{EI 2; II 5} Reporting as well as communication have to be transparent ^{II 5,6}
Verification	Verification is not yet in place, self-assessment ^{EI 3}	Verification is not yet in place, self-assessment ^{EI 3}	Self-assessment with possibilities towards external verification ^{EI 3}	External verification is in place ^{EI 3}	External verification is in place and there are possibilities for certification ^{EI 3}



<p>Materiality assessment</p>	<p>No assessment^{EI 2; O 9a}</p> <p>Organization should start with determining which (environmental) impacts it has^{EI 8}</p>	<p>Internal assessment^{EI 2}</p> <p>Assessment, but not verified^{O 9a}</p>	<p>Starting with an external assessment, with selection of stakeholders^{EI 2}</p> <p>Assessment is communicated, not verified yet^{O 9a}</p>	<p>Extensive stakeholder survey to assess materiality^{EI 2}</p> <p>Assessment is communicated and verified with external parties^{EI 8; O 9a}</p>	<p>Stakeholder engagement as well as societal engagement^{EI 2,16}</p> <p>Assessment is communicated, verified and maintained^{EI 8; O 9a}</p>
<p>Subjectivity & Transparency</p>	<p>Subjectivity and transparency will grow throughout the levels^{II 5,6}</p> <p>Transparency is important^{EI 4}</p>				

CC) Comparable Case; EI) External Interview; II) Internal Interview; O) Observation