



**UNIVERSITY  
OF TWENTE.**

**SUPPLIER PERFORMANCE MEASUREMENT SYSTEM DESIGN IN  
MANUFACTURING INDUSTRY**

Lappeenranta–Lahti University of Technology LUT & University of Twente

Double degree

Master's programme in Supply Management, Master's thesis

2023

Eveliina Toivakka

Examiners: Prof. Anni-Kaisa Kähkönen, LUT  
Junior Researcher Iryna Maliatsina, LUT  
Asst. Prof. Vincent Delke, UT  
Asst. Prof. Carolina Belotti Pedroso, UT  
Asst. Prof. Petra Hoffmann, UT

## ABSTRACT

Lappeenranta–Lahti University of Technology LUT & University of Twente

LUT Business School & Faculty of Behavioural, Management and Social Sciences

Business Administration

Eveliina Toivakka

### **Supplier Performance Measurement System Design in Manufacturing Industry**

Master's thesis

2023

97 pages, 18 figures, 9 tables, and 4 appendices

Examiners: Prof. Anni-Kaisa Kähkönen, Junior Researcher Iryna Maliatsina, Asst. Prof. Vincent Delke, Asst. Prof. Carolina Belotti Pedroso, Asst. Prof. Petra Hoffmann

Keywords: Supplier performance measurement system, Supplier performance management, Supplier relationship management, design framework

The poor performance of suppliers causes unnecessary costs for the buying company. To support supplier management and supplier performance management, a supplier performance measurement system should be built. Previous research on supplier performance measurement systems focuses on what is done in companies, leaving out the optimal way to design a supplier performance measurement system. The topic was approached through a literature review and interviews. The interviews were analysed using the Gioia methodology. The results revealed that in a large organization, a supplier performance measurement system should enable the sharing of information to different parts of the company, so that the supplier's performance can be reliably evaluated. This thesis presents a design framework that can be used to design a supplier performance measurement system.

## TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT & Twenten yliopisto

LUT-kauppakorkeakoulu & Käyttäytymis-, johtamis- ja yhteiskuntatieteiden tiedekunta

Kauppatieteet

Eveliina Toivakka

### **Toimittajan suorituskyvyn mittausjärjestelmän suunnittelu teollisuudessa**

Kauppatieteiden pro gradu -tutkielma

2023

97 sivua, 18 kuvaa, 9 taulukkoa ja 4 liitettä

Tarkastajat: Professori Anni-Kaisa Kähkönen, Nuorempi tutkija Iryna Maliatsina, Apulaisprofessori Vincent Delke, Apulaisprofessori Carolina Belotti Pedroso, Apulaisprofessori Petra Hoffmann

Avainsanat: Toimittajan suorituskyvyn mittausjärjestelmä, Toimittajan suorituskyvyn hallinta, Toimittajasuhteiden hallinta, Suunnitteluviitekehys

Toimittajan heikko suorituskyky aiheuttaa yritykselle tarpeettomia kustannuksia. Toimittajan suorituskyvyn mittaaminen mahdollistaa suorituskyvyn seuraamisen ja todentamiseen. Toimittajahallinnan tueksi tulisi rakentaa toimittajan suorituskyvyn mittausjärjestelmä, jota voidaan hyödyntää toimittajan suorituskyvyn hallinnassa. Järjestelmään liittyvä tutkimustieto keskittyy tällä hetkellä siihen, mitä yrityksissä tehdään, jättäen tutkimatta optimaalisen tavan suunnitella mittausjärjestelmä yrityksien tarpeisiin. Aihetta lähestyttiin kirjallisuuskatsauksen ja haastatteluiden avulla. Haastattelut analysoitiin Gioia-menetelmällä. Tulokset paljastivat, että suuressa organisaatiossa toimittajan suorituskyvyn mittausjärjestelmän tulisi mahdollistaa tiedon jakamisen yrityksen eri osiin, jotta toimittajan suorituskykyä voidaan luotettavasti arvioida. Työssä esitellään suunnitteluviitekehys, jota voidaan käyttää toimittajan suorituskyvyn mittausjärjestelmän kehittämiseen.

## **ACKNOWLEDGEMENTS**

First of all, I want to show my gratitude to my supervisors. Through your guidance, I have managed to finalize my thesis. Special thanks to Iryna, who had helped with two of my biggest projects by far and has been there in my corner while writing my theses. Also special thanks to Vincent who helped me to improve my work.

My gratitude goes out to my boss who had the patience to guide me through the whole process and providing support at any point where it was needed. And the case company for providing these excellent opportunities for students. I would like to give my thanks to all of my interviewees who provided excellent insights on my topic.

Then I would like to give thanks to my parents, my brothers, and my friends. My father has given unconditional support at every step of the way. He is the most patient person that I know allowing me to try without fear of failure. Thank you for helping out with my numerous moves as I followed my dreams. My mother has taught me to be independent and relentless. She has helped me through the university process. Then, I want to give special thanks to my friends, who have helped and supported me through the process. Thank you for being there. Especially Jesse, who has checked my work and has spent hours on end reading, re-reading, and commenting.

Writing a thesis was an experience, let's not do it again.

Espoo 18.9.2023

Eveliina Toivakka

## Table of contents

Abstract

Acknowledgements

1. Introduction.....	8
1.1 Research gap .....	10
1.2 Research goal .....	11
1.3 Definitions.....	12
2. Theoretical background .....	14
2.1 Purchasing in company's strategy .....	15
2.2 Supplier relationship management and supplier performance measuring .....	16
2.2.1 Tailoring performance measuring for strategic suppliers .....	19
2.3 A supplier performance measurement system lifecycle .....	22
2.3.1 Design phase of a supplier performance measurement system .....	24
2.3.2 Implementation phase of a supplier performance measurement system.....	27
2.3.3 Use and review phases of a supplier performance measurement system.....	29
2.4 Supplier performance evaluation.....	33
2.5 Supplier performance monitoring.....	36
2.6 Visualisation of the results .....	40
2.7 Communication .....	42
2.8. Supplier performance measurement system supporting supplier relationship management .....	45
3. Methodology .....	48
3.1 Research design.....	48
3.1.1 Case selection and description.....	49
3.2 Data collection.....	51
3.3 Data analysis .....	53
3.4 Reliability and validity.....	55
4. Findings.....	57
4.1 Managing internal process and design.....	57
4.1.1 Success factors for measuring system design.....	58
4.1.2 Quality management of SPMS .....	61
4.1.3 Dimensions with operational focus.....	63
4.2 SPMS as an enabler of supplier performance management.....	64
4.2.1 SPMS enabling intra-organizational learning .....	65
4.2.2 SPMS enabling inter-organizational relations and performance management ..	68

4.2.3 Combining all the dimensions .....	70
5. Discussion .....	71
5.1 Design framework .....	71
5.1.1 Primary Process .....	73
5.1.2 Supporting Processes .....	77
5.2 Theoretical Implications .....	78
5.3 Practical implications.....	79
6. Conclusions .....	80
6.1 Limitations and Suggestions for Future Research.....	80
Reference list .....	81

## Appendices

Appendix 1. Interview questions

Appendix 2. Quote examples of 1<sup>st</sup> order concepts

Appendix 3. Data structure of managing internal process and design

Appendix 4. Supplier performance measurement system role and management – Data structure of SPMS as an enabler of supplier performance management

## LIST OF TABLES

Table 1. Performance measurement system analysis (Neely et al., 1995)

Table 2. Supplier questionnaire themes for supplier evaluation (Wu & Blackhurst, 2009)

Table 3. Differences between diagnostic and interactive use (Henri, 2006; Maestrini et al., 2021).

Table 4. Existing performance evaluation articles.

Table 5. Existing supplier performance monitoring articles.

Table 6. Methodological choices

Table 7. Interviewees and reference in text

Table 8. Managing internal process and design themes

Table 9. Supplier performance measurement system as an enabler of supplier performance management themes

## LIST OF FIGURES

Figure 1. Division between strategic sourcing and operative procurement (Schiele, 2019)

Figure 2. Purchasing cycles between the company and supply market (Schiele, 2019)

Figure 3. Strategic and operational supplier relationship management (Modified from Lambert & Schwieterman, 2012)

- Figure 4. Supplier categorisation based on performance and suggested strategies (Lima-Junior & Carpinetti, 2016)
- Figure 5. Supplier relationship management's impact and their impact on economic value added (Lambert & Schwieterman, 2012).
- Figure 6. Phases for developing a supplier performance measurement system and aims of each phase (Maestrini et al., 2021)
- Figure 7. Supplier performance measurement system design phases (Luzzini et al., 2014)
- Figure 8. Example of a holistic supplier performance measurement system (Araz & Ozkarahan, 2007)
- Figure 9. Governance strategies for services (Ye et al., 2022)
- Figure 10. Spider graph, vendor monitoring sheet, and vendor performance dashboard (simplified from Pal & Kumar, 2008)
- Figure 11. Control chart for delivery performance (adapted from Morgan & Dewhurst, 2007).
- Figure 12. Communication and feedback nodes (Modified from Maestrini et al., 2018c)
- Figure 13. Supplier relationship management connections to supplier performance measurement system design
- Figure 14. Moving from induction to abduction in Gioia methodology (Magnani & Gioia, 2023)
- Figure 15. Second order themes and aggregate dimensions
- Figure 16. Managing internal process and design
- Figure 17. Supplier performance measurement system roles in an organization
- Figure 18. Supplier performance measurement system design framework

## **LIST OF ABBREVIATIONS**

- AHP – Analytical hierarchy process
- BU – Business unit
- ERP – Enterprise resource planning
- KPI – Key performance indicator
- MCDM – Multi-criteria decision-making
- PMS – Performance measurement system
- SCOR<sup>®</sup> – Supply Chain Operations Reference
- SPC – Statistical process control
- SPMS – Supplier performance measurement system
- SRM – Supplier relationship management

## 1. Introduction

Nowadays, world creates 5 exabytes of data that is expected to grow up to 463 exabytes by 2025 (Van der Sande, 2021). All in all, companies have more data available from various processes such as production, customer management, and purchasing, being able to make data-driven decisions. According to PwC (2023), significant benefits are expected from being a data-driven organization, as they outperform their peers in productivity and profitability. The data-driven approach is not limited to customer management as it can be applied to purchasing and supply management as well. There are possibilities to use the available data to manage suppliers' performance but it needs to be collected and refined to make profitable knowledge.

Poor supplier performance might increase the costs indirectly by 10-20 %, for example, due to unacceptable quality of parts or late delivery resulting in overtime (Belotserkovskiy et al., 2018). Going through the trouble of developing measuring capabilities in a company might pay off, since monitoring suppliers might have a positive effect on the supplier's operational performance (Maestrini et al., 2018b), ultimately benefiting the buyer. Performance measuring could be applied in the business context of supplier relationship management to help measure and reach strategic goals. Without formally measuring and monitoring the supplier's performance, companies cannot assess if the supplier is performing up to the company's standards (Simpson, Sigauw & White, 2002). Supplier performance evaluation models that are proposed by literature, are not necessarily feasible in companies due to their complexity (Luzzini, Caniato & Spina, 2014). On top of the other challenges, companies have difficulties identifying correct key performance indicators (KPI) to follow supplier performance or they lack resources or capabilities to monitor the performance (McKinsey, 2018).

Suppliers' performance can be assessed by their green performance (Luthra et al., 2017), innovativeness (Goldberg & Schiele, 2018), cost, quality (Nair, Jayaram & Das, 2015), and other criteria. Performance measurement can have multiple goals: choosing the best alternative from existing suppliers or potential suppliers, evaluating the performance of supplier development programs, or general monitoring of supplier performance (Luzzini et al., 2014). Performance management can provide means to identify quality deviations and perform root



cause analyses (Deloitte, 2017). Defining the strategic objectives of the performance evaluation models and defining the correct variables to measure suppliers are important for the effectiveness of supplier performance evaluation (Hawkins, Gravier & Muir, 2020). There are various models suggested by literature to choose the correct variables to measure. They are usually advanced mathematical models (Ho, Xu & Dey, 2010), such as AHP (Ulutas et al., 2016), DEA (Ross et al., 2009), or PROMETHEE (Araz & Ozkarahan, 2007). For example, fuzzy TOPSIS deals with imprecise data and qualitative and quantitative measurements to solve supplier selection problems or be used to assess the performance of the supplier (Chen, Lin & Huang, 2006). Also, models can be used to identify the most promising supplier for development (Araz & Ozkarahan, 2007).

Advanced models also have some drawbacks. First of all, the suggested models rarely answer to the realities of companies, as these models have some prerequisites to be feasible in use. The models assess the supplier based on quantitative and qualitative criteria which require different data collection practices e.g. additional evaluation by surveys (Zeydan, Çolpan & Çobanoğlu, 2011) or by asking suppliers to provide performance information. The collected data from suppliers might be incomparable due to subjectivity or different measuring practices, providing a challenge for supplier evaluation (Humphreys, McIvor & Chan, 2003; Humphreys et al., 2006). Also, the application of these models is a time-consuming process (Ho, Dey & Lockström, 2011) and might require additional training (Genovese et al., 2014). They are also highly context-dependent and might lead to ineffective results if data are missing (Talluri & Narasimhan, 2004). Literature has suggested more simple models, such as scorecards, that have been feasible in practical applications (Cao et al., 2022).

Current models in supplier performance evaluation have focused on supplier selection problems. In theory, these models could be extended to be used in supplier performance monitoring, although the fit of the measurements might be different for monitoring and supplier selection. Supplier monitoring might have more importance in established companies since they might have a stable supplier base, or the goal of monitoring is to reduce the number of suppliers (Luzzini et al., 2014). Even though these models exist and could be used, they require highly systematic data collection and eventually might be too extensive to use in real-life settings. More simple ways for supplier performance measuring analysis, such as a dashboard, might be more user-friendly. Irrespective of the decided analysis tool, companies

should have a structured supplier performance measurement system (SPMS) in place and understand how they should approach designing a SPMS to be able to gain advantages from supplier performance measuring.

## 1.1 Research gap

Hald and Ellegaard (2011) studied two cases that had designed and used a SPMS, and factors that influenced the design, implementation, and use of the system. Their study indicates that there are underlying dynamics that impact the outcome of supplier evaluation, such as how motivated suppliers are to improve their performance. Luzzini et al. (2014) studied how a supplier performance evaluation system can be designed through 13 cases, providing implications about the phases of the design. The studies are focused on the aftermath of designing and using such a system. Both studies are only focused on the outcomes of system design and usage, and they leave out the alternative ways to design, implement, and use the system, i.e., the focus is more on what is done but not what could or should be done.

Also, the constraints of data collection, use, and other factors that impact the design are not well understood. Supplier performance evaluation or monitoring has been mostly studied as a decision-making problem (Luzzini et al., 2014) leaving out all the data collection practices that make these models feasible to implement. In the data collection practices, the data quality might lead to the situation of “garbage in – garbage out”, where bad quality data will lead to an unreliable output of the system (Kilkenny & Robinson, 2018). Performance measuring should be about building good and realistic measurements, especially if the performance results are communicated to suppliers, and it is even more important if the performance measuring is used for performance-based contracting. For example, Hald and Ellegaard (2011) found that unrealistic measurements led to unmotivated suppliers, highlighting the importance of the fairness of the evaluation that stems from the data quality. The intended use has an impact on the design. Maestrini et al.’s study (2021) divided use into diagnostic and interactive, which both have a positive impact on supplier performance improvements. However, it has not been previously studied qualitatively, how the users would like to use SPMS.

To benefit from the supplier performance measurement systems, it has been found that communication of the results is important. It is not insignificant how the communication will take place. Suppliers might find the results unfair or not reflecting the actual situation (Hald & Ellegaard, 2011), suggesting that they will not want to act based on the measurements (Maestrini et al., 2018c). The importance of communication or so-called socialisation practices is shown to mediate the performance impact of the monitoring (Cousins, Lawson & Squire, 2008). In the design phase, a company can establish communication procedures for supplier performance results and prevent unwanted communication between buyer and supplier, e.g., unmotivating performance measuring.

Even though articles have identified what kind of design phases there are, they have not studied the design process before the implementation of a supplier performance measurement system, which provides valuable information about the company's struggles in the design phase. The performance measuring system design might have some constraints or opportunities that are not yet identified by literature. Also, the performance measurement systems are rarely designed for the whole supplier base, which is the aim of the case company in this study, making it an interesting case to study.

## **1.2 Research goal**

This thesis aims to study the design process of a manufacturing company and understand the underlying issues that a company might face when designing a supplier performance measurement system for supplier performance monitoring. The focus of the study is to identify how companies should approach the design of SPMS. Therefore, the following research questions are formed:

- 1. How to design a supplier performance measurement system?*
- 2. What are the key factors of a supplier performance measurement system?*

The main research question aims to capture the system from data collection to communication of the results. The first research question is focused on the design phase of such a system which is mainly answered with the literature review. However, as the literature review also provides the structure for the interview questions, it is anticipated that interviews raise issues

that should be considered during the design process. Therefore, the second question reflects more on the various factors that impact the design, implementation, and use of such system.

As supplier performance monitoring provides an additional tool for the sourcing function, the context of the case company has to be considered. As the goal of SPMS is not to have the most advanced system in place, but to serve the needs of the company, a single case study provides an excellent opportunity to study the phenomenon and identify factors that should be considered in other companies as well. On the other hand, a single case study limits the generalizability of the results. The case company has identified its strategic suppliers through a supplier categorization that will be considered during the interviews, e.g. if there are differences in the measurements for strategic suppliers. Case company is also a large organization in the manufacturing industry which provides rich context in terms of purchases.

### **1.3 Definitions**

Supplier performance evaluation is understood and used in various ways in the literature. The supplier is evaluated across multiple metrics to assess various performance aspects, such as sustainability, cost, delivery performance, or a combination of many metrics that are then called performance. Supplier performance evaluation is mainly used in literature to solve supplier selection from new or old suppliers (Patrucco, Frattini & Di Benedetto, 2022). Assessment is sometimes used as a term instead of evaluation in this context. The main difference from monitoring is the timing of the activities. Evaluation can take place once or periodically, for example, once before an event and a second time after the event ends. Monitoring happens continuously or more regularly.

Supplier performance monitoring can be understood in various ways as well. Sometimes, monitoring is referred to as a continuous practice, even in real-time. In the supplier performance evaluation context, when the assessment takes place twice, it might be referred to as monitoring (Araz & Ozkaharan, 2007). For example, when supplier development programs take place, supplier performance evaluation is suggested to take place twice – before and after the program, and it could be called monitoring. In this thesis, when the same supplier is assessed by the same or similar metrics twice, it is called monitoring and when the supplier

is assessed once, it is called evaluation. Also, when supplier performance is followed continuously, e.g., by the data that is available from ERPs, it is referred to as monitoring.

Measuring has some concepts that should be defined as well. In the literature, performance dimensions, e.g., delivery performance, are compiled from multiple criteria such as time from order to delivery, on-time delivery, and delivery reliability (Lee, 2009). Dimensions are sometimes referred to as attributes or main criteria. Criteria are sometimes called sub-criteria, metrics, and sub-factors.

Supplier performance evaluation and monitoring models do not function in isolation from other processes in a company. In literature, a performance measurement system needs to have two features: 1) performance measures, and 2) supporting infrastructure (Franco-Santos et al., 2007). The latter might include methods of data collection and information systems to support the measuring, or processes to deal with data, such as analysis (e.g., evaluation models), interpretation, or visualization. A supplier performance measurement system is referred to as a “set of metrics measuring the efficiency and effectiveness of suppliers’ actions and the goodness of the relationship with them” (Maestrini et al., 2017, p. 301). The practical implementation of the system might take place in an ERP module or vendor portals, or it might not be integrated into systems at all, such as a supplier performance spreadsheet in Excel (Luzzini et al., 2014). Supplier performance evaluation systems have been called by many names: Vendor evaluation system (Luzzini et al., 2014), Supplier performance measurement system (SPMS) (Maestrini et al., 2017), or just simply supplier evaluation (Hald & Ellegaard, 2011). In this thesis, the wording SPMS refers to the whole process from collecting the data to the interpretation practices and results communication with the supplier.

## 2. Theoretical background

In this section, the theoretical background for this study is discussed. As pointed out by Gioia, Corley and Hamilton (2013), the literature review for Gioia methodology should not aim to provide an exhaustive literature review, as the theory is supposed to be built from the analysis that uses the discussion between theory and data to build a new theory in an abductive manner.

Many concepts related to supplier performance are overlapping. For example, supplier performance evaluation can be seen as a critical step for strategic supplier relationship management. In literature, performance evaluation is often suggested to be evaluated by using mathematical models. However, some literature defines the process more fully, e.g., discusses an application of performance measurement system. The literature is focused on these performance evaluations, where it is seen as a multi-criteria decision-making (MCDM) problem. There is also a stream of literature about supplier performance measurement systems, where models can be used differently to eventually manage the performance of the supplier. Even though this thesis is not about the usage of these advanced MCDM models, they are discussed in the literature review, as they are about supplier performance evaluation or monitoring and are widely studied. Furthermore, these articles might have a case application providing insights about the usage of an (advanced) SPMS.

The first two sub-chapters are more descriptive in nature and discuss what supplier relationship management is, and how it is connected to strategy. As supplier performance measuring is a possible method for supplier relationship management, and eventually fulfilling strategic goals, it is important to understand how it is connected to strategic supplier relationship management as well. Supplier performance measuring does not work in isolation and therefore, the emphasis of the following sub-chapters is on supplier performance measurement systems, starting from lifecycle and ending in the communication chapter. As the literature around supplier performance measuring is focused on these advanced decision-making models, supplier evaluation and monitoring models are discussed in two chapters after the lifecycle chapter. After that, the visualisation of the results is discussed. It is not irrelevant how the results are communicated as they mediate the performance impact from the SPMS (Cousins et al., 2008). Therefore, the last chapter discusses communication practices.

## 2.1 Purchasing in company's strategy

Historically, purchasing was seen as a function that companies had to do, and the strategic value was not recognized (Ellram & Carr, 1994). Nowadays, the value of purchasing is becoming even more recognised, and the process is often divided into strategic sourcing and operative procurement. The division is shown in Figure 1. Strategic sourcing has a more long-term focus on planning supply, selecting suppliers, and contracting suppliers. Strategic sourcing aims to effectively manage the supply base by identifying and choosing the correct suppliers (Talluri & Narasimhan, 2004). Operative procurement contains more day-to-day activities, ensuring that the supply is ordered, delivered, and paid.

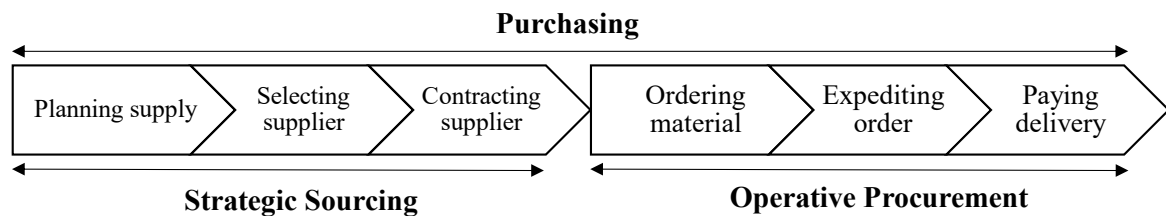


Figure 1. Division between strategic sourcing and operative procurement (Schiele, 2019)

When justifying the increasing strategic importance of purchasing, the percentage of revenue spent on materials and services is brought into the discussion (Emiliani, 2010). As purchasing is responsible for a big part of companies' costs, it is not irrelevant how the purchasing is planned and executed. However, to contribute to a company's profitability, purchasing should be integrated into the company's strategy (Ellram & Carr, 1994). Traditional goals of purchasing have been buying at the lowest possible cost and ensuring sufficient supply and quality. As purchasing is starting to reflect more company's strategy, the price-focused mindset is changing to consider other factors as well, such as sustainability (Pagell, Wu & Wasserman, 2010) and innovation (Luzzini et al., 2015).

Purchasing could be put into the company's strategy context by integrating it into the purchasing year, as suggested by Schiele (2019) in Figure 2. Suppliers are selected, contracted, and evaluated in the category sourcing cycle (Schiele, 2019). New suppliers emerge from the supply market and at some point, some suppliers do not fulfil the needs of the company

and are eventually phased out. Suppliers are not equally important to the company and therefore, the management of suppliers should be different depending on various factors, such as trust, the importance of products, and profitability. However, the supplier's point of view should also be considered, as the buying company might not be an important customer for the supplier (Mortensen, Freyta & Arlbjørn, 2008; Jääskeläinen, 2018). These aspects define the relationship that exists between the buyer and supplier. Even though some aspects might be given or hard to change, a company can influence at least some or most of the factors, making the supplier relationship management a necessity.

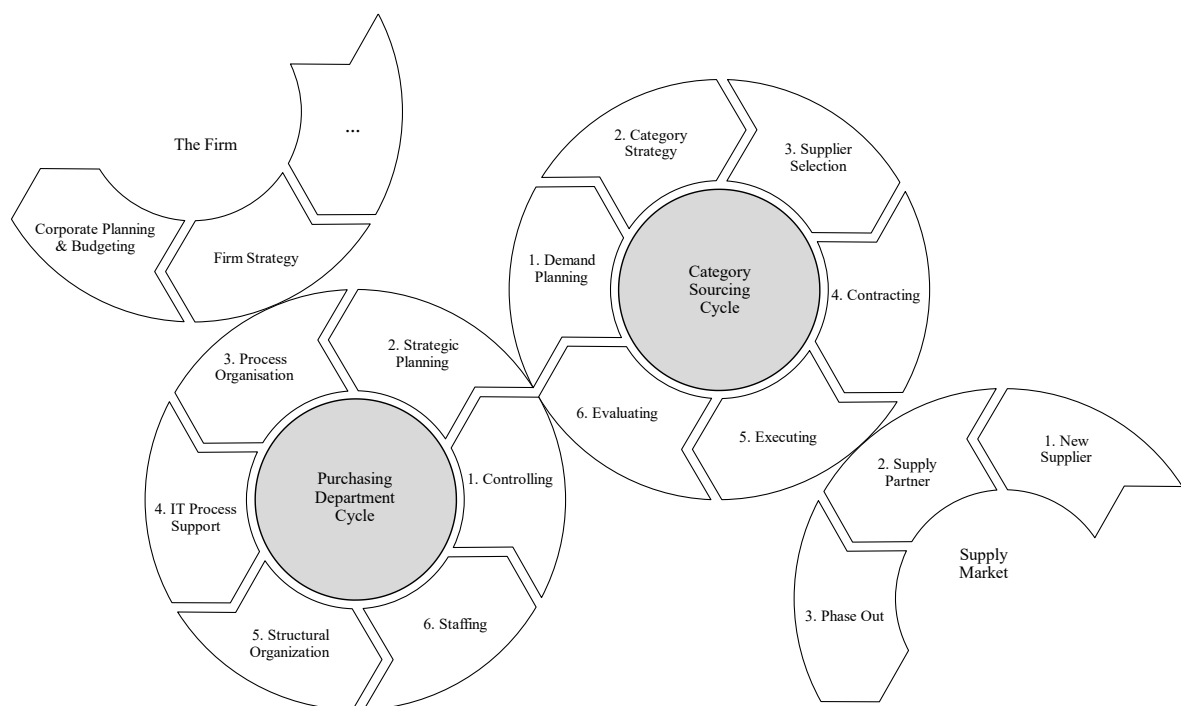


Figure 2. Purchasing cycles between the company and supply market (Schiele, 2019)

## 2.2 Supplier relationship management and supplier performance measuring

Supplier relationship management (SRM) is defined as a “business process that provides the structure for how relationships with suppliers are developed and maintained” (Lambert & Schwieterman, 2012, p. 337). Relationship management is a way to handle the supply chains of the company (Lambert & Enz, 2017). SRM can be approached in a structured manner that performance measuring supports. The process of SRM could be divided into two sub-processes: strategic and operational (See Figure 3). The goals of both processes might be



different. Lambert and Schwieterman (2012) suggest that SRM should focus on how to divide created value among suppliers and suggest that strategic relationship management should aim for defining guidelines to share process improvement benefits with suppliers. Operative procurement, on the other hand, would be responsible for implementing it according to guidelines. Schiele (2019) suggests that the operative side would only consider ordering and paying, and other actions would be considered on the strategic side. Park et al. (2010) suggest that the end goal of the SRM is continuous improvement of the supply base and suggest that the managers should be in charge of the SRM process.

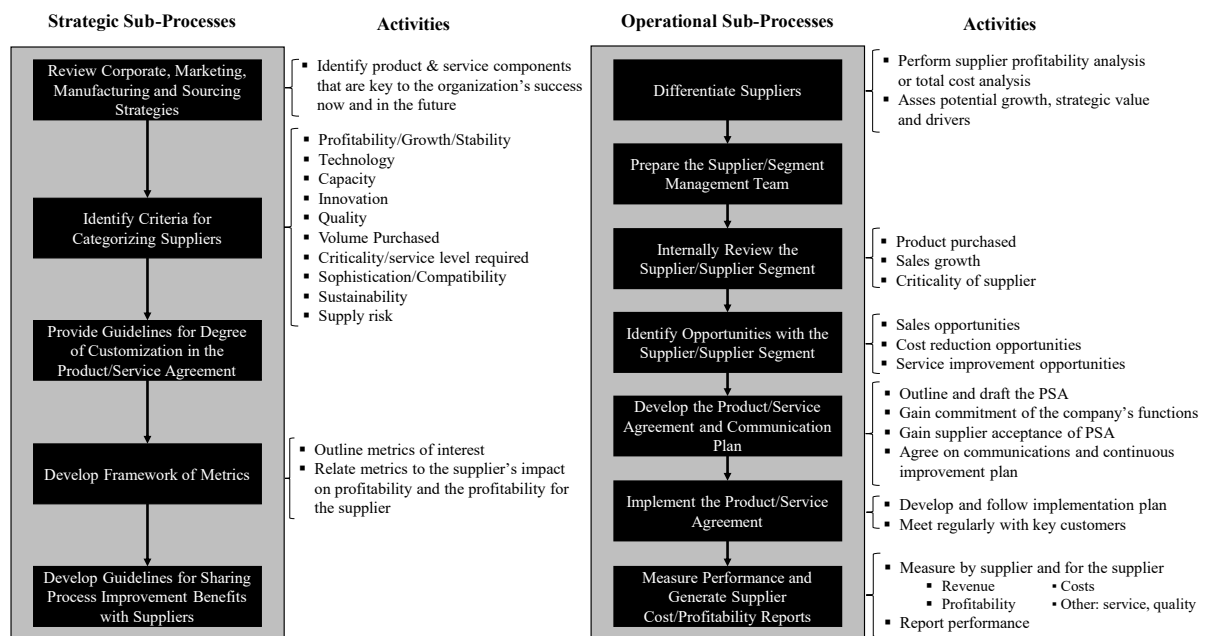


Figure 3. Strategic and operational supplier relationship management (Modified from Lambert & Schwieterman, 2012)

The first step in Lambert and Schwieterman (2012) strategic sub-process is to go through the company's strategies to be able to define which suppliers are critical for the company. The topics discussed could be, for example, future opportunities, threats, or risks. The next step would be to categorize suppliers by identifying important criteria to achieve strategic goals. Criteria vary across different product groups or needs of a team. The third step considers the development of guidelines on how much there is room for alternative treatments for different suppliers. For example, an important category might be allowed to share more

information with its supplier. The fourth step considers developing the framework of metrics that will be used to measure the performance of suppliers. Metrics should be aligned with the strategic objectives of other functions as well. The fifth step would be the guideline about how process improvement benefits should be shared with suppliers. (Lambert & Schwieterman, 2012) Operational SRM is the implementation of SRM according to the guidelines.

In SRM, the categorisation of suppliers might act as a foundation for which suppliers are critical, i.e., strategic, and how suppliers should be treated. To identify which suppliers are critical for supplier relationship, the company could quantify different variables to understand how the supplier should be managed. One widely used tool for supply management is Kraljic's matrix (1983), which divides products based on supply risk and profit impact into four different categories: leverage, strategic, non-critical, and bottleneck items. One way to use this matrix is to develop supplier relationships according to the categorization. Park et al. (2010) propose supplier relationship management strategies by evaluating the material and supplier performance. The model suggests the categorization of suppliers into either improvement, maintenance, collaboration, or prime categories and the result could be used to identify suppliers for development programs.

Supplier evaluation supports formation of supplier relationship management strategies. For example, Lima-Junior and Carpinetti (2016) propose a model that uses performance measurements as a base, and the importance of these measurements is based on decision-makers' opinions. The metrics are further aggregated into two performance dimensions: cost and delivery performance. Based on the results, the supplier is categorised across four categories which suggest a strategy for that supplier. The criteria for supplier evaluation, the performance dimensions, and the classification matrix with action plans are shown in Figure 4.

It should be noted that supplier evaluations should not be the only factor defining SRM, even though it is suggested as a necessary step for SRM. For example, Cannon and Perreault (1999) found that 3 % of supplier relationships that were categorized as "mutually adaptive" or "customer is king" were less than two years old. This suggests that supplier relationships should be given time to evolve to gain importance in the supplier's mind. Newer relationships might be overlooked in supplier performance measuring, as the focus might be on historic data, thus leading to a situation where new or recent suppliers are not evaluated or

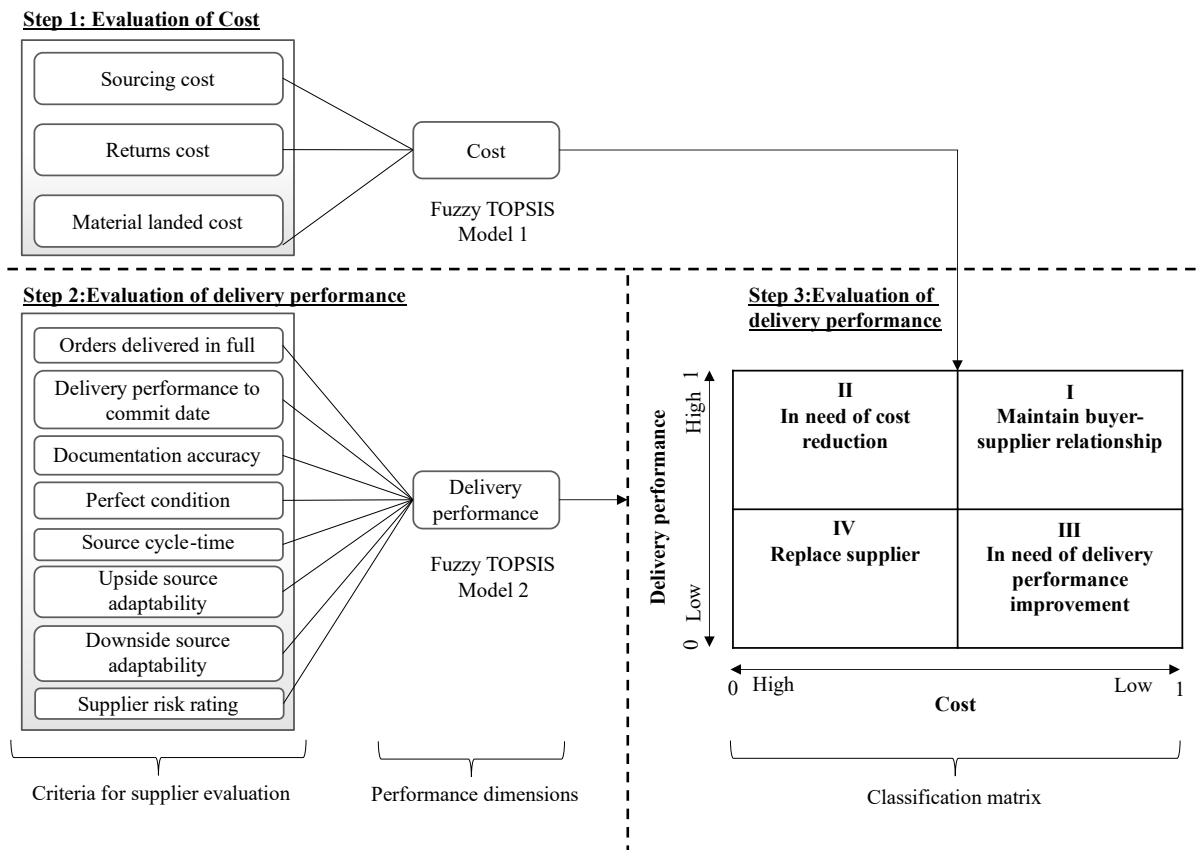


Figure 4. Supplier categorisation based on performance and suggested strategies (Lima-Junior & Carpinetti, 2016)

monitored due to lack of performance information, inhibiting the development of good relationships in the first place. Apart from that notion, the supplier relationship management suggests developing differing measuring for strategic suppliers.

### 2.2.1 Tailoring performance measuring for strategic suppliers

Having a supplier monitoring system at place might enable cooperation relationship with suppliers which eventually benefits the company's performance (Carr & Pearson, 1999). Some level of differentiation in SPMS across suppliers might be beneficial due to differing nature of products and services. As SRM suggests, not all suppliers are made the same. The same applies to supplier performance measuring. One possible way to diversify supplier performance measurement is to focus on strategic suppliers, enabling the focus shift from performance measuring to performance management. Focusing only on strategic partners is suggested to be the most advantageous for collaborative supplier performance measuring (Jääskeläinen & Thitz, 2018).

There are various ways to tailor the process for strategic suppliers. Lambert and Schwieterman (2012) reflect, through strategy and segmentation, which suppliers will be monitored and suggest a monetary value-based (economic value added) approach to identify best alternative measurements. They suggest that the SRM's impact should be identified, and measures developed accordingly. The economic value-added model and these quantifiable SRM impacts are shown in a Figure 5. In their model, economic value-added monitoring is used to share the added monetary value with strategic suppliers.

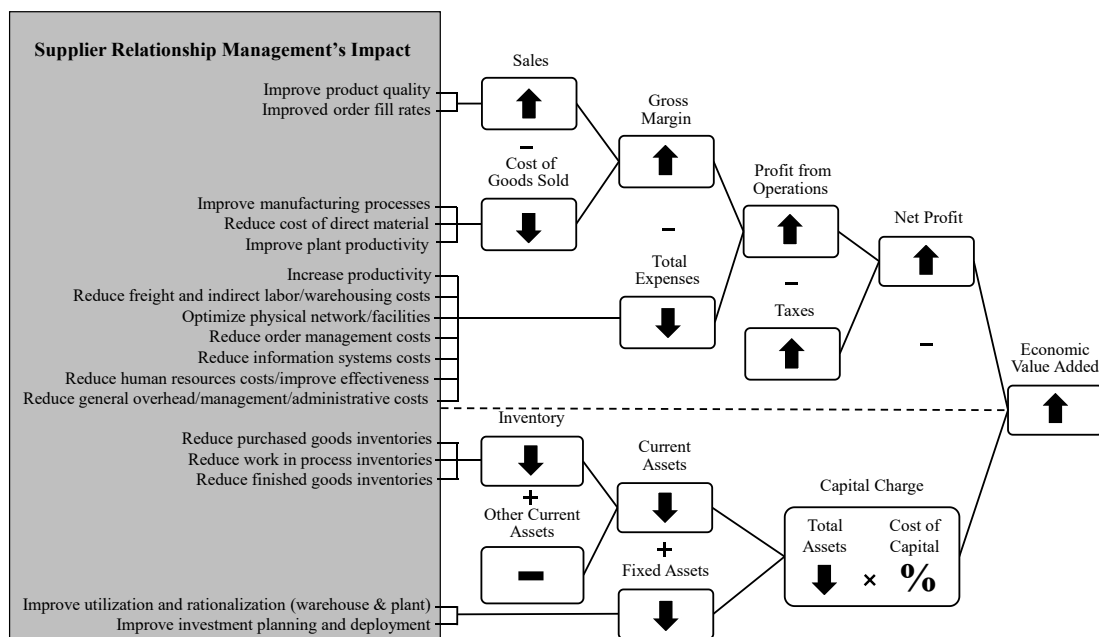


Figure 5. Supplier relationship management's impact and their impact on economic value added (Lambert & Schwieterman, 2012).

For supplier performance measuring, measurements could be divided into strategic and operational performance dimensions (Nair et al., 2015). Operational metrics are, for example, cost, quality, delivery, and flexibility. Suppliers' strategic metrics are product development capability, technology management capabilities, investment capability, and global supply capabilities. Nair et al. (2015) suggest that a supplier, that is chosen and monitored based on operational criteria, affects the purchasing performance in cost, quality, delivery, and flexibility dimensions. On the other hand, strategic performance evaluation of suppliers did not affect delivery and flexibility dimensions. To make the measuring effective, Hawkins et al. (2020) suggest that measurements should be defined by sourcing teams, including how criteria are measured and ensuring that the thresholds of the values are unambiguous. These

guidelines should be provided to suppliers, e.g., in contracts to improve the justification of measuring.

Traditionally, supplier performance measuring has been around these financial metrics (Choy et al., 2007). Purely financial metrics are not sufficient for supplier performance measurement, as they do not i) provide forward-looking perspective, ii) relate to strategic performance, or iii) connect effectiveness to efficiency (Bullinger, Kühner & Van Hoof, 2002). Integrating both non-financial and financial metrics has been found to allow learning and dialogue to take place (Dossi & Patelli, 2010). Another way to divide measurements is based on whether the measurements are lagging (reactive), or leading (proactive). Example of lagging measurements are financial metrics, delivery and quality performance. Leading measurements (e.g., risk management, organizational, environmental and social practices) might help buyers to improve their supply chain's performance in terms of quality, delivery and costs (Dey et al., 2015). Performance measuring should take a more future-oriented viewpoint and therefore, leading metrics are suggested to be used for strategic suppliers. Leading metrics might provide some examples of best practices (e.g., superior risk management or social practices) and these might be beneficial to benchmark in other companies.

There is a strong emphasis on positioning the company's strategy as the source for the best criteria. This might not reflect the daily life in companies. For example, Kotula et al. (2015) found that even though strategic goals, such as improved quality, were found to be a critical success factor for case companies, the most important criteria for strategic supplier evaluation and selection were price or cost related, such as riskiness and supplier's performance, suggesting that there was a misalignment between strategy and strategic supplier evaluation. On the other hand, Jääskeläinen (2018) found that during development of measurements, the significance of costs was acknowledged in companies, but not emphasised. Costs cannot be dismissed completely, but if the purpose of the performance measuring is to bring improvements to supply chain, it is not enough to capture the complexity of the performance (Bullinger et al., 2002). Performance should not be approached only through its financial dimension, as non-financial performance might be even more important if, for example, customers justify buying the company's products or services based on value (e.g., superior quality or sustainable products) and not just price.

As sustainability has been gaining importance in companies' strategies, involving sustainability targets to supplier performance evaluation is natural. However, it is not as straightforward as it sounds. It has been suggested that companies should especially follow suppliers' compliance with laws and regulations (Chiarini, 2017). Also, ISO 14001 is an important certificate to have when trying to improve a supplier's environmental performance, according to their study. If the environmental impact of suppliers is quantified and measured, it needs to consider suppliers, because if only the first-tier suppliers are measured, they might have outsourced polluting parts of their production (Genovese et al., 2014). When improving sustainability, the supplier must be willing to disclose information about its emissions and consumption during its processes. Information disclosure is important for trust and transparency between the buyer and supplier. Therefore, it is a critical criterion for sustainable supplier selection (Govindan et al., 2021) and a prerequisite for monitoring.

Alternatively, for different goals, e.g., sustainability or risk management, supplier performance measuring can be based on the relationship or supplier characteristic. The measuring might include purchase characteristics as well. One of the ways to distinguish measuring is based on what is purchased, e.g., how critical the purchased product is. There is a difference between supplier monitoring for the maintenance of the company headquarters' sauna and the maintenance of a factory's production line. The study of Jääskeläinen and Thitz (2018) suggests that collaborative measurements might be the most beneficial for project-based businesses due to the increased need for discussions and communication between buyer and supplier. Their study suggests that the starting point of these collaborative measurements is the joint targets between buyer and supplier. One of the possible methods to distinguish between these is to use portfolio models such as a purchasing cube. All in all, measuring is a balancing act between standardization and accuracy, i.e., how well the measuring reflects the purchased product, relationship, and supplier.

### **2.3 A supplier performance measurement system lifecycle**

The roots of supplier performance measurement systems lie in performance measurement system literature. Neely, Gregory, and Platts (1995, 81) defined a performance measurement system (PMS) as a "set of metrics used to quantify both the efficiency and effectiveness of actions". PMS could be evaluated through three different levels. The first level refers to

individual performance measures, which is examined by analysing what is measured and whether the measurements offer any benefits. The second level of PMS is the system as an entity and is also referred to as the set of performance measures. To analyse the appropriateness of the system, one should analyse if it contains relevant information and if the goals have been defined. At the highest level, the PMS entity is studied in the organizational environment. Assessing the suitability of the system, one should consider if it is fit for the organizational environment, such as if it is aligned with the strategic goals and organizational culture. (Neely et al., 1995) To analyse the performance measurement system, Neely et al. (1995) propose the following questions that are listed in Table 1.

Table 1. Performance measurement system analysis (Neely et al., 1995)

Individual measures	Performance measurement system	The Environment
What performance measures are used?	Have all the appropriate elements (internal, external, financial, nonfinancial) been covered?	Whether the measures reinforce the firm's strategies?
What are they used for?	Have measures which relate to the rate of improvement been introduced?	Whether the measures match the organization's culture?
How much do they cost?	Have measures which relate to both the long- and short-term objectives of the business been introduced?	Whether the measures are consistent with the existing recognition and reward structure?
What benefit do they provide?	Have the measures been integrated, both vertically and horizontally?	Whether some measures focus on customer satisfaction?
	Do any of the measures conflict with one another?	Whether some measures focus on what the competition is doing?

Similar questions could be applied in SPMS context. Genovese et al. (2014, p. 1199) constitute the supplier evaluation problem from two constructs: "the definition of models and methods to analyse and measure the performance of a set of suppliers (vendors) on a set of dimensions (criteria) in order to improve customer competitiveness". To be able to evaluate suppliers' performance, the company should have structured SPMS in place (Bruno et al., 2012). Figure 6 depicts the cyclical nature of SPMS development. Hald and Ellegaard (2011) suggest three phases for the development of a supplier measuring system. Their first step considers design. In this step, the key objectives of measuring are defined, and measurements are selected. The second step is the implementation of the system that defines the procedures and makes sure that data can be collected regularly. The third step is the use. Supplier's

performance is assessed, and action plans are done based on the assessment. Maestrini et al. (2018a) suggest a fourth step, which is review.

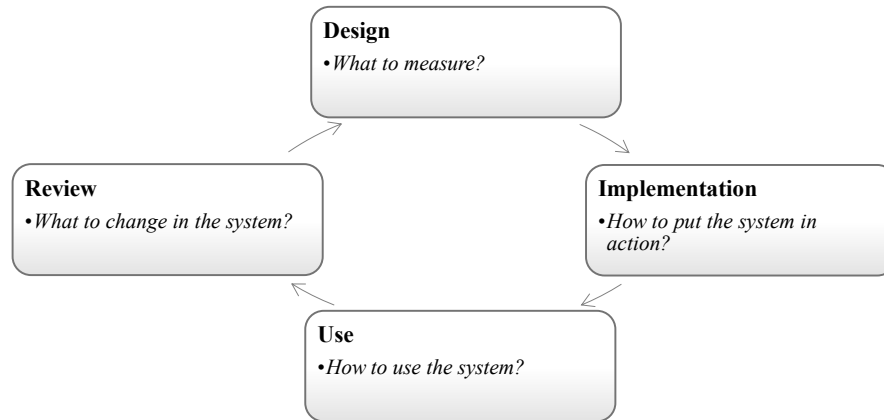


Figure 6. Phases for developing a supplier performance measurement system and aims of each phase (Maestrini et al., 2021)

It is noteworthy, that phases might overlap depending on the applied process, e.g., some measurements are implemented before designing others, suggesting a non-linear progression for the measurement system (Bourne et al., 2000). To provide structure to the review, the suggested lifecycle phases are used. The design phase is discussed first.

### 2.3.1 Design phase of a supplier performance measurement system

Design phase could be defined as a step, where the identification of relevant objectives, e.g., use and planned measurements, takes place. The successfulness of SPMS depends on the design phase as it will establish how well the implementation phase will work. Luzzini et al.'s (2014) study focuses on the design phase of a supplier performance measurement system. They further divide the design phase into three key components, which are shown in Figure 7 and discussed further in the following paragraphs.

The strategic alignment is divided into four components: i) system objective, ii) commitment from top management, iii) Units involved in SPMS design/execution, and iv) KPI and weight definition. In Luzzini et al.'s (2014) study, companies reported system objectives to be, for example, supply base reduction, supplier development, and supplier scouting. The system



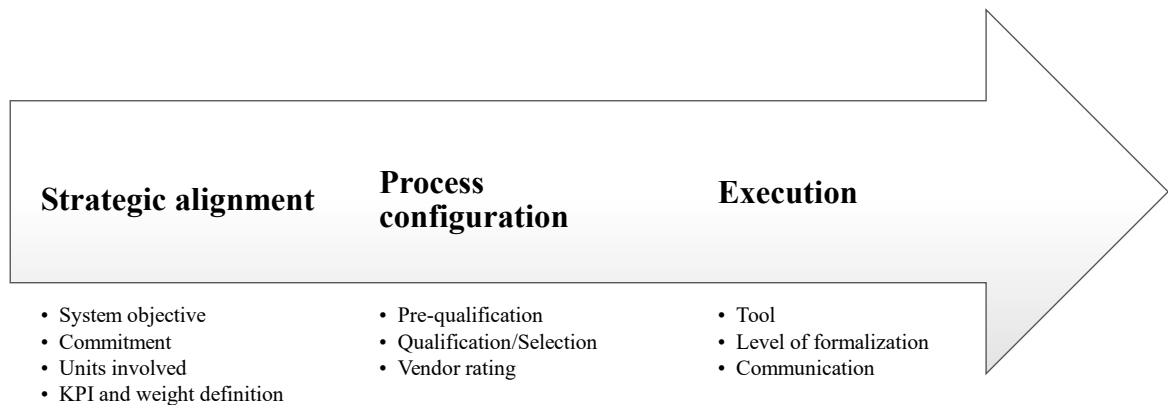


Figure 7. Supplier performance measurement system design phases (Luzzini et al., 2014)

objective was centred around process standardization through more formal and defined supplier management processes and evaluation. (Luzzini et al., 2014). However, the realised benefit might be different from the objective, suggesting that they are not necessarily identified. Standardizing a SPMS is a challenge if there is an extensive number of suppliers. The process cannot be the same for a supplier that provides a critical, highly technical product and a supplier that provides standard products (Bruno et al., 2012).

The first step in the SPMS design process is identifying key criteria in a company (Luzzini et al., 2014; Hald & Ellegaard, 2011). Companies can approach this problem by identifying suitable criteria from academic literature and practitioners (Genovese et al., 2014). Supply Chain Operations Reference (SCOR<sup>®</sup>) model's metrics are widely adapted in supplier evaluations (Lima-Junior & Carpinetti, 2016; Genovese et al., 2014). Mainly the identification of relevant KPIs is suggested to be derived from the strategy (Neely et al., 1995) or from the company's decision-makers (Genovese et al., 2014). Chosen KPIs can either be defined independently (e.g., who holds the area of expertise designs) or have shared design (e.g., the units try to reach consensus about KPIs) (Luzzini et al., 2014). Literature suggests these multi-criteria decision-making models for choosing and weighting KPIs (Ho et al., 2010; Dutta, Jaikumar & Arora, 2021). These models are discussed more in their own sub-chapter.

It is not easy to define correct KPIs, as the unit of measurement influences the results, and measurements might influence each other in unanticipated ways, adding to the complexity of measuring. For example, in logistics, the question could be whether the delivery performance is assessed on delivery, pallet, or item level (Thunberg & Persson, 2014). If they

evaluate delivery reliability, the results might vary depending on the unit of measurement, e.g., is a delivery late if one item is missing from the pallet. Some performance attributes have a relationship with each other. If a company wants to have high sourcing flexibility, i.e., they can make changes to their orders, it might have a damaging effect on cost performance, if not planned carefully. There is no linear relationship, but a curvilinear relationship, between sourcing flexibility and delivery performance, suggesting that companies should aim towards either high or low sourcing flexibility (Wagner, Grosse-Ruyken & Erhun, 2018). Sourcing flexibility should be considered early on during supplier selection, as it is not necessarily cost effective to ask a supplier to perform flexibly afterwards.

The main purposes of SPMS could be roughly divided into three different processes according to Luzzini et al. (2014). From the system perspective, SPMS can be used to collect general and prequalifying data about suppliers. Also, it can be used for collecting critical data about a supplier for supplier qualification or selection. The third process is supplier rating, where the supplier is evaluated through continuous and comprehensive evaluation practices. Companies might use the SPMS for all three different purposes that require the definition of their own processes.

The execution step, which will define how the system is planned to be managed and used, is divided into three components: i) supporting tool, ii) level of formalization, and iii) communication (Luzzini et al., 2014). Supporting tool refers to IT-infrastructure, e.g., non-integrated platforms, ERP, or vendor portals. The level of formalization refers to how defined the use is, for example, whether the procedures are strictly defined or not. On the other hand, defining the process does not mean that it is enforced. The third and most important step is to design the communication practices, as the supplier's operation might be improved according to the results. As suggested by Hawkins et al. (2020), supplier evaluation can be even weaponized by buyers, who might threaten the suppliers with bad ratings. Harmful practices might lead to decreased benefits from supplier measuring, as they produce unreliable information.

### **2.3.2 Implementation phase of a supplier performance measurement system**

In essence, implementation of performance measurement might refer to capturing already existing data, processing it, and forming visualisations that present information. The implementation phase considers “which systems and procedures are put in place to collect and process the data that enable the measurements to be made regularly” (Bourne et al., 2000, p. 758). Implementation of SPMS might include developing new data collection procedures and other supporting factors, e.g., reporting practices. Communication and reporting factors are discussed in their own subchapter. The implementation phase influences, for example, how resource intensive the SPMS is (e.g., is the process automatic, semi-automatic, or manual) (Hald & Ellegaard, 2011). It could be argued that the implementation phase could be a technical description of how the system should be produced, e.g., on the programming level. However, as the focus is more on the managerial level, the focus of this subchapter is more on best practices or identified influential factors.

Data collection practices are important to define to gain information from the SPMS. Quantitative data is available from ERPs and does not necessarily need new processes to collect it. On the other hand, qualitative data is another matter, as one of the ways to produce performance information about suppliers is to i) evaluate them or ii) ask them to self-evaluate. How it should be conducted depends on the purpose of the measuring. For example, for supplier selection, the data needs to be collected only once, and this is usually done via a form that is filled out by the supplier. Information collection might be indirect (ask the supplier to provide performance information) or direct (the supplier is being evaluated by the buyer, e.g., by audits). The data collection might be different for every case. For example, one might want to do an audit for a problematic supplier to put pressure onto the supplier to change the processes (Purdy & Safayeni, 2000). The physical presence of the evaluator might give the buying company a priority in the supplier’s processes as well. Evaluation might be a qualitative or quantitative survey, or open text based.

Qualitative data collection has other factors influencing it as well. For example, the main question is – Who evaluates? Supplier or buyer? An employee from operative procurement or strategic sourcing? There are various approaches to consider. Pal and Kumar (2008, p.

394) commented on how the evaluator influences the rating: "... ratings of course, is dependent on evaluators state-of-mind, whims and wits, knowledge and thoroughness about the organization's functionality, bias and vested interests." They suggest that if it is possible, there should be a group of evaluators. The evaluator, or group of evaluators, has an impact on evaluation (Buffa & Ross, 2011). As pointed out by Purdy and Safayeni (2000), supplier evaluation models usually disregard the social and organizational bias present in data collection, and Hawkins et al. (2020) identified that evaluators might give good ratings to avoid disputes with suppliers. Hald and Ellegaard (2011) found that buying personnel might continuously give bad ratings to suppliers, even though they were satisfied with the outcome to pressure them. Steward et al. (2010) found that there were national differences in what is expected from the supplier. Also, if supplier performance is assessed on technology-related performance, the person who is interpreting the data should be able to understand what good performance is in this context. When considering sustainability, companies might have different units in how they measure environmental metrics (Genovese et al., 2014). Evaluators should carefully consider their bias when evaluating. Suppliers might seem overly helpful and cloud the judgment of the evaluator (Clauss & Tangpong, 2019).

The data collection practices should be well defined if suppliers are asked to evaluate themselves. For example, in Genovese et al.'s (2014) study, suppliers evaluated themselves, and the questionnaire was too open, resulting in unusable data. In a study by Talluri et al. (2013), suppliers self-evaluated their capabilities, while purchasing personnel evaluated their performance outcomes. Both were done as questionnaires. Faraz et al. (2018) suggest sending questionnaires, that contain evaluations based on the Likert scale, to suppliers. Suppliers that evaluate themselves might want to hide their bad performance and give good ratings (Hawkins et al., 2020).

Frequency of measuring will also depend on what is being measured. Data from ERPs could be assessed daily, but qualitative matters are another thing. Wu and Blackhurst's (2009) case company collected data quarterly from suppliers with a survey that contained 36 questions. The themes used in the questionnaire are shown in Table 2 to show possible metrics for data collection. Other studies suggest that evaluation should take place periodically, from once in a month up to once in a year (Luzzini et al., 2014). For supplier performance rating, the data collection might depend on, for example, how standardised the delivered products are.

If the product is repetitive direct material, the data collection might be monthly, if the methodology for evaluation is standardised (Luzzini et al., 2014). Having frequent evaluations will influence the workload that employees and suppliers are experiencing.

Table 2. Supplier questionnaire themes for supplier evaluation (Wu & Blackhurst, 2009)

Questionnaire group	Questions related to
Basic supplier information	Information on locations, sales information, labour relations, & pending acquisitions
Material sourcing information	Supplier certification and sourcing agreements
Quality information	Reject rates, quality control procedures
Operations information	Process capabilities, lead time information and information exchange capabilities
Engineering information	Design capabilities, partnership programmes
Total cost information	Information on raw material, labour, overheads and margins

When building measurements or models, e.g., supplier performance rating, one should consider the minimum allowed values for the performance criteria. In the supplier selection case, a supplier might obtain a low-performance valuation in critical criteria even if they were being proposed as the best alternative by the evaluation model (Bai et al., 2019). Similar results were found by Talluri and Sarkis (2002). However, they found that suppliers might obtain the best evaluation score from monitoring by only focusing on a few criteria that were found important by suppliers. In the end, the supplier might be underperforming in a way that is harmful to production, but this might be ignored by the data. It is not merely a problem for advanced methods, but simple KPIs that are produced by weighting various variables. For example, a performance index might have this issue as well. To address this issue, Rezaei et al. (2022) use constraints, that a supplier, who has performed worst in a criterion, cannot outrank a supplier who has not scored the lowest in any criteria.

### **2.3.3 Use and review phases of a supplier performance measurement system**

The use of SPMS refers to collecting, reviewing, and acting upon the performance data (Hald & Ellegaard, 2011). Even though SPMS could have varying purposes, e.g., supplier selection

and supplier development, the use is divided into two broad categories: *diagnostic* and *interactive use* (Henri, 2006; Maestrini et al., 2021). The diagnostic use is a more traditional purpose of PMS referring to controlling, that ensures strategic goals are achieved (Henri, 2006). It is achieved by measuring variables that follow the successfulness of predefined goals. Eventually, diagnostic use provides a direction, i.e., informing the supplier of what is to be expected. The interactive use of SPMS provides a platform for dialogue, identifying new opportunities and potential for win-win situations (Maestrini et al., 2021). Interactive use might be more recommended for strategic suppliers and long-term oriented measurements, such as innovativeness and sustainability (Maestrini et al., 2021). Both uses can co-exist in the SPMS. Due to their differing objectives, diagnostic and interactive SPMS can cause dynamic tensions (Maestrini et al., 2021). Table 3 shows a comparison of these two uses.

Table 3. Differences between diagnostic and interactive use (Henri, 2006; Maestrini et al., 2021).

	Diagnostic use	Interactive use
Objective in SPMS	Aligning supplier behaviour with buyer's purchasing strategy	Supplier involvement
Goal setting	Intended, pre-established	Emerging through dialogue
Use	Mechanic, controlling, and tracking; measurement tool	Opportunities seeking, solving potential conflicts mutual adjustments and improvements; a strategic management tool
Means	Following measurements, identifying deviations in performance, improvement plans, providing explanations, penalties, or incentives	Allowing communication to happen, proactive engagement, and establishing joint improvement plans
Approach	Negative, mistakes and variations	Positive, learning, and improving

Diagnostic use is more focused on simply monitoring, tracking goals, and using it to identify non-performance or other problems, and asking suppliers to solve these problems. One example of this use is supplier selection cases, where supplier performance might be monitored and evaluated, but information is solely used for internal decision-making or maybe sending the results to the supplier. Effectively, the use is diagnostic if the encouragement for open discussion is missing and common relational issues are ignored (Maestrini et al., 2021). Nevertheless, diagnostic SPMS use has a positive impact on supplier performance improvements

(Maestrini et al., 2021). Specifically for improving supplier quality performance, monitoring the performance of suppliers might be key (Prajogo et al., 2012). All in all, diagnostic use might be more suitable for more objective measurements, such as operational dimensions (Maestrini et al., 2021), eventually enabling the identification of issues that might provide best outcome for improvement.

For more interactive use, one could integrate buyer-specific metrics into the system, allowing a bi-directional exchange of information and providing a platform for discussion. Also, it is important to note that the supplier's performance is not solely dependent on the supplier, as the buyer-supplier relationship forms a dyad and eventually, the buyer's performance influences the supplier's performance. In Ross and Buffa's study (2009), they utilised data regarding purchase order stability, that measures the percentage of orders, where the buyer made no changes. They also consider how accurate and comprehensive the buyer's forecast was. The same metrics were used in the following study by Ross, Kuzu, and Li (2016). They found that timely and accurate information sharing by the buyer allows suppliers to perform significantly better. Suppliers, however, might be different in the sense of how their performance reacts to these differences in buyer's information sharing. The interactive use has been suggested to be a great tool for managing supplier's performance in project-related purchasing, such as new product development (See Le Dain, Calvi & Cheriti, 2011 for a framework, Patrucco et al., 2022 for a project based SPMS lifecycle).

The use of supplier performance measuring might be influenced by relationship factors as well. If the buyer is not an important customer to the supplier, they might disregard the results from the measuring all together (Maestrini et al., 2018c). The supplier relationship might be defined by market-related (availability of alternative suppliers and supply market dynamism) and situational (importance of supply and complexity of supply) determinants (Cannon & Perreault, 1999). Eventually, the relationship influences how supplier performance can be evaluated, due to willingness to share information. The buying company might not have direct product information from the suppliers, and the assessment might rely on indirect sources or assessing process performance directly, by audits for example (Purdy & Safayeni, 2000). Over time, the relationship might evolve, and the evaluating company does not need to assess the supplier as closely as in the beginning. Eventually, the assessment of performance might be more focused on indirect sources of information (Purdy & Safayeni,

2000). Using relationship metrics in SPMS might bring alternative benefits as well. It has been suggested that relationship-related metrics are good for supplier performance evaluation, as they can be used to assess if the supplier is willing to give preferential treatment (Ho et al., 2011). It is not straightforward how the relationship is defined, and which buyers does the supplier see as important. For example, in Maestrini et al.'s (2018c) sample, one of the company's sales managers said that even though they are a multinational company, customer satisfaction is important to their strategy and therefore, they assessed almost all the performance results from their customers.

A SPMS should reflect the buyer-supplier relationship characteristics during the development of the system, as previously discussed. Ultimately, the use of SPMS has an impact on the buyer-supplier relationship, that should be taken into consideration. Diagnostic use decreases the level of trust between the buyer and supplier (Maestrini et al., 2021). The emphasis on control might eventually mean that the supplier is required to obey the buyer's demands, leading to mistrust and opportunistic behaviour. On the other hand, trust positively mediates the performance improvements from interactive use, suggesting that interactive use fosters an environment of trust (Maestrini et al., 2021). All in all, diagnostic use of SPMS does not mean that there are fewer benefits *per se*, but it might lead to decreased trust, that is eventually harmful for SPMS benefits. Thus, it should be communicated to the supplier that the performance measuring aims for mutual benefits, and it should also be emphasised to internal employees that SPMS is not an exercise for power, but a means to "collect objective and clear information about supplier performance, to drive future improvements" (Maestrini et al., 2021, p. 1257).

The fourth phase in the development of SPMS is reviewing the SPMS. In a broad sense, the last phase refers to periodically assessing if the SPMS should be changed or improved. SPMS should be reviewed and updated from time to time to ensure that the SPMS is still aligned with the company's strategy. It was found that periodical review practices improved the supplier's performance across delivery, innovation, and sustainability dimensions (Maestrini et al., 2018a). Predefined review plans might be beneficial as management gathers to review the system as a whole (Bourne et al., 2000). Also, having periodical reviews helps to answer the current environment's ever-changing demands.



## 2.4 Supplier performance evaluation

As previously discussed, SPMS forms the frame where different supplier evaluation models can be applied. Supplier evaluation is defined as “the process of quantifying the efficiency and effectiveness of supplier action” (Hald & Ellegaard, 2011, p. 890). Supplier performance evaluation models have varying purposes, such as supplier selection or simply assessing performance. They are also be used to rank suppliers or quantify supplier performance by combining qualitative and quantitative criteria (Tsai, 2009). Supplier evaluation can be used to categorise suppliers (Lima-Junior & Carpinetti, 2016). The supplier evaluation models discussed in this chapter are more focused on a single case of evaluation, while continuous supplier monitoring is discussed in the next chapter.

It is challenging to decide the most influential factors on which to base the evaluation. The information that is used to assess suppliers might be imprecise, and the importance of measurements is dependent on the preferences of decision-makers (Chen et al., 2006), eventually having an impact on the supplier performance evaluation (Ross & Buffa, 2009). Therefore, supplier evaluation models usually propose a methodology that are used for identifying weights for different metrics. For example, it is simple to rank a few metrics based on importance, but ranking alone is not enough to solve the complexity of these problems. Decision-makers might see some variables as important or relatively more important than other metrics and so on. It can be argued that importance varies across employees. For example, the quality manager and procurement might have opposing views on importance. Therefore, it is suggested that the importance of metrics is a multi-criteria group decision-making problem. To assess the group’s opinions, a fuzzy ELECTRE III is suggested (Shen, Xu & Xu, 2015). The analytical hierarchy process (AHP) uses pair-wise comparisons for identifying weights for metrics. Muralidharan, Anantharaman, and Deshmukh (2001) suggest using AHP combined with confidence intervals to identify variability in decision-makers’ opinions. Similarly, TOPSIS is used to identify the optimal solution for measurement weights (Chen et al., 2006). To make the measurements reflect the subjectivity of decision-making, they use fuzzification to be able to use expressions rather than only use numbers. Models can be also combined. Parthiban, Zubar, and Katarak (2013) propose a model that uses a fuzzy SWOT analysis for supplier prequalification and DEA for final supplier selection.

The possible metrics for supplier evaluation are well understood in supplier evaluation literature (Talluri, DeCampos & Hult, 2013). Supplier evaluation models usually integrate various measurements into the decision-making. These measurements are either qualitative or quantitative. Some metrics are binary in their nature. The most used criteria are quality, delivery, price/cost, manufacturing capability, and flexibility (Ho et al., 2010). Other metrics are, for example, environmental impact, relationship, expertise, and financial condition. The importance of the criteria depends on various factors, such as the criticality of the product. Criteria and models for supplier evaluation might differ in what is used for supplier monitoring (Tsai, 2009).

There are different modelling approaches for supplier evaluation and some existing papers are show in Table 4. Even though a plethora of alternative models exist, the results might be the same. A fuzzy TOPSIS, a fuzzy VIKOR, and a fuzzy GRA were tested in a supplier selection problem, and the results were identical (Banaeian et al., 2018). However, models vary on how easy or time-consuming they are to use (Guneri & Kuzu, 2009).

Table 4. Existing performance evaluation articles.

Goal & main technique(s)	Main measurements	Reference
Supplier selection based on two years of performance data <i>Bayesian estimators</i>	<ul style="list-style-type: none"> <li>• Product Quality,</li> <li>• On-time delivery,</li> <li>• Cost Performance,</li> <li>• Supplier reliability</li> </ul>	Arnoğlu, Sarkis & Dhavale (2021)
Supplier selection based on social sustainability <i>A grey BWM &amp; TODIM</i>	<ul style="list-style-type: none"> <li>• Work health and safety,</li> <li>• Training education and community influence,</li> <li>• Contractual stakeholders' influence,</li> <li>• Occupational health and safety management system,</li> <li>• The interest and rights of employees,</li> <li>• The rights of stakeholders,</li> <li>• Information disclosure,</li> <li>• Employment practices</li> </ul>	Bai et al. (2019)
Supplier categorization in efficient/inefficient, selection <i>Fuzzy SWOT, DEA</i>	<ul style="list-style-type: none"> <li>• Strength/weakness criteria: Quality, Delivery, Productivity, Service, Costs,</li> <li>• Opportunities/threats criteria: e.g. Technological capabilities, Application of conceptual manufacturing, Environment management</li> </ul>	Parthiban et al. (2013)
Supplier selection <i>Fuzzy TOPSIS</i>	<ul style="list-style-type: none"> <li>• Profitability of supplier,</li> <li>• Relationship closeness,</li> <li>• Technological capability,</li> <li>• Conformance quality,</li> <li>• Conflict resolution</li> </ul>	Chen et al. (2006)

Apart from strictly mathematical models, other approaches exist as well. Supplier performance evaluation can also be used to assess technological competencies. Kronemeyer, Kotzab, and Moehrle (2022) suggest using semantic analyses to analyse suppliers' and buyer's patent portfolios in order to assess the similarities and broadness of the applied patents. The evaluation informs the buyer if the supplier has very similar patents and if they have patents on multiple sub-technologies. The suppliers are categorised, and their placement in different categories can be visualised in a matrix. As patent information is freely available, the method can be also used to assess other markets as well (Kronemeyer et al., 2022).

As the object of the supplier performance evaluation might differ, there are suggestions to alter the performance evaluation model as well. Supplier performance evaluation can take a cost-based approach, as Roodhooft and Konings (1997) suggest. The supplier is anticipated to make some errors in deliveries, quantities, and quality. These errors have a cost that is estimated by the buying company, and the results are combined with the price difference between the suppliers. The supplier that has the least amount of costs is chosen to be the supplier. Some models might want to find the best supplier based on environmental criteria or other aspects of sustainability. For example, for quantifying environmental impact, energy accounting might be applied (Tian & Sarkis, 2020). Some models are tailored for every selection case. Vokurka, Choobineh and Vadi (1996) differentiate evaluation based on whether the supplier offers a commodity or a quality product.

One example of how to integrate these models into practice is to integrate them into a holistic system and build processes around them. For supplier measurement, Araz and Ozkaharan (2007) suggest a strategic supplier evaluation and management system that is built from three different systems: supplier evaluation, supplier sorting, and supplier management. The proposed system is shown in Figure 8. Araz and Ozkaharan's (2007) system includes a systematic evaluation of suppliers, how the suppliers should be developed, and monitoring of their performance as well. They use multi-criteria sorting methods PROMETHEE and PROMSORT, which belong to methods used for MCDM problems.

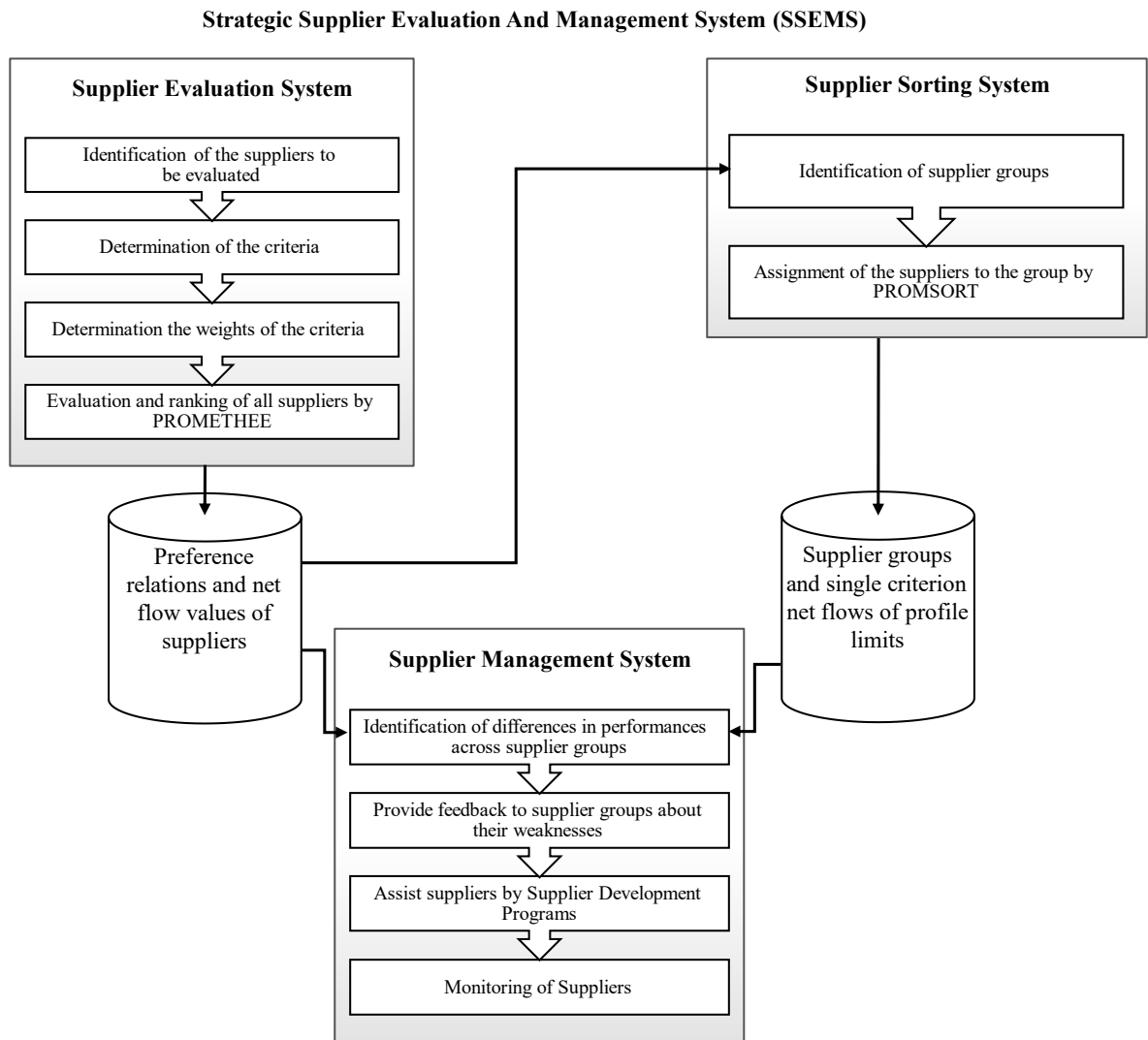


Figure 8. Example of a holistic supplier performance measurement system (Araz & Ozkarahan, 2007)

Supplier evaluation models have some shortcomings. As they are based only on a chosen set of suppliers, they might leave out some alternatives. On the other hand, the models perform better if the company has previously had a relationship with a supplier, giving the company more data to be able to evaluate supplier fit (Sarkis & Talluri, 2002). This suggests that one has to be aware of the bias, that supplier performance evaluation favours existing suppliers.

## 2.5 Supplier performance monitoring

As supplier performance evaluation models' goals are focused on supplier selection, the time focus of evaluation models is usually once. Monitoring can have various goals, but one is to

understand the influence of SRM activities (Park et al., 2010), such as supplier development programs (Araz & Ozkarahan, 2007). Monitoring can be used, for example, to improve supplier quality (Niazmand, Mirzazadeh, & Rezaie, 2014), identify weak or strong suppliers (Bruno et al., 2012), and assess ongoing performance. It can also be used as a continuous risk management tool (Moretto et al., 2019).

Supplier evaluation and monitoring are different in frequency, goal, and data collection, as previously discussed. As the goal of monitoring is different from evaluation, different metrics should be tracked. Using the same criteria for supplier selection and monitoring might lead to unusable results due to differing goals (Lima-Junior & Carpinetti, 2016). However, monitoring and supplier evaluation might have overlapping goals as well, such as identifying underperforming suppliers for supplier development, i.e., supplier selection. In theory and practice, supplier performance evaluation models could be used for monitoring suppliers as well. However, data collection provides a challenge due to the systematic collection of the data. It is easy to use quantitative measurements for monitoring, as they are easily available from ERP systems. As quantitative measurements do not tell everything, qualitative metrics are suggested to be important as well. On the other hand, qualitative measurements need to be regularly collected to be able to use them in supplier performance monitoring. However, even though collecting the data is a challenging task, it can be rewarding, as it has been found that monitoring supplier performance positively influences the supplier's operational performance (Maestrini et al., 2018b). Also, supplier's opportunistic behaviour might be mitigated by monitoring, improving the supplier's operational performance.

Advanced mathematical approaches, similar to those used in supplier evaluation, are used to identify the importance of metrics. However, as the goal is to identify performance over-time, more visual approaches, such as control charts, might be suggested. For quantitative supplier performance monitoring, classical statistical analyses and the use of control chart techniques are suggested (Morgan & Dewhurst, 2007). The benefits of control charts are that they will easily visualize the results from supplier performance measuring and can bring value through root-cause analyses. Also, they can be easily applied to a multitude of suppliers. Applying qualitative metrics in control charts is not necessarily feasible for numerous suppliers due to the need to collect data through surveys (Faraz et al., 2018), eventually

leading to a work-intensive process. A few examples of supplier performance monitoring are shown in Table 5.

Table 5. Existing supplier performance monitoring articles.

Goal & <i>main technique(s)</i>	Main measurements	Reference
Monitoring performance <i>BCC DEA</i>	<ul style="list-style-type: none"> <li>• Total costs,</li> <li>• Number of shipments,</li> <li>• Number of bills arrived without error,</li> <li>• The number of shipments to arrive on time,</li> <li>• Ratings for experience and credence (e.g., access, communication and trust)</li> </ul>	Talluri & Sarkis (2002)
Monitoring performance <i>Statistical process control (SPC) charts</i>	<ul style="list-style-type: none"> <li>• Supplier service (%),</li> <li>• On shelf availability (%)</li> </ul>	Morgan & Dewhurst (2007)
Monitoring quality of supplier <i>Fuzzy SQFE</i>	<ul style="list-style-type: none"> <li>• Mean unit demerit,</li> <li>• Level of none quality,</li> <li>• Coefficient of quality attitude,</li> <li>• Level of quality attitude,</li> <li>• Level of delivered quality</li> </ul>	Niazmand et al. (2014)
Monitoring the supplier relationship stability <i>SPC chart</i>	<ul style="list-style-type: none"> <li>• Legitimacy and compatibility (e.g., trust, power &amp; mutuality),</li> <li>• Social relations (e.g., communication extent),</li> <li>• Economic and shared values (e.g., co-manufacturing),</li> <li>• Learning bonds (e.g., staff exchange, training)</li> </ul>	Faraz et al. (2018)

Supplier monitoring results can be used in strategic decision-making. Monitoring has strategic value, if it is used to understand why a certain supplier is failing during a period and how to prevent it in the future (Talluri & Sarkis, 2002). The results can be used as a baseline, when the supplier is asked to improve their performance (Dey et al., 2015). In the supplier monitoring system, information that brings strategic value to the company should be integrated. For example, one can integrate operational performance information and financial information into a system, to give early warning signs if the supplier is failing (Moretto et al., 2019). Monitoring provides evaluations over time and sometimes, the results are old. The data reliability should be questioned when using old ratings, as suppliers' and buyer's processes evolve, and evaluators change (Hawkins et al., 2020).

Suppliers should have a good understanding of what is anticipated from them, as it impacts how they see the performance measuring (Hawkins et al., 2020). For example, some performance issues could be tackled in contracts and sometimes, it is more beneficial to monitor the performance. Contract specificity and monitoring have been found to improve service performance; however, the results depend on measurement ambiguity (difficulty to define performance metrics in the contracting phase) and task complexity (Ye et al., 2022). Therefore, Ye et al. (2022) suggest varying governance strategies for services that are shown in Figure 9 along with an example of such service.

<b>Task Complexity</b>	<b>High</b>	Primary: contract specificity Secondary: monitoring Example: IT outsourcing	Both contract specificity and monitoring are primary Example: Legal process service
	<b>Low</b>	Neither contract specificity nor monitoring is primary Example: Cleaning services	Primary: monitoring Secondary: contract specificity Example: Office design service
		<b>Low</b>	<b>High</b>
<b>Measurement Ambiguity</b>			

Figure 9. Governance strategies for services (Ye et al., 2022)

Supplier performance monitoring enables performance-based contracting. In services, contract structure might refer to the definition of responsibilities, performance criteria, and incentives. It has been found that the perceived performance from services is influenced by the contract structure and follow-up management (Zou et al., 2019). It is noteworthy, that in their study, incentives refer only to contractual penalties in the case of contract breach. In general, rewarding suppliers for their performance might increase opportunistic behaviour, ultimately decreasing the positive impact of supplier performance monitoring. On the other hand, rewarding suppliers for good performance increased the operational performance of the supplier (Maestrini et al., 2018b). Performance-based contracting requires balancing between motivating and unmotivating goals. If the supplier performance goals are set too high, or the data is unreliable, it might lead to unmotivated suppliers (Hald & Ellegaard, 2011). Rewarding certain performance might lead to suppliers behaving so that they only reach set goals, resulting in counterproductive outcomes from the measuring and incentives.

Industry context also influences the whole process. For example, the pace is different in fast-moving consumer goods when compared to a factory building costing billions of euros. In the retail context, only quantitative supplier performance measuring might be seen as a “stick”, as the focus of the supplier performance monitoring is short (Morgan & Dewhurst, 2007), and the relationship aspects are usually ignored. Performance might vary over time, but the supplier might still be valuable to the buying company, and this might be overlooked in monitoring. Buyers might be impacted by insights, that are provided by monitoring and ultimately show, that the buyer and supplier have a similar understanding of the supplier's performance. For example, in the context of services, follow-up management (e.g., performance reviews, meetings, and revision of contracts) has a positive impact on perceived performance by buyers (Zou et al., 2019).

## 2.6 Visualisation of the results

Supplier performance evaluation and monitoring results can be visualised in different ways. The results are straightforward to visualise if the model produces results in a single metric e.g., performance index. In the case of supplier selection, the results are usually shown as a table that informs the buyer of supplier ranking (Bai et al., 2019). Following a single metric does not necessarily bring value to the decision-makers. More advanced visualisations displays trends and comparisons of suppliers. For supplier evaluation and monitoring, Pal and Kumar (2008) suggest spider graphs, individual performance monitoring sheets, and vendor performance dashboards. These are illustrated in Figure 10.

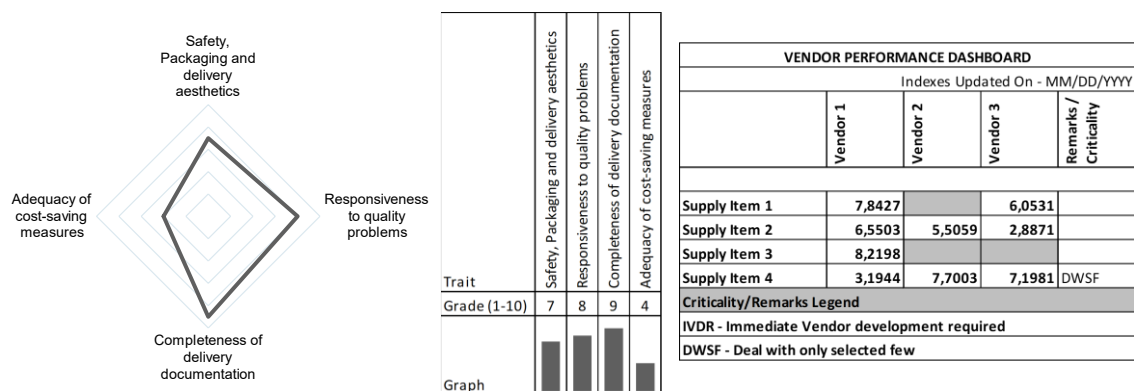


Figure 10. Spider graph, vendor monitoring sheet, and vendor performance dashboard (simplified from Pal & Kumar, 2008)



Pal and Kumar (2008) evaluated suppliers and their performance in supplying certain items. The spider graph and monitoring sheet offer similar but more in-depth information. The overall performance index is used in a dashboard, where suppliers that offer the same item can be compared. Visualisation of results eases the assessment of supplier performance evaluation. Morgan and Dewhurst (2007) suggest that control charts are a good way to visualise the results, as they provide information about the history as well. Also, the buyer will be alerted if the results are unusual. Over time, supplier monitoring provides suppliers with valuable feedback from the buyer. If, for example, a supplier is underperforming during a period or multiple periods, suppliers can perform a root cause analysis of what has happened during those period(s) (Talluri & Sarkis, 2002). An example of a control chart is given in Figure 11. Visualisation of the results might ultimately provide implications about the supplier performance trend. Visualisations supports communication to the supplier.

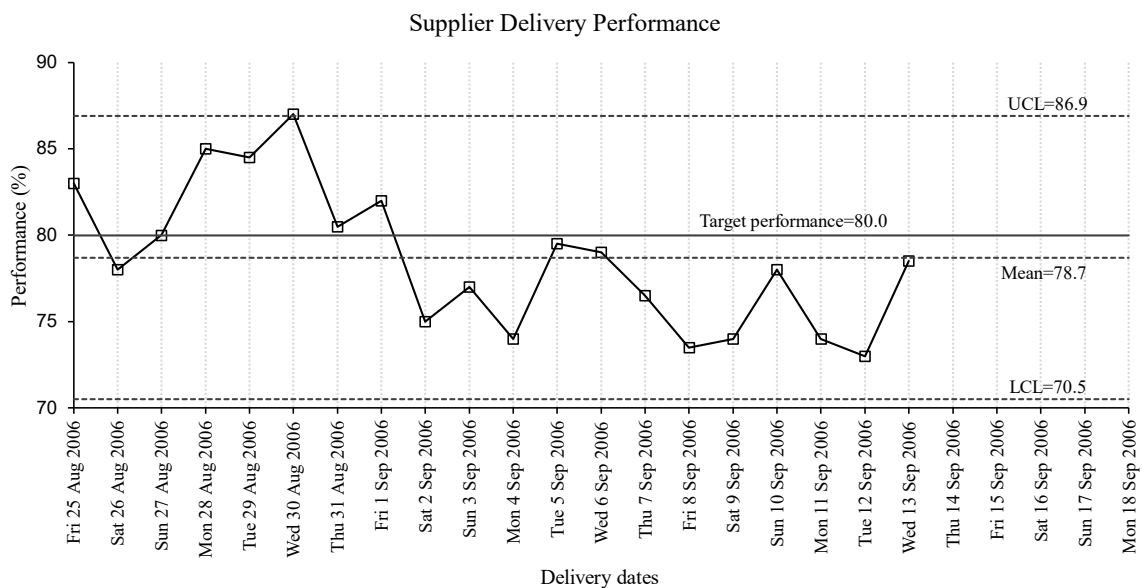


Figure 11. Control chart for delivery performance (adapted from Morgan & Dewhurst, 2007).

Visualisations have their time and place when developing a SPMS, as they transform data into information. As the companies work in dynamic environment, which changes in an increasing pace, the information should be easily available in a correct form regardless of the goal of the SPMS. Demand for clear and concise information is not unjustified, as it has impact on the communication as well.

## 2.7 Communication

Maestrini et al.'s (2018c, p. 304) results suggest that there is

*“a relationship between the evaluation content and suppliers' commitment: the clearer the information, the higher the suppliers' interest.”*

To sum it up, a SPMS provides valuable insights about the supplier base. However, the value of SPMS can be increased by using it as an inter-organizational tool for improving suppliers, and the relationship between buyer and supplier. It is important to design the SPMS to correspond to the needs of the users, as it acts as a primary interface between the buyer and supplier (Luzzini et al., 2014). The use might vary from diagnostic to interactive, but the results should be communicated to the supplier. However, it is not insignificant how the communication takes place, as it influences the supplier's reaction to the communication (Maestrini et al., 2018c). As Maestrini et al. (2021) suggest, it should be communicated that the goal of the SPMS is the common good.

The existence of extranets will give a communication channel for the results and the process can be automated, at least for the quantitative measurements. Just sending the results does not necessarily translate to benefits. It has been identified that the buying firm's benefits from a SPMS come through socialization practices, such as meetings (Cousins et al., 2008). The frequency and benefits associated with these meetings might differ in the different phases of supplier relationships. I.e., at the beginning of the relationship, meetings might be beneficial for trust-building, but when the relationship is maintained, as frequent meetings are not as beneficial. Frequent communication of strategic criteria might not be beneficial, and it might even have a negative effect on the supplier's commitment (Prahinski & Fan, 2007). It is noteworthy, that without building a good buyer-supplier relationship, the benefits from supplier performance monitoring, incentives, competitive pressure, or direct involvement do not lead to performance improvements (Benton Jr, Prahinski & Fan, 2020). Similar results were found by Prahinski and Benton (2004), who suggest that without supplier's commitment, the communication of results will not influence supplier's performance. But on the other hand, communication quality has a positive effect on the supplier's commitment that ultimately leads to better performance (Prahinski & Fan, 2007). Diagnostic use might also decrease trust in the relationship (Maestrini et al., 2021).

The communication practices of the company should be well-defined. This way, the effect of personal judgement is mitigated. Personal judgement can have a positive effect, for example, if the system is failing and therefore, the results are not sent (Hald & Ellegaard, 2011). It can also have a negative effect if someone decides that the results do not have an effect on the supplier and therefore, decides not to communicate the results. It has an impact on the perceived assessment accuracy if the requirements are well defined with the supplier, eventually improving the benefits from the evaluation (Hawkins et al., 2020). Supply performance results might lead to unconscious bias, for example, in cases where the company has a stellar performance history. If a critical component manufacturer has an almost perfect performance history and they make a mistake, their contract is more likely to be terminated than a supplier that has marginally acceptable performance (Chen, Rungtusanatham & Goldstein, 2019).

The reactions and communication from performance measurement might be different. Maestrini et al. (2018c) categorise communication into four categories: no sharing, synthetic sharing, performance sharing & explanation, and joint design. The reaction modes is divided into three categories: indifference, passive interest, and active interest with acceptance/objections. These are depicted in Figure 12. No sharing refers to a situation, where the information is not shared, e.g., due to the ability to simply change the supplier, or it is shared when the results are unacceptable. Synthetic sharing refers to a situation, where a supplier is given a rating, but the reasoning behind the rating might not be disclosed. It might be a source of power for the buyer. Performance sharing with explanation happens, when the buyer delivers a complete set with explanations of how they are compiled. The fourth communication node is joint design, where the buyer and supplier develop the metrics together. The suppliers' reactions to the buyer's communication might be indifferent due to unclear reports, e.g., receiving only numbers as feedback, or the supplier having a lot of power over the buyer. In the case of passive interest, the supplier might assess the results, but pay no mind to them, e.g., due to having a lot of heterogenous suppliers. Third, the supplier might be acting upon the results with acceptance, but also with objections, e.g., if the results are far from their own measuring results.

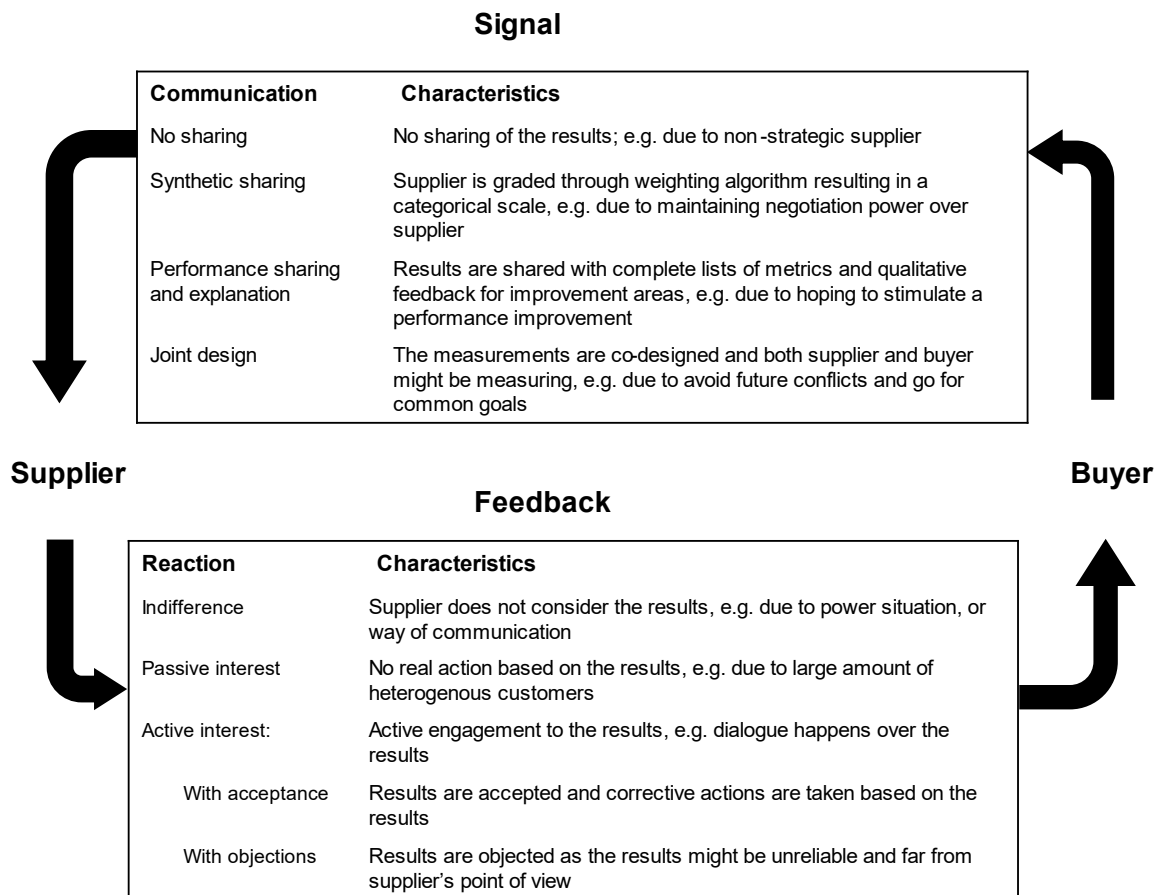


Figure 12. Communication and feedback nodes (Modified from Maestrini et al., 2018c)

The danger of communicating supplier performance results is evident. For example, providing only a score might annoy the supplier (Maestrini et al., 2018c). The result might even implicate that the supplier is overperforming, and the goal of continuous improvement is forgotten. The performance of the supplier might be dependent on the buyer's capabilities to deliver information, e.g., the accuracy of forecasts (Ross et al., 2016). On the other hand, the performance of the system might be dependent on factors that could be solved during the design phase. For example, suppliers might have problems acting upon the results of the supplier performance measuring if the data quality has failed, or if the measuring is unfair (Hald & Ellegaard, 2011). It might have a damaging effect on the relationship. In essence, supplier performance monitoring has the goal to influence the behaviour of the supplier. Based on the monitoring, the supplier might want to have excess inventory to perform well in the measuring, however, it can lead to increased costs for the buyer (Morgan & Dewhurst, 2007), which might not be the goal of the SPMS.

## 2.8. Supplier performance measurement system supporting supplier relationship management

Various aspects that are connected to the supplier performance measurement system were discussed in the previous chapters. The SPMS use supports the performance management of suppliers and further supplier relationship management. Eventually, SPMS functions in an environment that is influenced by the company's strategy and the context of the company. The company's strategic goals might define the purpose and needs first for the supplier relationship management which in turn defines the needs and goals for supplier performance management. Eventually, a tool for performance management is the whole SPMS that supports the SRM and performance management of the supplier. On the other hand, the design of SPMS should consider the context e.g., power balance between the buyer and suppliers, as resources might be wasted on suppliers that are not interested in the buyer or performing according to their suggestions, so tailoring the SPMS for strategic suppliers is evident. In case of more powerful supplier investing in communication and other relational practices is beneficial (Jääskeläinen, 2021). Similarly, the company's existing relationships with suppliers influence the SRM.

Concluding the literature review, Figure 13 illustrates the connections between SRM, SPMS, company's strategy and context. Luzzini et al.'s (2014) model forms the base for the SPMS design. However, apart from their model, the literature suggests that rather than first defining the measurements and then the purpose (supplier preselection, selection and monitoring), one should consider the purpose and need simultaneously as they define the measurements and the measurements are not universal. As for the first phase, following requirements are suggested to be mapped out:

- Purpose/Need
  - Supplier selection, monitoring, development
  - Interactive use, diagnostic use
- Data
- Measurements
- Responsibilities
- Frequency of measuring

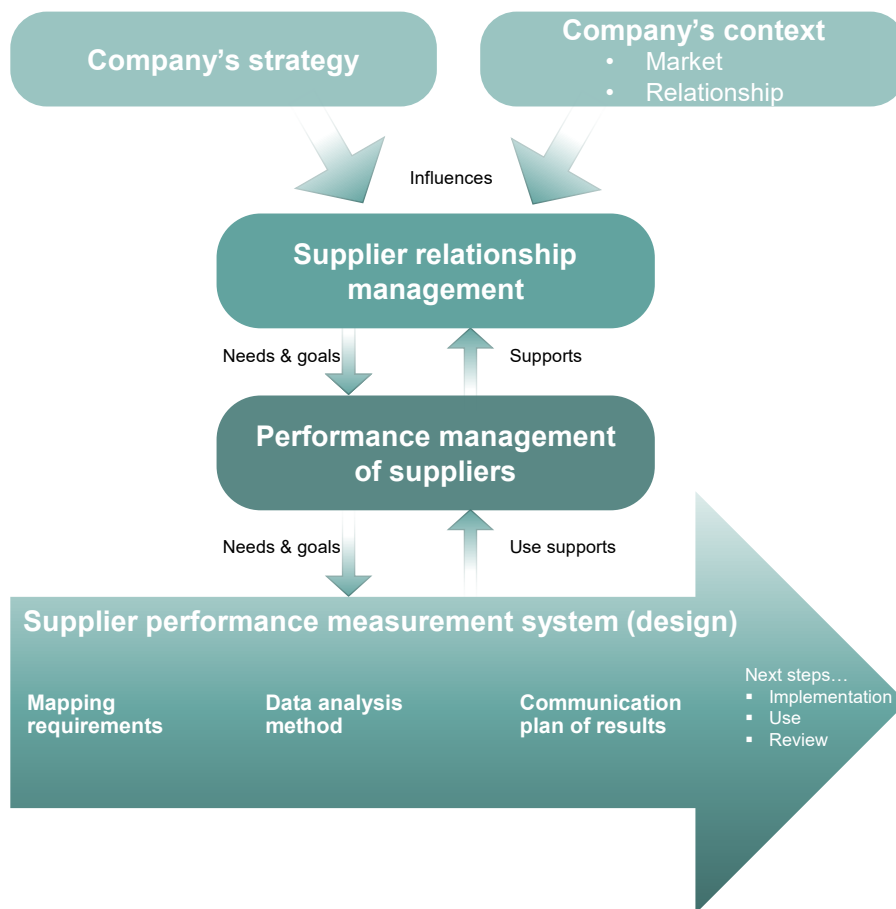


Figure 13. Supplier relationship management connections to supplier performance measurement system design

Moreover, the purpose is not limited to the intended use, as it might be good to consider whether the purpose is diagnostic (measurements are defined before measuring) or interactive (dialogue defines measurements and goals). Interactive use might be in turn heavier to implement as it requires additional discussion between buyer and supplier. The supporting tool (IT infrastructure) should be also flexible enough to cater the varying needs of interactive use. Available data and measurements are good to consider, as new ways to collect data might be even too costly and diminish the benefits from the SPMS. Secondly, it has been suggested that evaluators and other responsibilities should be defined before the implementation. Also, frequency of measuring might be ideal to define to maintain systematic use of SPMS. It has been suggested that the cycle should vary from 1-12 months depending on the goal and available data.

The next phase is more or less about the ways that the data can or should be made into information. The literature suggests various decision-making models. However, the reality

in the companies might be that they are less capable of using the advanced methods. More simple methods for data analysis might be chosen such as scorecards and dashboards. However, if a company has the capabilities to go for advanced models, they might want to choose these models to implement the data analysis for them. However, if there is no high-quality data available, simpler choices may be good to avoid “garbage in – garbage out” -situations and focus on the quality of the data. Tools should be defined in this phase to support the SPMS implementation. As the discussion is about performance management of the suppliers, the results should be communicated, and the communication based on the results should be defined beforehand to avoid unwanted behavior from the persons who communicate with the supplier. After the design, it moves to the implementation and use, eventually reaching need for revision.

### 3. Methodology

This study adopted a qualitative research approach to identify underlying issues regarding the supplier performance measurement system, and more specifically, the design of such a system, taking an explorative stance. First, the research design and case selection are described, combined with the case description. Second, the collection of data is explained. The last sub-chapter explains the chosen data analysis method.

#### 3.1 Research design

This study adopts a single case study approach. It is a method of choice, as the intent is to study the phenomenon in its applicable context and is more intensive in nature, as the purpose is to provide as many insights from the case as possible (Eriksson & Kovalainen, 2016). Single case study provides an excellent opportunity to gain insights from the case company that has a need for practical application of the SPMS. The strong point of a single case study is that it “leads researchers to see new theoretical relationships and question old ones.” (Dyer & Wilkins, 1991, p.614). Even though SPMS might be theoretically new concept but supplier performance monitoring has long traditions in companies. Therefore, a single case study provides an excellent opportunity to study the research problem. Other, methodological choices are summarised in Table 6.

Table 6. Methodological choices

Research philosophy	Interpretivism
Research approach	Abductive
Methodological choice	Qualitative
Research strategy	A single case study with a grounded theory approach
Time horizon	Cross-sectional
Techniques and procedures	Literature review, Semi-structured interviews with purposeful sampling, Data analysis with Gioia methodology

The study aims at transferability, which refers to “a single observation can represent a principle that applies to many different contexts” (Gioia, 2021, p. 21). Companies might have a specific way they want to approach the development of SPMS design. This study is focused more on the ex-ante, i.e., before developing such a system, providing other companies with some information on what they might want to consider while developing a SPMS.



Business research is founded in multiple disciplines and therefore, it is a challenge to define exact practices for research methods. Due to the foundations of business research, multiple methods and philosophies coexist (Saunders, Lewis & Thornhill, 2016, p. 126). The research philosophies vary across their ontological (nature of reality or being), epistemological (what constitutes acceptable knowledge), and axiological (role of values) approaches. The chosen data analysis methodology has some implications about which research philosophy one might follow. If the research is purely inductive, e.g., the researcher tries not to apply any previous knowledge and tries to be free from any philosophical influence (Berthelsen, Lindhardt & Fredriksen, 2017). However, as the research applies an abductive approach, and puts a strong emphasis on the data and understanding of interviewees, this research follows the philosophy of interpretivism. The influence of pragmatism cannot be completely ignored, as the topic and research questions are pragmatic in their nature.

The time horizon is cross-sectional, as the data is collected during one period of time and no follow up is provided. Techniques and procedures followed during this research are literature review, semi-structured interviews, and a single case study with a grounded theory approach using Gioia methodology. The semi-structured interviews and data analysis are explained more in-depth in the following sub-chapters. In interpretive grounded theory, the literature review acts as a guide for the researcher. Due to a lack of previous knowledge, the literature search started as systematic, while applying the words: supplier\*, performance and measur\*. Other synonyms were applied as well. However, after writing the initial version of the literature review, it was clear that other terms were connected as well, e.g., strategic supplier relationship management and purchasing strategy in general. Afterward, the systematic way of conducting a literature review was discarded. The initial literature review provided a foundation for semi-structured interviews and their questions. The literature review was left incomplete before data collection, as suggested by Corbin and Strauss (2008), to be able to improve the literature review after the researcher knows what emerges from the data. After the interviews and data analysis, the literature review was finalised.

### **3.1.1 Case selection and description**

It has been argued in the literature that there is no single best way to design an SPMS (Luzzini et al., 2014). As an SPMS is context specific and a deep understanding of context

is important, a single case study has been chosen. The case company is a good choice due to the size and varying needs due to its large number of suppliers. As there is a large variety of suppliers, there is large variety of needs that is studied. Therefore, it may reflect the usage of SPMS in other companies as well, providing a good foundation for transferability. Also, being an established company they have experience in supplier management which can be supported by SPMS. The richness of the case is captured by involving a variety of specialists, that have experience in working with a variety of suppliers (e.g., suppliers who offer services, infrastructure, and indirect or direct products).

The case company is a manufacturing company with multiple billion euros in revenue. The company uses multiple systems that collect supplier-specific data, that could be utilized to follow supplier performance. They already have some key performance indicators (KPIs) that are used to monitor suppliers; however, they are hard to reach, and the company's measurement capabilities are not in optimal use. Also, they have not yet identified, which variables could be useful for the case company to follow and could bring strategic value to the decision-making. In the company, they use a supplier classification system that provides some foundation for measuring. In a broader sense, the current supplier performance measuring is not systematic and does not cater to the varying needs of the case company.

This study and the interviews have a secondary goal of producing buying companies with knowledge about how they should build a holistic SPMS. It ultimately benefits the outcome of this thesis as they have invested resources for this process and provides committed interviewees leading to accurate data sources. This could influence the results; however, the company's interest is not to influence the employees and their answers, but to see the ideas of employees and what has been identified in the theory. The case company left the researcher with an open assignment and only provided help with, e.g., identifying key employees for interviews. The researcher was employed in the case company during the study but did not have ties to the case company prior to the assignment. An alternative report was produced to fulfil the case company's needs and the thesis was not influenced by the case company.

## 3.2 Data collection

### *Interview questions:*

In Gioia methodology, the focus is not to validate existing theories, but to build new or elaborate existing theories around a phenomenon. Therefore, in this study, the interview questions were left open, leaving the interpretation of the questions to the interviewees to see how they understand supplier performance measuring. For example, the purpose of supplier performance assessment was seen differently by the interviewees. Interviewees were not provided with existing theories about supplier performance measuring, and it is highlighted that the interview is about their opinion and point of view. It was pointed out that the point of the interview is not to measure their expertise on the subject, or how much they know specific terminology, but actually identify their views and opinions on the matter.

Questions and themes were identified based on the initial literature review. Appendix 1 shows the interview questions and reasoning why questions were included. As there were participants, who were not directly involved with suppliers, the same questions were still used to keep the same structure. However, the interviewees were informed that they can rephrase, or say if the question is not relevant to them. Usually, interviewees provided interesting insights about the question's topic, even though they did not directly answer the question. In total, only a few questions were skipped during all the interviews. Some deviations in interview questions were used for experts in sustainability and safety.

### *Sampling:*

Purposeful sampling was used to identify interviewee candidates. The goal for the sampling was to have interviewees, who were i) knowledgeable on the topic and ii) varied in expertise. The relevant interviewees were identified with the help of a category manager, and other employees also suggested relevant interviewee candidates as well. Expertise in this context mainly refers to inviting interviewees that had different main procurement areas, such as project-related, services, or direct materials. Interviewees were category managers, purchasing specialist, sustainability, and quality specialists. Seniority level varied from specialist to vice presidents.

### *Interviews:*

After the initial round of interviews (21) and analysis, it was noticed that there are still issues that are highly related to this topic, but knowledgeable persons have not been included in the sample. Three additional persons were invited to the interview. In total, 24 interviews were conducted. The semi-structured interviews were used as a main source for data collection. In some cases, the interviewees were later contacted through Teams to clarify their responses or provide additional information. Semi-structured interviews were chosen, as they were recommended as the main data collection method for the Gioia methodology.

The data was collected with 24 interviews between May and July 2023. The interviews were conducted in Finnish (22) and English (2). In total, there were 1 210 minutes of interview recordings. They were transcribed with the help of Word Online and Teams. The transcriptions were checked manually and corrected. It resulted in 735 pages of data. The tenure of interviewees (in the case company or subsidiary, the total is calculated) varied from 5 months to 32 years. The average tenure in the sample was 9 years and 9 months. To improve the anonymity of the interviewees, their position in the company is not disclosed. Table 7 contains interviewees speciality areas and length of the recording.

Table 7. Interviewees and reference in text

Interviewee and reference in text	Area	Length
Interviewee 1	Sourcing	53 min
Interviewee 2	Sourcing	43 min
Interviewee 3	Sourcing	42 min
Interviewee 4	Sourcing	38 min
Interviewee 5	Sourcing	50 min
Interviewee 6	Sourcing	38 min
Interviewee 7	Sourcing	42 min
Interviewee 8	Sourcing	42 min
Interviewee 9	Sourcing	1 h 8 min
Interviewee 10	Sourcing	26 min
Interviewee 11	Sourcing	42 min
Interviewee 12	Maintenance	1 h 36 min
Interviewee 13	Logistics	48 min
Interviewee 14	Sourcing	56 min
Interviewee 15	Logistics	52 min
Interviewee 16	Logistics	34 min

Interviewee 17	Logistics	50 min
Interviewee 18	Sourcing	42 min
Interviewee 19	Logistics	50 min
Interviewee 20	Logistics	1 h 2 min
Interviewee 21	Safety	54 min
Interviewee 22	Sourcing	42 min
Interviewee 23	Sourcing	1 h 27 min
Interviewee 24	R&D	1 h 3 min

---

The interview transcriptions were not translated, but the quotes presented in chapter five and Appendix 2 have been translated from Finnish to English. Although effort has been made to convey the meaning and context of the original quotes as accurately as possible, some minor differences may be present in the translated quotes, as some phrases would be difficult to translate directly. The sentence structure also often had to be restructured to fit the target language, while finding exact matches in meaning proved to be challenging for some individual words. Additionally, to make them more easily understandable, the quotes were edited by removing some unnecessary repetition and colloquial expressions from them. However, the context and meaning of the quotations should still be intact.

### 3.3 Data analysis

After conducting the interviews, initial outcomes were noted down by the researcher for later use to support the data analyses. It is by no means an interview diary, but it helped with upcoming interviews and also during the naming of the codes and identifying relevant observations to be coded. Also, during transcribing, another set of notes was usually generated. These notes helped the data analysis process, as they provided some initial thoughts from the interviews.

As Saldaña pointed out (2013, p. 8) "... coding is a cyclical act", this research also followed the same mindset. Rather than trying to aim for perfect results at once, the coding was given time to develop through recoding, reading, and renaming. Giving time for the coding process helps with the quality of the results, as the researcher is a novice in coding. As suggested by Saldaña (2013, p. 17), novice coders should not be too strict about what they code, as they might ignore what might be important in the data. Through experience, the researcher

understands what is important in the data, and what is irrelevant or trivial. To help with coding NVivo was used.

Due to using the Gioia methodology in this thesis, the research varies across inductive and abductive logic. As described by Eriksson and Kovalainen (2016, p. 24), abductive approach refers to “the process of moving from the everyday descriptions and meanings given by people to categories and concepts that create the basis of an understanding, or an explanation of the phenomenon described.”. The process of how the approaches change between induction and abduction is depicted in Figure 14.

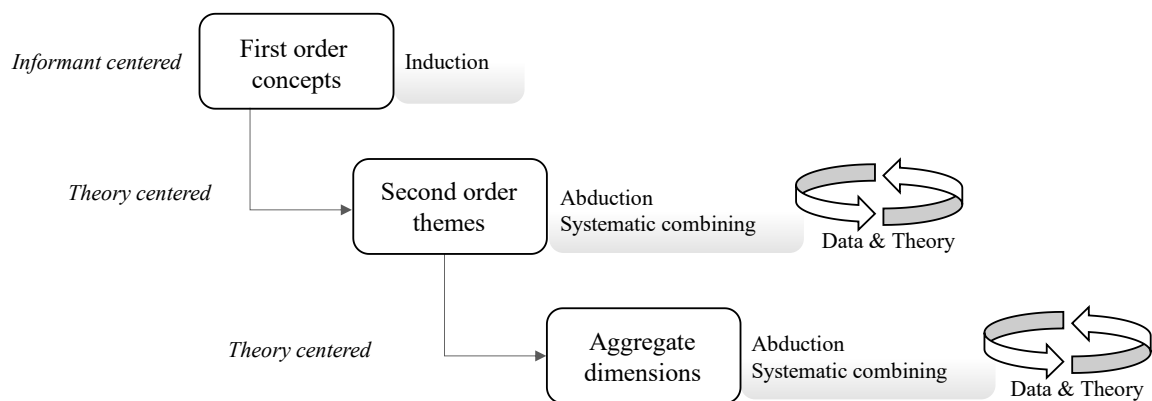


Figure 14. Moving from induction to abduction in Gioia methodology (Magnani & Gioia, 2023)

The first concept of coding aims to code the phenomena as the interviewee (knowledgeable agent) sees it and uses similar wording (Gioia et al., 2013). In the first round of coding, the interesting phrases were coded, or in vivo coded, based on how well they could be worded as close as possible in interviewees' terms. Over time, the in vivo codes were coded into new or existing codes. The first round of coding resulted in over 700 codes, and further filtering was required. Due to the richness and extensive amount of data, the coding and recoding were conducted multiple times, increasing the quality of the codes and ensuring relevance to the topic. This process helped to familiarize the researcher with the data and underlying issues better. Final codes with quote examples are shown in Appendix 2.

Second-order themes were emerging from the first-order concepts, trying to find theoretical implications in the codes that might be ignored in the current theory. By combining the informants' and researcher's voices, the results provide a more “qualitatively rigorous”

demonstration of data-to-theory connections and gives some confidence that any creative insights are rooted in the informants' experience (Gioia, 2021, p. 24). Aggregate dimensions were identified from these second-order themes, the emergent data structures is shown in Appendices 3 and 4. The second order themes and aggregate dimensions are shown in Figure 15.

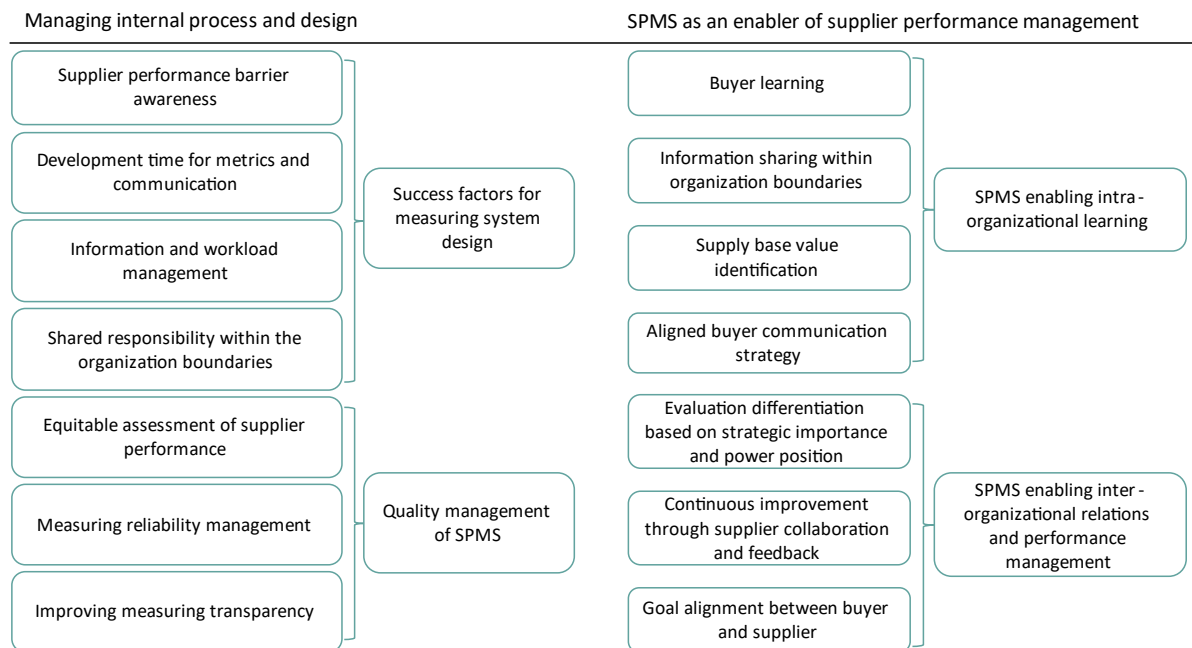


Figure 15. Second order themes and aggregate dimensions

In total, four aggregate dimensions were formed. There were division of aggregate dimensions that had a clear focus on practical, more operational point of view and the second had focus more on the intended use. Managing internal process and design is focused on factors that influence the operational use of SPMS. And on the other hand, SPMS as an enabler of supplier performance management is focused on the possible uses that were identified in the interviews.

### 3.4 Reliability and validity

Validity is a challenge in qualitative studies, as it is dependent on the data and the researcher themselves. Also, the validity is increased by providing quotes from the data, as it provides evidence of the chosen codes. The study was conducted under the supervision of more

experienced researchers. They improved the validity of the study by validating the codes and suggesting alternative wordings. The validity of the results is impacted by choosing the single case study. Single case study provides an excellent opportunity to gain in-depth information, but it limits the generalizability of the results (Kähkönen, 2014). On the other hand, the aim of the study is transferability, which means that the aim is to generate insights that can be applied in a variety of contexts, i.e., providing knowledge that could be considered, when in a similar situation. This has been achieved by interviewing experts that are responsible for a variety of purchases and other functions, which in turn improves the validity of the study as well.

There were various steps, where the study aimed for improved reliability. Yin (2009, 45) describes reliability as description of the research process, so that it would be possible to follow same procedures so that the results would not change. Gioia methodology was a suitable choice for the research problem as the phenomena is context dependent and the topic is subjective. Semi-structured interviews provided an excellent opportunity to capture the interviewees' opinions and point of views around the studied topic. This study aimed to increase the reliability by carefully describing the used methods and different phases in the study. The number of interviewees were satisfactory, as over the last interviews, the saturation was reached, and the number of new insights were rare. The interview questions, and the example of codes and quotes are presented in the appendices as well.



## 4. Findings

Through the analysis, a data structure is formed. For the clarity of the findings section, the data structure is divided into two: *Managing internal process and design* and *SPMS as an enabler of supplier performance management*. Through analysis, a total of four dimensions were formed. This section elaborates on emerging themes and dimensions.

### 4.1 Managing internal process and design

The first aggregate dimension is success factors for measuring system design, which is a collection of factors that interviewees suggested to be important to supplier performance monitoring throughout the interviews. The second dimension is about the quality management of SPMS, where the themes are related to how the quality of SPMS can be either increased or decreased by measuring practices. A summary of the themes and short descriptions from this data structure are shown in Table 8.

Table 8. Managing internal process and design themes

Aggregate dimension & theme	Description
<i>Success factors for measuring system design</i>	
Supplier performance barrier awareness	It is not straightforward that measuring describes the supplier's actual performance. There might be various factors that influence the results, e.g., forecasting, ordering habits, too tight schedules or different measuring practices.
Development time for metrics and communication	Supplier performance measuring and communication should be given time to develop based on identified good practices and results.
Information and workload management	If workload and/or amount of information is too extensive it might lead to non-use.
Shared responsibility within organization boundaries	A single person should not be responsible for the whole process but it should be a shared responsibility of within the organization.
<i>Quality management of SPMS</i>	
Equitable assessment of supplier performance	Assessment should aim to avoid unfairness that might rise from measuring practices and feedback.
Measuring reliability management	SPMS should aim to produce reliable information that is usable to manage the performance of the supplier.
Improving measuring transparency	Measuring should aim to increase the transparency between buyer and supplier.

### 4.1.1 Success factors for measuring system design

Throughout the discussions with interviewees, some concerns were raised about issues that influence the success of such system. The aggregate dimension is compiled from themes that are related to the buyer's practices that should be managed in the design phase. Managing the system so, that the amount of information is not too extensive, and the process is not increasing the workload too much, has been seen as important.

*Supplier performance barrier awareness:*

It was highlighted in the interviews, that suppliers' performance might not be only about how they perform, but other factors might also influence their performance. This might be due to the buying company's information-sharing challenges or measuring practices. The key is to be aware of such issues and hopefully, eventually fix these. Depending on the reasons, it is not an easy challenge to tackle. One of the challenges related to the case company's industry is forecasting, which in turn influences how much suppliers reserve capacity:

*"When the volume changes come in a short cycle, the time span for arranging more capacity with the current suppliers or obtaining additional capacity from new suppliers becomes a bit too short. So perhaps partly through our actions and partly through the nature of our operations, we contribute to these challenges." [Interviewee 20]*

Without understanding what influences their performance measurements, the feedback given might be unfair. Too tight schedules might also be a barrier to supplier's performance. This might be an information sharing issue, or due to planning:

*"How we plan that. That is, how we plan, for instance, our own production and deliveries and such. If our schedule planning is too strict, the supplier might not be able to perform accordingly" [Interviewee 16]*

The case company has long and good relationships with suppliers, which were seen as an important factor in understanding the supplier's performance. For example, the varying performance is not necessarily about the supplier itself, but practices for measuring might be different in different places. These factors have to be understood to avoid unnecessary conflicts with suppliers when, for example, comparing the suppliers:

*"A car or truck is registered as having arrived when it enters through the gate to the factory area, but what greatly affects the comparability of arrival data is the presence of a lunch restaurant in the port area, which means that very often, the driver might take a break there in the port area and have lunch. And then, of course, the departure is registered when they drive out of the factory area. So, if we, for instance, compare that to the port of [a city], where when a vehicle arrives, it unloads and departs. And if, without knowing these facts, we would start comparing these two locations like: '[previously mentioned city's port] has totally miserable vehicle cycle times, so please do something'. [angrily] ... But on the other hand, as we know these things, we will not go on to make the mistake of directly comparing suppliers with each other using some of these particular indicators [that are not comparable]." [Interviewee 17]*

#### *Development time for metrics and communication:*

Having some level of guidance about communication might be beneficial to ensure that the communication actually happens, and that the results are not just sent to the supplier without an open discussion. It might be more beneficial to give time to develop measuring and communication practices through experience, rather than aiming to be perfect at once:

*"I think that the [communication] practices form over time based on conversations. When it is identified what produces added value and what does not." [Interviewee 21]*

The theme is also related to the management of the amount of information, as having an extensive amount of data might not be beneficial. Also, in the beginning, it might be more beneficial to focus on making it work, rather than making it complex. When companies want to have an exhaustive system at place, it might lead to excessive amount of information:

*"That is, people in different areas have this and that many things that they would like to know and measure. 'It would be nice to know this and that. And these and those.' ... Then it forms, what we also have and can be seen, such a huge amount and a huge jungle of different information, different metrics and different reports. And perhaps with that, what is essential and where we really want to put our effort into becomes lost to us. That less is more. ... if it helps us to move forward, we should prefer simpler, even if they are as simple as possible, they often help more than very advanced, theorized ones. Even in basic matters, there is usually quite a lot to do." [Interviewee 18]*

#### *Information and workload management:*

In the SPMS design, it is noteworthy that workload might be inhibiting the use, and decreasing the potential benefits from using the SPMS. There is the data that is directly available from the system and does not need additional work from the users, but if there is an additional, manual evaluation done outside of the system, it increases the workload. Some of the interviewees had experience with evaluations that were collected by using Excel sheets. If

the process is automated to the extent that is possible, it will decrease the amount of manual work. The frequency of evaluation influences the workload as well. Due to the infrequently changing environment, the yearly evaluation frequency was seen as enough. Also, the SPMS role was not seen as an alerting system:

*"There is nothing that major happening on the suppliers' side. If something does happen there, it is either bankruptcies or something else. Hopefully, we will hear it through someone else or through some other channel [than the performance monitoring]." [Interviewee 9]*

Another factor that is inhibiting the use of such a system is the amount of information. In the current business environment, the data available has increased tremendously. It was suggested by multiple interviewees that the amount of information should be managed, and maybe a solution for it is to have one summary page of the main results, and then the possibility to look behind the measurements if needed.

*Shared responsibility within the organization boundaries:*

Supplier performance measuring, if it is quantitative or qualitative, is constructed by the data that is either collected manually or provided automatically by information systems. However, there is always someone responsible for a piece of information. It was recognized that some incidents go without reporting. Also, the company's employees influence the data quality at various points of the organization. In the eyes of interviewees, they pointed out that someone has to be in charge of completing the process, and in the case company, it would be the person responsible for the supplier. Hence, the evaluation should be in the hands of the supporting organization to avoid multiple evaluations across business units (BU). They named key evaluators from the case company, including the person who initiated the need for purchase, and employees who actually work closely with the supplier. However, it cannot be the responsibility of a single person that the data quality is right, and everything is in the system as it should. It should be designed so, that the results are an effort of multiple persons in the company:

*"Then there's the operational procurement, but also the technical organization in that they record the reclamations, etc. The fact is that everyone does their own part, and handles their plots and entries into the systems ... But if that [only one] person was responsible for ensuring that all the reports are functioning, that the numbers are correct, that the data is imported correctly, etc., then no, it is too large of a responsibility to bear. It is not really possible." [Interviewee 3]*

### 4.1.2 Quality management of SPMS

The second dimension is about practices mainly related to the measuring practices of the company, that can either increase or decrease the quality of SPMS. Sometimes, the quality is influenced by the evaluator's bias of the buyer and also, evaluation is more or less a subjective opinion. Feedback should aim to be equitable. Improving measuring transparency and reliability can be used to improve the discussions with the suppliers, ultimately benefiting the perception of quality of the SPMS.

#### *Equitable assessment of supplier performance:*

Practices might lead to unfairness towards the supplier, which can be tackled by managing the quality of measurements and feedback. To avoid bias in numerical data, one should consider making the measurements as neutral as possible. For example, if reclamations are used to indicate the quality of the supplier and the number is followed, it should be divided by, for example, number of deliveries to make measuring more neutral. If only the number of reclamations is followed, the buyer might be negative towards a supplier whose volume is huge and quality is top notch, leading to unfair treatment of the supplier. Also, one of the interviewees pointed out, that the results should be presented in a sensitive manner, as the buyer's data quality might be failing, not the supplier's performance:

*"... I would at least gently ask 'how do your on-time delivery numbers to us look like?' or 'what is your supply security for [the case company]?' ..." [Interviewee 10]*

If the supplier's performance is evaluated with a subjective opinion, it might be problematic due to the evaluator's bias or other factors that influence the outcome. For example, it was noted that recent problems with suppliers might be highlighted in the performance evaluations. Also, filtering does not mean that the problematic issues should not be brought up, but the way that they are presented should be constructive, not negative just for the sake of being negative:

*"Perhaps here, we must also begin by going through the positives, so that we should not start with 'this is what is amiss with you'. Let's leave the bad part out and state 'here is where you should improve'." [Interviewee 5]*

#### *Measuring reliability management:*

The reliability of the measuring should be examined from time to time. Interviewees brought up issues related to the evaluator's distance from the supplier and to the frequency of measuring. The evaluator might be unreliable if they do not work closely with the supplier, as they just evaluate what they have heard from elsewhere. The outcome from the performance measuring might be unreliable, if the assessment of the results is too infrequent:

*"Once a year is, perhaps to some extent, a bad practice, because we are already a bit reactive as we are talking about a period that is already in the past, and we might not necessarily have the regularity to influence it." [Interviewee 21]*

Reliability is increased with systematic and extensive data collection, which would also benefit the objectivity of measuring. To improve the reliability of measuring, it was suggested that the buying company should ensure that incidents are reported properly and have bilateral evaluations with the suppliers to bring depth into the measuring. Properly reported incidents, such as reclamations, could show that rather than being a single incident, there is a problem that should be solved:

*"[the employee thinks:] 'I will just call that supplier now, leaving no trace of it anywhere', even though our supplier manager does not know at all that the supplier performed poorly or made a mistake or something else [as it is not marked anywhere]. And the next time they go on to negotiate with the supplier, everything will be fine and there is no problem, even if there really was one. And then of course, it might be that you only take care of those individual cases, even though the same supplier has received complaints from several parties." [Interviewee 14]*

#### *Improving measuring transparency:*

Properly reported reclamations influence the transparency of the measuring in the first place. It provides proof to the discussions, as they can be used to show where and how the performance is insufficient and also, where the supplier has succeeded. The performance feedback changes from purely subjective opinion to more objective, if it can provide written proof of the supplier's performance, e.g., by showing that some incidents happen continuously, rather than being a single incident that might make the observation subjective. To avoid uncertainty

related to supplier performance measuring, it should be clear how suppliers are assessed, and how the results are used:

*"Thinking from the supplier's point of view, if I just got some results and did not have the ability to go through them or I was not told what they were based on, it could be confusing ... As in, they would get evaluated, but in reality, only the evaluator would know what they are based on and why this is being done or where the numbers come from?" [Interviewee 2]*

Some interviewees took another stance towards the measuring of the supplier. There were some existing measuring practices in the case company, where the supplier provided the data about their performance. Due to this, suppliers have to assess their own performance when they provide the data, and have a moment of reflection on it. On the other hand, one of the interviewees suggested two-sided measuring to show, that the buying company demands the same things from itself that it demands from the suppliers:

*"So that we are able to say that we do it this way and we want you to do it like that as well." [Interviewee 3]*

#### **4.1.3 Dimensions with operational focus**

The previously described aggregate dimensions have a clear focus on the operational side of SPMS. These themes describe concrete actions that companies can do to influence a successful adoption of a SPMS. Further, the success factors are more so describing factors that should be considered during the design phase, while quality management is more or less about factors to consider during the implementation of the system (Figure 16).

Success factors for design are compiled from design choices that should be influenced before implementing such a system. It does not mean that these factors should not be considered while updating an existing SPMS. If later on it is noticed, for example, that the SPMS has problems in use, the reason might be found from these factors. It is noteworthy, that interviewees sometimes reflected on the use of SPMS based on their experience, as 1) some of them had developed such systems for use in the case company and/or 2) their previous workplace had one.

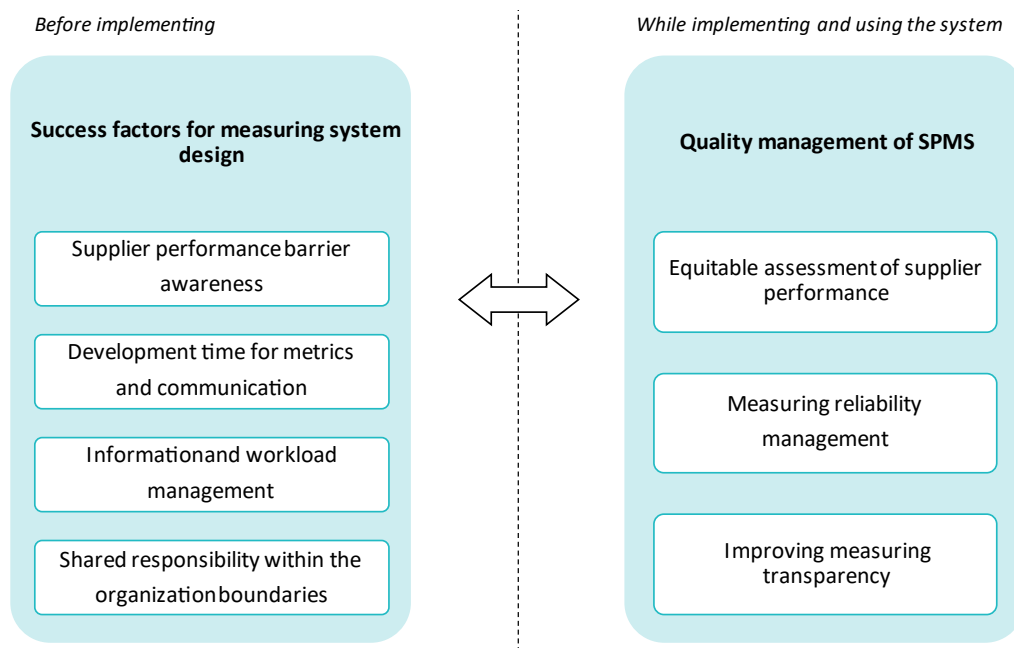


Figure 16. Managing internal process and design

## 4.2 SPMS as an enabler of supplier performance management

Through the analysis of data, it was found that SPMS can serve multiple roles in an organization. First, it can enable employees to learn from the supplier performance results (in the case that the performance measurement system would be widely available inside the organization), further enabling the buying company to manage and share information more effortlessly. If the usage and communication are defined, it can make the company's communication appear more aligned to suppliers. Secondly, it can help the buying company to establish more open inter-organizational relations with the suppliers. As goals can be agreed upon with suppliers, it can help with goal alignment between the supplier and buyer, further enabling the supplier performance management. In Table 9, the themes and short descriptions are summarized.



Table 9. Supplier performance measurement system as an enabler of supplier performance management themes

Aggregate dimension & theme	Description
<i>SPMS enabling intra-organizational learning</i>	
Buyer learning	Results might show areas that buying company should improve e.g. processes or habits (e.g. better information sharing).
Information sharing within organization boundaries	SPMS enables information sharing within the organization boundaries as the results are available to multiple persons rather than being tacit knowledge in organization.
Supply base value identification	Measuring can be used to indicate where best capabilities are in the supply base.
Aligned buyer communication strategy	Defining communication can align the buying company's communication practices, making sure that suppliers receive feedback of their results.
<i>SPMS enabling inter-organizational relations and performance management</i>	
Evaluation differentiation based on strategic importance and power position	Not all suppliers are made equal - suppliers should be chosen to be evaluated based on strategic importance and it should consider power situation as well.
Continuous improvement through supplier collaboration and feedback	Improvements in supplier's performance is reached through collaboration practices (e.g. meetings) and giving feedback.
Goal alignment between buyer and supplier	If the measurements are done in an agreement, they can help the buying company and supplier to aim for common goal.

#### 4.2.1 SPMS enabling intra-organizational learning

This dimension illustrates, what kind of an impact SPMS can have inside an organization. SPMS's only use is not to follow the supplier's performance. It allows the buying company to learn from working with the supplier. It is not only restricted to people who work closely with the supplier, as it can help with information sharing for other parts of the organization as well. As a company's size get larger, it can help the buying company to identify potential value opportunities from the existing supplier base. On the other hand, it helps the buyer to align the organization.

##### *Buyer learning:*

The buyer learning theme is about what the organization can learn from working with the suppliers. Discussions can also show areas of improvement for buyers, e.g., when the buyer is inhibiting the performance of the supplier. Evaluation indicates what could be done differently next time when working with a supplier, or if some things should be tackled in the

contracts. All in all, it is more or less about how the buyer can change their working practices to ensure the best outcome for the buyer and supplier. Person responsible for safety noted about managing the supplier's working practices through instructions:

*"However, as we may not be full professionals in the supplier's field, then of course we want to trust the supplier and cooperate to build that practice in such a way that it also works in practice, so that we do not just make instructions that exist on paper but are not actually followed" [Interviewee 21]*

In the interview, there was a question about comparing suppliers to others. The interviewees noted that comparing could be used to identify the best- and worst-performing suppliers. If the suppliers are similar, and also in other situations, it should be identified why they are better or worse than others. Again, it is not necessarily only about developing the supplier, but might be about how the buying company could improve its own processes.

*"Is there anything that we could perhaps do better. For example, is there something to improve at the level of communication or otherwise, so that we could reach the same results as these other suppliers that are performing better" [Interviewee 15]*

*Information sharing within organization boundaries:*

As the case organization is a large company, information sharing is something that should not be considered self-evident. Therefore, a lot of interviewees saw performance evaluation as an opportunity to hear performance feedback from internal customers. Satisfaction with the supplier might be different in different places. One of the interviewees commented that some business units might be really happy with the supplier, while others might be struggling to cope with the same supplier. Interviewee from logistics commented the following about the importance of the business unit's employees' opinions on the supplier's performance:

*"... those business units are like the number one priority, where feedback needs to be asked from, so that in a way, you have to get them to commit and then collect the feedback that tells you which suppliers work, and which do not." [Interviewee 13]*

As in any company, reclamations or other performance deviations will happen from time to time. As was previously discussed, these can be used as cases to learn from. However, the companies can also take a proactive stance in information sharing. As these deviations happen, the system could warn other factories as well, so that they can look if they have the

same vulnerability. As information is shared and available later on, it can help integrate new employees into the organization.

*"If you think, for example, that the responsibilities suddenly changed and someone else came over to do my work and I wasn't around to share my own experiences, that new person would then, in a way, have to learn the same things all over again." [Interviewee 4]*

#### *Supply base value identification:*

Depending on what is measured in the company, the information could be used to identify value potential from the supply base. As indicated in the interview, it should be measured somehow, what is the innovativeness of the suppliers, and whether they can bring value to the company. Some suppliers have a long history with the company, and they might have rigid working habits and are not necessarily willing to develop their processes. Investing in development programs for rigid suppliers might be a waste of money in the first place. Therefore, it is important to identify suppliers, who have the potential to develop further.

*"That a supplier would be able to bring their own views forward, telling us that they have this kind of a thing, and they would be ready to develop these types of things. Those are the things that, in my opinion, should also be measured to determine how capable a supplier is. Not only in that, what they normally do from Monday to Friday, but also in what additional value they can bring to our process" [Interviewee 10]*

Suppliers work as their own entity, which means that they have their own plans and struggles, that the buyer might be completely ignorant about. Interviewee from procurement theorized, that the buyer could have indications about the supplier's future performance through measuring:

*"... they are also the two first indications that will let you know if a supplier starts struggle, if the quality or reliability of the delivery start to suffer ..." [Interviewee 3]*

#### *Aligned buyer communication strategy:*

In large organizations, it is beneficial to have a communication plan, as it improves how the organization looks to outsiders as well. Communication practices could be linked to governance models to ensure that the treatment is more equal. It will also improve learning, as everybody has to learn only a single way of measuring. On the other hand, having a communication plan would ensure that the communication takes place as has been agreed upon.

*"And it develops it towards systematization, so that everyone does not have to reinvent the wheel and do things in their own way and carry those operating methods with them [from previous workplaces] all the time. And this way, it is done in a standardized way and more time is left for the essentials, when learning the common work practices." [Interviewee 3]*

#### **4.2.2 SPMS enabling inter-organizational relations and performance management**

The dimension is focused on the role that SPMS plays between the buyer and supplier. It was identified that there has to be some level of differentiation in the measuring process, and the differences mainly stemmed from power position and strategic importance (in the case company, they have implemented a categorization model for the supplier base). Similarly to the previous dimension, a theme of information sharing was emerging, and the feedback was seen as important. However, the focus was more on how measuring assists with bilateral information sharing, and how suppliers can fulfil the needs of the buyer, i.e., how performance can be managed. It was identified that there should be some kind of alignment of performance goals between the supplier and buyer to bring benefits from the SPMS.

*Evaluation differentiation based on strategic importance and power position:*

The metrics, which are easily available from the system, should be used for the whole supplier base, and no differentiation is needed. However, they noted that not all of the suppliers are made the same, and there should be evaluation differences between important and less important suppliers. One way to understand the criticality of suppliers is to divide based on if the supplier influences the processes or production:

*"In my opinion, it is worth it to simply separate the performance evaluation into those suppliers that have a direct impact on our process and production, and those that do not." [Interviewee 9]*

The market situation influences the differentiation as well. It was identified through the interviews, that the company's sourcing function benefits from the company's size, as the suppliers are mainly willing to cooperate with the buyer's company. It was also pointed out in multiple interviews, that they have long and good relationships with the suppliers as well. On a further note, they might sometimes face regional capacity constraints, as the company is significant in its size:

*"... however, there is always a limited amount of capacity, the number of drivers and cars cannot suddenly increase." [Interviewee 20]*

*Continuous improvement through supplier collaboration and feedback:*

Companies might be too eager to send data to the suppliers without any follow-up practices. Sometimes, the importance of communication is forgotten even though it is vital for the whole process. It was identified, that supplier performance measuring provides a platform for inter-organizational information sharing, and that having regular communication practices would ensure it. More or less, the focus should be on what works and what does not work, according to the interviewees. Information sharing should be based on bilateral communication, where the buyer and supplier openly discuss about performance.

*"It could be good to have frequent follow-up meetings where you also present what has been working well, what can be improved, how do we get more of these win-win situations? How can we help each other to grow and so on." [Interviewee 7]*

Comparing the suppliers might increase the ability to share information to the supplier. When comparing suppliers in, for example, a supplier selection situation, the outcome might point out areas where suppliers might need to improve themselves to be more competitive.

*"... the suppliers usually want to know how we view them, and then they typically go back and try to fix those things [that we notify them about]" [Interviewee 6]*

*Goal alignment between buyer and supplier:*

Goal alignment has significance in supplier measuring. First of all, the measuring can be used to communicate the buyer's desires, so it is not insignificant if the supplier is informed about these metrics. The existence of measuring does not necessarily mean that the supplier is pressured to act upon the buyer's demand, but that they are made aware of the hoped goals, that can enable the alignment of the two companies. As suggested, the end goals of the organization should be well-defined to make measurements for follow-up. Sometimes, the measurements might have contradictory goals, and this should be considered to see what is really important. In an ideal situation, the measurements are agreed with the supplier:

*"Well, ideally, and in general, we have the same metrics in use that the supplier also uses for self-evaluation and development. We should not have our own metrics that we do not share with*

*the supplier. Rather, these are our common metrics, where we define what is the target level in our operational life, and then we both work towards this common goal." [Interviewee 16]*

### 4.2.3 Combining all the dimensions

In chapter 4.1.3, the managing internal process and design themes were shown in a figure, as the focus of the dimensions were operational, i.e., how to make the system work in a company context. Inter-organizational and intra-organizational aggregate dimensions have a more strategic focus, but still provide some implications about the operational side of the SPMS (Figure 17). Ultimately, the SPMS provides a tool to manage the performance of the suppliers through guiding the organization. However, if the use is only on intra-organizational learning, i.e., the results are not shared outside of the organization, the outcome is more or less about improving the supply base. Through sharing the results, the focus can shift from identifying and changing suppliers, to actually trying to improve the existing suppliers.

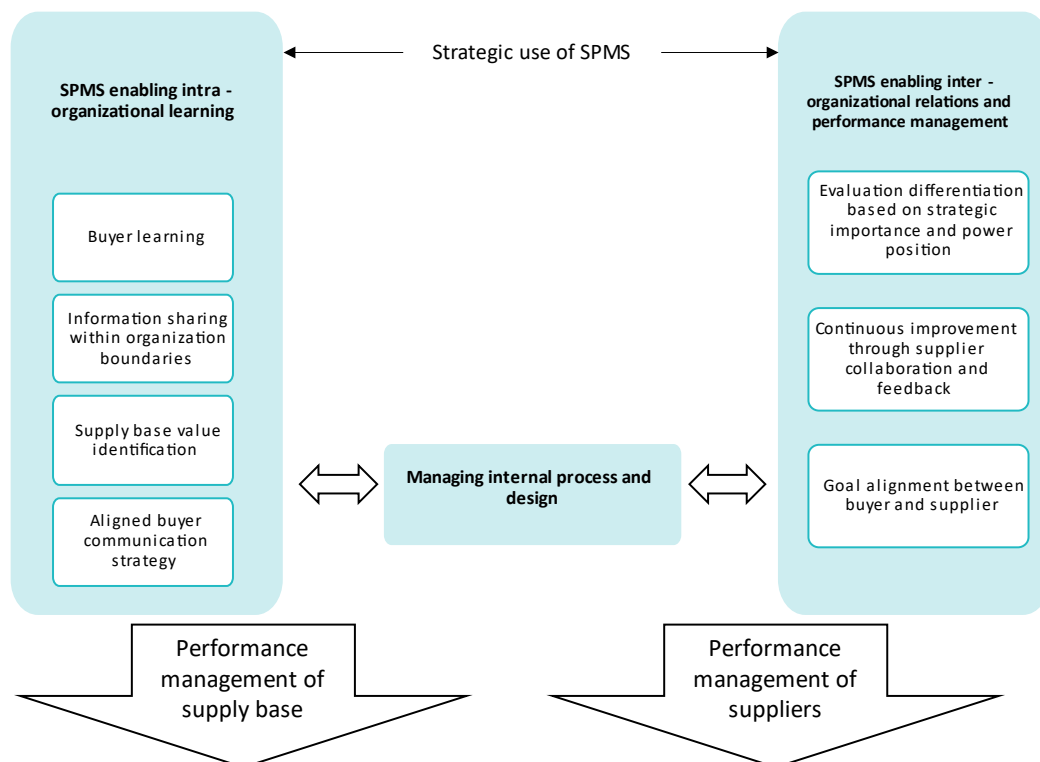


Figure 17. Supplier performance measurement system roles in an organization

## 5. Discussion

This research was guided by the first research question: *How to design a supplier performance measurement system?* SPMS design has been previously depicted as a straightforward process that goes from one step to another. Apart from what has been previously presented, it is more dynamic process, as the various factors that are included in design (e.g., available data, measurements, planned use) are interconnected. Moreover, a company should consider designing a primary process (Data collection, collation and use) and a number of supporting processes (Frequency, stakeholders, and an owner of the communication process) to be able to design feasible SPMS. The primary process and supporting processes provide an answer to the second research question: *What are the key factors of a supplier performance measurement system?*

This study approached the SPMS design without predefining the intended use of SPMS (such as supplier development, categorization, or selection), allowing interviewees to contemplate on the subject more broadly. Throughout the interviews, the interviewees connected supplier evaluation to selection, monitoring, or other uses that present the operational use of SPMS. However, the SPMS was not only limited to performance measuring. While using the system, it allows information sharing inside or outside of the company that was found to be critical in this study. The presented framework combines both, as a large organization has a need to share information inside and outside of the company to be able to manage their suppliers but also to improve their internal information sharing practices.

### 5.1 Design framework

The resulting framework divided the SPMS design into two areas: primary and supporting processes. For the sake of simplicity, the figure presents a linear process, even though later phases might bring requirements to previous phases, or supporting processes might lead to changes in the primary process. The primary process starts from the available data (making processes for collecting data or using existing data), measurements, and defining the needs and goals. The following phase is the data processing and eventually, the produced information can be used in the decision-making and/or communication. The supporting processes

consist of deciding the collecting frequency, identifying the relevant stakeholders, and defining a process owner of the model. The design framework is presented in Figure 18.

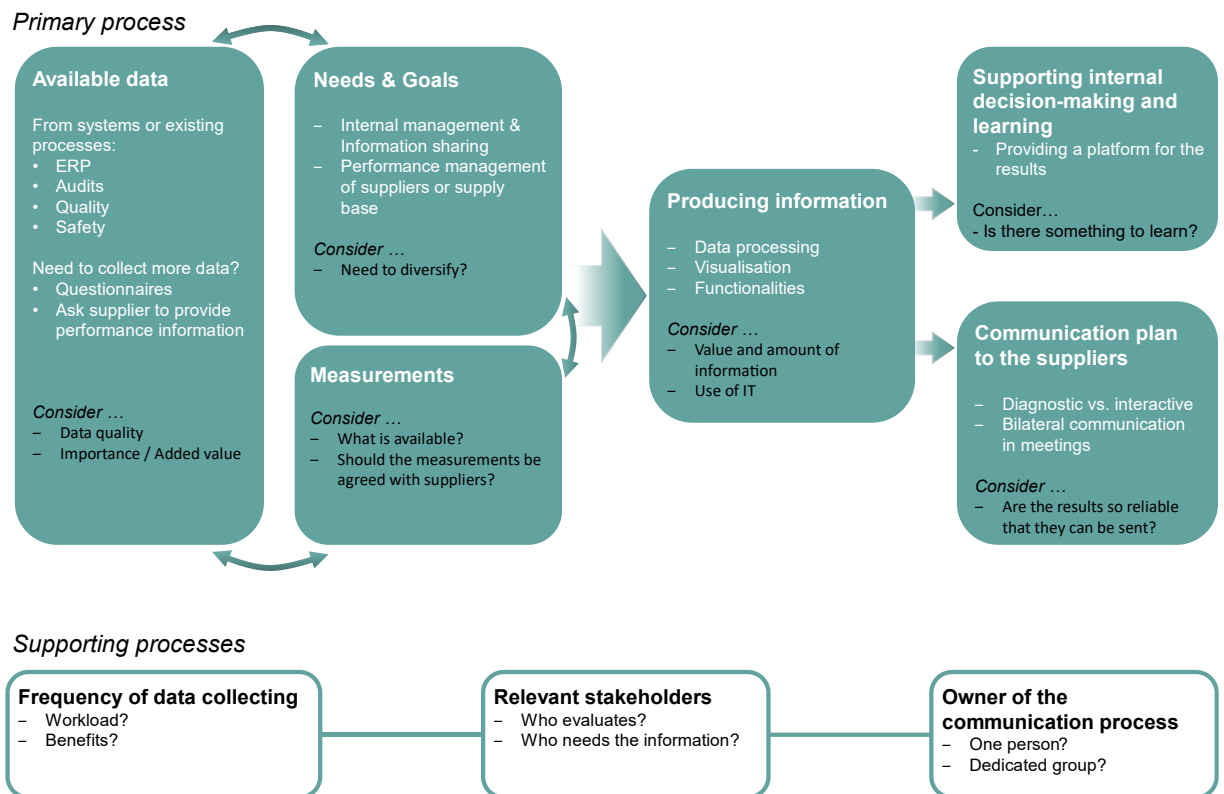


Figure 18. Supplier performance measurement system design framework

The primary process follows three steps that can be found from advanced decision-making models and also from Luzzini et al.'s (2014) study. Evaluation and monitoring models usually discuss the metrics, the ways to produce information, and the decision-making problems that can be solved with these models (Dutta et al., 2021). Luzzini et al.'s (2014) study considers these aspects as well. However, they take a practical point of view, embedding the SPMS into strategy, and also discuss the need to define communication and stakeholders. To build upon these studies, the starting point in practice is more or less defining the primary process as it goes from data to actual use of the SPMS. Supporting processes are critical to make it more systematic by, for example, defining how often the results are produced and used, finding the relevant stakeholders to evaluate, and defining someone who is in charge of the process, e.g., by communicating the results to the suppliers. They are supporting as



they are not as critical as the primary process. For example, even though the frequency is not systematic, it does not mean that there are not any benefits gained from SPMS.

### **5.1.1 Primary Process**

The primary process starts with careful consideration of available data, measurements, and needs and goals of the measuring. These three are overlapping in this process, as they are intertwined, meaning that each of these factors either constraint or establish goals for each factor. Further, the data is processed through certain logic to produce usable information for decision-making and communication.

#### **Available data:**

As a starting point for an SPMS, it should be assessed what kind of data is available. These vary from company to company, but in a manufacturing company, an ERP is a good source of data, and other processes might also produce supplier specific data, e.g., safety management. SPMS does not have to be constrained by the available data, and alternative sources for data should be considered. One can collect data from various people in an organization, which was found important during the interviews, as it enables hearing feedback from employees that you are not working closely with and allows information sharing to happen inside of the company. The supplier can also provide data that the buyer cannot measure and does not have visibility to.

In a perfect world, data quality is considered before using the system, to maximize performance outcome of the measuring. Reliability is one of the quality factors that stems from measuring results that actually reflect the supplier's performance. Unreliable results might be due to different measuring practices, workload, amount of information and also, there might be barriers to the supplier's performance that the buyer should be aware of. Evaluators' bias should be considered as well, as the evaluators might behave in their own interest and threaten suppliers with bad ratings (Hawkins et al., 2020). On top of quality factors, the data collection is an investment from the company, and the importance and added value of the measuring should be considered. As pointed out by Neely et al. (1995), the measurements can be too costly to have in regard to the information gained.

### **Needs & Goals:**

Companies should consider their needs and goals when developing a SPMS, as it influences 1) what data is needed and 2) what kind of measurements there should be. One should not skip defining the needs and goals, and go straight to measuring, as the measurements are not universal (Lima-Junior & Carpinetti, 2016). Goals can be, e.g., for supplier selection, supplier development, or monitoring. It was identified that SPMS is not limited to those concepts. In a large organization, the SPMS can serve multiple needs. Eventually, SPMS can enable internal management and information sharing, which can be a valuable outcome for a large company. It is essential in large companies, where employees, who manage suppliers, might be far from the actual day-to-day work and do not witness the performance of the supplier. Further, the intra-organizational use enables performance management of the supplier base, where the focus is not to develop the existing suppliers but select new suppliers and keep or phase out existing suppliers to gain performance advantages from the supplier base. Also, it can be used to identify potential from the supplier base. Inter-organizational use enables systematic information sharing between the buyer and the supplier, which in turn, enables the supplier performance management.

Apart from the more organizational point of view, the actual use might be defined into two categories suggested by literature. The focus of SPMS might be diagnostic or interactive, which in turn influences the data and measuring (Maestrini et al., 2021; Henri, 2006). Diagnostic could be referred to as more formal measuring, e.g., are the goals are reached. And the interactive is more open, e.g., what could be improved. Depending on the approaches, both have their time and place in SPMS. However, interactive use might be heavier to implement, as it is not as defined. Thus, its use might be limited to a few suppliers, e.g., strategic suppliers.

All in all, one should consider the diversification of the SPMS based on strategic importance and power, according to this research. It is more or less about the management of the workload, as having frequent qualitative evaluation is extra work. Strategically important suppliers are not necessarily the suppliers with the most spend, but strategic importance can also stem from poor performance that further influences the whole production. Sometimes, the supplier might be hard to replace, and therefore, the ideal solution might be to try to develop

the supplier. Power has significance over how the measuring takes place. Based on this research, all of the ways that power influences the use of SPMS cannot be exhaustively recognized. However, through the interviews, it was noted that the case company has relatively much power in the area, and mainly the suppliers are interested in cooperation, thus possibly increasing the benefits gained from SPMS. As the SPMS is an investment of the company's resources, one should carefully choose suppliers where the greatest potential lies, e.g., suppliers that are interested in development as well. Power imbalance does not mean that there are no benefits gained from the SPMS, as Jääskeläinen (2021) identified, that the powerful supplier might be interested in strategic development with the buyer, despite the power imbalance. Also, the supplier's strategy might be serving the customer in the best way (Maestrini et al., 2018c), suggesting that they might be interested in the performance feedback, irrespective of the power situation. It was also identified in the interviews, that it is better to say something about the buying company's desires than to be silent.

### **Measurements:**

Based on the literature review, the measurements should reflect the company's strategy (Neely et al. 1995). When developing an SPMS, there should be careful consideration of the importance and added value of measurements. To identify relevant measurements, one can ask the specialists, or the measurements could be agreed in consensus with the supplier. Without specifically commenting on what should be measured, one could think through the goal of the measuring, how defined the measuring should be. For example, the evaluator could be asked to grade perceived cooperation from 1 to 5 or take a more open approach, e.g., comment on how well the cooperation has worked. The defined approach is easy for quantitative data (e.g., delivery precision) that is available from the system in the first place. On the other hand, qualitative data might provide more insights. However, the amount of qualitative evaluation should be considered carefully, as it easily increases the amount of work, which might lead to non-use.

### **Producing information:**

Just collecting or measuring something is not enough to make valuable information. The data processing can result in, for example, a dashboard (Pal & Kumar 2008) or a ranking. In

literature, the visualisations were rarely brought up, even though it is critical that resulting data is presented in informative manner. A list or a ranking of suppliers might not be most beneficial for supplier performance management due to a lack of individual feedback. On the other hand, lists can provide identification of underperforming suppliers, where the supplier development activities would be most beneficial (Araz & Ozkarahan, 2007). Someone else than the end-user should be responsible for data collection and processing, e.g., it should be a shared responsibility within the buying company, as the responsibility might become unbearable for one person if they have to make sure that data is reliable and informative. It was noted that the amount of information should be managed, as sometimes, the solutions might contain so much information that the result is not informative anymore. It should be considered, what is valuable information, e.g., worth to show. Based on the interviews, it was suggested that maybe there should be a summary page and the possibility to go behind these measurements if needed. It should be also considered how the usage of IT supports the SPMS, as it can significantly decrease the need for manual work, which might have a positive impact on the outcome.

### **Supporting internal decision-making and learning:**

The produced information can be used to support decision-making in the company. For example, a lot of the performance evaluation models are used for supplier selection (Patrucco et al., 2022). From the interviews, it was identified that not only decision-making is enabled by the use of SPMS, but also learning. It can help to identify potential from the supplier base, which in turn might help to make the supplier base more competitive.

The company might consider using SPMS for the purpose of sharing information internally. For example, as companies tend to use the same suppliers across multiple regions, it might be beneficial to share the known causes for disruptions across the organization, to learn and look, if the same problem with this supplier exists in a different region. If the information is available for later use, one can learn from these cases. However, some level of criticality might be beneficial, as the information is 'old news' and does not reflect the current situation. On the other hand, it cannot be ignored, that the results might suggest that the supplier is struggling, e.g., financially, which in turn means, that they should be under scrutiny to avoid disruption by bankruptcy.

### **Communication plan to suppliers:**

If the goal of SPMS is supplier performance management, communication should take place. Communication with suppliers might be only synthetic, as described by Maestrini et al. (2018c), where the communication is uni-directional and the results are sent without any follow-up. To get more optimal results from the SPMS, it might be more beneficial to aim for bilateral communication. It has been found, that so-called socialization practices (e.g. meetings) are important to gain improved supplier performance from SPMS (Cousins et al., 2008). Similarly, it was found in the interviews, that performance measuring provides a base for communication and inter-organizational information sharing. It can be used to identify pain points, where either the supplier or the buyer should improve. As previously described, the communication can take a diagnostic point of view, e.g., the discussions will be around what can be done to reach set goals. The discussion is more or less on the follow-up of the measurements. On the other hand, the core of interactive use is communication regarding ‘where we can find these win-win situations’, rather than following some specific measurements even though results provide a foundation for discussions.

Communication can be formalized with meetings and other follow-ups regarding the results. Meetings allow the supplier to comment on their results and provide their insights. As the communication influences the outcome from the use of SPMS (Cousins et al., 2008; Maestrini et al., 2018c), it is not irrelevant how it takes place. In the interviews, it was suggested, that one should be careful to avoid unfairness. For example, the data quality has significance if the results can be used to manage the performance of the suppliers. Therefore, it should be considered if the quality of the results is sufficient enough that it can be used in communication. On the other hand, the data quality should not completely inhibit the use of SPMS, as it can be a platform where dialogue happens (e.g., if the buyer’s results are wrong, the supplier can just point that out). They can also point out if the results are due to the actions of the buyer, i.e., the buyer’s action is barrier to the supplier’s performance.

### **5.1.2 Supporting Processes**

There are some other processes that should be defined in SPMS. There were a few factors raised during the analysis and literature review. These processes support the whole SPMS.

The **frequency of data collection** should be defined in SPMS to maintain a systematic way of using it. For data available from the systems, the results could be available daily. For data collection from employees or suppliers, the frequency is less, e.g., once a year. The frequency might be increased later on, but when designing an SPMS, it might be beneficial to start with yearly assessments. It was seen as beneficial to start, and then see how it goes and affects the workload. On the other hand, it might influence the benefits and the reliability, if the frequency is too low and the results are essentially too late. However, the purpose of SPMS is not necessarily to be an alerting system.

Identifying **relevant stakeholders** should be approached in two ways: who evaluates, and who needs the information (i.e., the end user of the system). First of all, it is important to identify who should be evaluating the performance of the suppliers. These are the people who witness the performance of the supplier. Secondly, it is important to identify the end-users, as they might have varying needs for the information making this connected to the primary process. The SPMS might be different if the intended use is more on an operational level when compared to a case where the use is strategic.

SPMS should have a defined **owner of the communication process**, meaning the person, who has the responsibility of communication and who possibly oversees that the information is there to be communicated. In a large organization, there might be a group of people that is responsible for supplier performance management (Lambert & Schwieterman, 2012). It does not mean that the person or group is responsible for the data quality, but rather that they act as gatekeepers, ensuring that the quality of the communicated information is top-notch. The people responsible for the data quality are elsewhere. It might lead to a too heavy workload, if data quality is also the responsibility of the person who communicates the results.

## 5.2 Theoretical Implications

This study approached the design from another point of view than previous studies on SPMS design. The stance of this thesis was to identify what should be done, making the focus shift from what is done, to the ideal solution. Luzzini et al.'s (2014) study pointed out, that SPMS realized benefits might be different than intended in organizations. However, it might be due to a lack of knowledge of how the SPMS can be designed and used. The study's approach

unravelling insights about the SPMS design and quality management of it, that were not previously discussed in the literature. First of all, SPMS's role in an organization is larger than anticipated. End-users would like to use it for internal information sharing which in turn allows the internal decision-making and supplier management to be grounded in data. Also, the SPMS should be used to manage supplier's performance through communication of the results and dialogue. This study contributed to the literature by providing a new model for SPMS design that provides deep practical insights about the process and how it should be designed.

Measurements are connected to the use and therefore, they should be considered early on. Hald and Ellegaard (2011) pointed out that unrealistic measurements led to unmotivated suppliers showing the importance of measurements and understanding the use of the system. It is a problematic situation if the realised benefits are different than the intended as the design does not reach its full potential as measurements are not universal. In companies, producing new data is a challenge due to workload and therefore, it should be considered what is available and what performance information can be easily connected to the supplier all while maintaining the quality of measuring. Sustainability performance is becoming mandatory for large companies to follow from suppliers making it worthwhile to have functional SPMS in place.

### **5.3 Practical implications**

The illustrations of SPMS design process can be used to guide companies on what are the phases of SPMS design, and also point out some critical factors that should be considered during the design. The framework should guide how to design best-in-class SPMS making the investment in SPMS worthwhile. In practice, there should be a strong focus on the chosen measurements to avoid unreliable results from SPMS. Nevertheless, the measurements should not be only about numerical results as textual feedback and communication are important to gain performance improvement results from SPMS. The quality of the SPMS should be considered while designing it. The starting point should be simple to see what provides value in a company's setting and avoid making the process too heavy workload-wise as it might lead to non-use.

## 6. Conclusions

Two broad categories that are not previously discussed in the context of supplier performance measurement systems emerged from the analysis: intra-organizational and inter-organizational use. In a large organization, the role of SPMS is not necessarily only the management of suppliers, but also enabling internal information sharing. People who work in procurement do not necessarily have good visibility on how suppliers are performing, clouding the judgement of management. Also, single observations of the supplier's performance might lead to bias, so a systematic use of SPMS produces more objective information about the supplier's performance. Further, it enables relations between the buyer and the supplier through better information sharing, therefore suggesting the role of SPMS to be bigger than merely a performance measurement system. The communication of the results was seen as key, and the focus should be on dialogue, and trying to find win-win situations and discover, what could be improved in the cooperation. SPMS can indicate how the supplier should be managed, e.g., the hopes and needs of the buying company can be communicated to the supplier.

### 6.1 Limitations and Suggestions for Future Research

There are some limitations that should be considered in future studies. This study was conducted as a single case study, that could be extended by involving the supplier's point of view in the discussion. There is room for a study, that could uncover what kind of performance feedback helps the supplier to improve, and what are the best practices of measuring from the point of view of the supplier. Also, if the supplier has an indifferent reaction to SPMS, studies should try to qualitatively uncover the reasons behind the reaction.

Literature has been focused on developing various models to solve the multi-criteria decision-making problem that performance evaluation can be. It could be an interesting case to take a design science approach to implement an advanced model as a part of SPMS, giving insights about what kind of data collection habits there should be to make monitoring work, and bring benefits into a company's decision-making or supplier performance management.



## Reference list

- Araz, C., & Ozkarahan, I. (2007). Supplier evaluation and management system for strategic sourcing based on a new multicriteria sorting procedure. *International journal of production economics*, 106(2), 585-606.
- Arnoğlu, M. Ö., Sarkis, J., & Dhavale, D. G. (2021). Selection of suppliers using Bayesian estimators: a case of concrete ring suppliers to Eurasia Tunnel of Turkey. *International Journal of Production Research*, 59(18), 5678-5689.
- Bai, C., Kusi-Sarpong, S., Badri Ahmadi, H., & Sarkis, J. (2019). Social sustainable supplier evaluation and selection: a group decision-support approach. *International Journal of Production Research*, 57(22), 7046-7067.
- Banaeian, N., Mobli, H., Fahimnia, B., Nielsen, I. E., & Omid, M. (2018). Green supplier selection using fuzzy group decision making methods: A case study from the agri-food industry. *Computers & Operations Research*, 89, 337-347.
- Belotserkovskiy, R., Sewak, J., Teodorian, A., & Lietke, B. (2018). *Contracting for performance: Unlocking additional value*. McKinsey & Company. <https://www.mckinsey.com/capabilities/operations/our-insights/contracting-for-performance-unlocking-additional-value>
- Benton Jr, W. C., Prahinski, C., & Fan, Y. (2020). The influence of supplier development programs on supplier performance. *International Journal of Production Economics*, 230, 107793.
- Berthelsen, C. B., Lindhardt, T., & Frederiksen, K. (2017). A discussion of differences in preparation, performance and postreflections in participant observations within two grounded theory approaches. *Scandinavian Journal of Caring Sciences*, 31(2), 413–420.
- 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.
- Bruno, G., Esposito, E., Genovese, A., & Passaro, R. (2012). AHP-based approaches for supplier evaluation: Problems and perspectives. *Journal of purchasing and supply management*, 18(3), 159-172.
- Bullinger, H.-J., Kühner, M., & Van Hoof, A. (2002). Analysing supply chain performance using a balanced measurement method. *International Journal of Production Research*, 40(15), 3533-3543.
- Buffa, F. P., & Ross, A. D. (2011). Measuring the Consequences of Using Diverse Supplier Evaluation Teams: A Performance Frontier Perspective. *Journal of Business Logistics*, 32(1), 55–68.
- Cao, Y., Umberger, G., Shelton, A., & Butts, T. (2022). Development of a Supplier Scorecard for a Medical Device Company to Formalize Its Supplier Monitoring and Re-evaluation Process Toward ISO13485: 2016. *Journal of Pharmaceutical Innovation*, 17, 1295–1304.

- Cannon, J. P., & Perreault Jr, W. D. (1999). Buyer–seller relationships in business markets. *Journal of marketing research*, 36(4), 439-460.
- Carr, A. S., & Pearson, J. N. (1999). Strategically managed buyer–supplier relationships and performance outcomes. *Journal of operations management*, 17(5), 497-519.
- Chen, C. T., Lin, C. T., & Huang, S. F. (2006). A fuzzy approach for supplier evaluation and selection in supply chain management. *International journal of production economics*, 102(2), 289-301.
- Chen, Y. S., Rungtusanatham, M. J., & Goldstein, S. M. (2019). Historical supplier performance and strategic relationship dissolution: Unintentional but serious supplier error as a moderator. *Decision Sciences*, 50(6), 1224-1258.
- Chiarini, A. (2017). Environmental policies for evaluating suppliers' performance based on GRI indicators. *Business Strategy and the Environment*, 26(1), 98-111.
- Choy, K.L., Chow, H. K., Lee, W., & Chan, F. T. (2007). Development of performance measurement system in managing supplier relationship for maintenance logistics providers. *Benchmarking : an International Journal*, 14(3), 352–368.
- Clauss, T., & Tangpong, C. (2019). Perception-based supplier attributes and performance implications: a multimethod exploratory study. *Journal of Supply Chain Management*, 55(4), 34-66.
- Corbin, J., & Strauss, A. L. (2008). *Basics of qualitative research : techniques and procedures for developing grounded theory* (3<sup>rd</sup> edition.). SAGE.
- Cousins, P. D., Lawson, B., & Squire, B. (2008). Performance measurement in strategic buyer-supplier relationships: the mediating role of socialization mechanisms. *International Journal of Operations & Production Management*, 28(3), 238-258.
- Deloitte. (2017). *Performance Management in Supply Chain and Operations – Steering value chain activities towards exceptional performance*. <https://www2.deloitte.com/de/de/pages/operations/articles/performance-management-in-supply-chain-and-operations-steering-value-chain-activities.html>
- Dey, P.K., Bhattacharya, A., Ho, W., & Clegg, B. (2015). Strategic supplier performance evaluation: A case-based action research of a UK manufacturing organisation. *International Journal of Production Economics*, 166, 192–214.
- Dossi, A., & Patelli, L. (2010). You learn from what you measure: financial and non-financial performance measures in multinational companies. *Long Range Planning*, 43(4), 498-526.
- Dutta, P., Jaikumar, B., & Arora, M. S. (2021). Applications of data envelopment analysis in supplier selection between 2000 and 2020: A literature review. *Annals of Operations Research*, 315, 1399–1454.
- Ellram, L. M., & Carr, A. (1994). Strategic purchasing: a history and review of the literature. *International journal of purchasing and materials management*, 30(1), 9-19.
- Emiliani, M. L. (2010). Historical lessons in purchasing and supplier relationship management. *Journal of Management History*, 16(1), 116–136.

- Eriksson, P., & Kovalainen, A. (2016). *Qualitative methods in business research* (2<sup>nd</sup> edition.). Sage Publications.
- Faraz, A., Sanders, N., Zacharia, Z., & Gerschberger, M. (2018). Monitoring type B buyer–supplier relationships. *International Journal of Production Research*, 56(18), 6225–6239.
- Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D., & Neely, A. (2007). Towards a definition of a business performance measurement system. *International Journal of Operations & Production Management*, 27(8), 784–801.
- Genovese, A., Lenny Koh, S. C., Kumar, N., & Tripathi, P. K. (2014). Exploring the challenges in implementing supplier environmental performance measurement models: a case study. *Production Planning & Control*, 25(13-14), 1198–1211.
- Dyer, W. G., & Wilkins, A. L. (1991). Better Stories, Not Better Constructs, to Generate Better Theory: A Rejoinder to Eisenhardt. *The Academy of Management Review*, 16(3), 613–619.
- Gioia, D. (2021). A systematic methodology for doing qualitative research. *The Journal of Applied Behavioral Science*, 57(1), 20–29.
- Gioia, D., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, 16(1), 15–31.
- Govindan, K., Aditi, Dhingra Darbari, J., Kaul, A., & Jha, P. C. (2021). Structural model for analysis of key performance indicators for sustainable manufacturer–supplier collaboration: A grey-decision-making trial and evaluation laboratory-based approach. *Business Strategy and the Environment*, 30(4), 1702–1722.
- Goldberg, J., & Schiele, H. (2018). Early Supplier Integration: Assessing Supplier Innovation Ideas. *IEEE Engineering Management Review*, 46(3), 94–102.
- Guneri, A. F., & Kuzu, A. (2009). Supplier selection by using a fuzzy approach in just-in-time: A case study. *International Journal of Computer Integrated Manufacturing*, 22(8), 774–783.
- Hawkins, T.G., Gravier, M. J., & Muir, W. A. (2020). The role of supplier performance evaluations in mitigating risk: Assessing evaluation processes and behaviors. *Industrial Marketing Management*, 87, 2–17.
- Henri, J. F. (2006). Management control systems and strategy: A resource-based perspective. *Accounting, organizations and society*, 31(6), 529–558.
- Ho, W., Xu, X., & Dey, P. K. (2010). Multi-criteria decision making approaches for supplier evaluation and selection: A literature review. *European Journal of Operational Research*, 202(1), 16–24.
- Ho, W., Dey, P. K., & Lockström, M. (2011). Strategic sourcing: a combined QFD and AHP approach in manufacturing. *Supply Chain Management: An International Journal*, 16(6), 446–461.
- Humphreys, P., McCloskey, A., McIvor, R., Maguire, L., & Glackin, C. (2006). Employing dynamic fuzzy membership functions to assess environmental performance in the supplier selection process. *International Journal of Production Research*, 44(12), 2379–2419.

- Humphreys, P., McIvor, R., & Chan, F. (2003). Using case-based reasoning to evaluate supplier environmental management performance. *Expert systems with Applications*, 25(2), 141-153.
- Jääskeläinen, A. (2021). The relational outcomes of performance management in buyer-supplier relationships. *International Journal of Production Economics*, 232, 107933.
- Jääskeläinen, A. (2018). Comparison of performance measurement in different purchasing and supply management practices. *International Journal of Productivity and Performance Management*, 67(8), 1290–1309.
- Jääskeläinen, A., & Thitz, O. (2018). Prerequisites for performance measurement supporting purchaser-supplier collaboration. *Benchmarking : an International Journal*, 25(1), 120–137.
- Kilkenny, M. F., & Robinson, K. M. (2018). Data quality: “Garbage in – garbage out.” *Health Information Management Journal*, 47(3), 103–105.
- Kotula, M., Ho, W., Dey, P. K., & Lee, C. K. M. (2015). Strategic sourcing supplier selection misalignment with critical success factors: Findings from multiple case studies in Germany and the United Kingdom. *International Journal of Production Economics*, 166, 238-247.
- Kraljic, P. (1983). HBR. *Harvard business review*.
- Kronemeyer, L. L., Kotzab, H., & Moehrle, M. G. (2022). Analyzing technological competencies in the patent-based supplier portfolio: introducing an approach for supplier evaluation using semantic anchor points and similarity measurements. *International Journal of Operations & Production Management*, 42(11), 1732-1759.
- Kähkönen, A-K. (2014). Conducting a case study in supply management. *Operations and Supply Chain Management: An International Journal*, 4(1), 31-41.
- Lambert, D. M., & Enz, M. G. (2017). Issues in Supply Chain Management: Progress and potential. *Industrial Marketing Management*, 62, 1–16.
- Lambert, D. M., & Schwieterman, M. A. (2012). Supplier relationship management as a macro business process. *Supply chain management: an international journal*, 17(3), 337-352.
- Lima-Junior, F. R., & Carpinetti, L. C. R. (2016). Combining SCOR® model and fuzzy TOPSIS for supplier evaluation and management. *International Journal of Production Economics*, 174, 128-141.
- Le Dain, M. A., Calvi, R., & Cheriti, S. (2011). Measuring supplier performance in collaborative design: proposition of a framework. *R&d Management*, 41(1), 61-79.
- Lee, A. H. (2009). A fuzzy AHP evaluation model for buyer–supplier relationships with the consideration of benefits, opportunities, costs and risks. *International Journal of Production Research*, 47(15), 4255-4280.
- Luthra, S., Govindan, K., Kannan, D., Mangla, S. K., & Garg, C. P. (2017). An integrated framework for sustainable supplier selection and evaluation in supply chains. *Journal of Cleaner Production*, 140, 1686–1698.

- Luzzini, D., Amann, M., Caniato, F., Essig, M., & Ronchi, S. (2015). The path of innovation: purchasing and supplier involvement into new product development. *Industrial Marketing Management*, 47, 109-120.
- Luzzini, D., Caniato, F., & Spina, G. (2014). Designing vendor evaluation systems: An empirical analysis. *Journal of Purchasing and Supply Management*, 20(2), 113–129.
- Maestrini, V., Patrucco, A. S., Luzzini, D., Caniato, F., & Maccarrone, P. (2021). Supplier performance measurement system use, relationship trust, and performance improvement: a dyadic perspective. *The International Journal of Logistics Management*, 32(4), 1242-1263.
- Maestrini, V., Luzzini, D., Caniato, F., Maccarrone, P., & Ronchi, S. (2018a). The impact of supplier performance measurement systems on supplier performance: a dyadic lifecycle perspective. *International Journal of Operations & Production Management*, 38(11), 2040-2061.
- Maestrini, Luzzini, D., Caniato, F., & Ronchi, S. (2018b). Effects of monitoring and incentives on supplier performance: An agency theory perspective. *International Journal of Production Economics*, 203, 322–332.
- Maestrini, V., Maccarrone, P., Caniato, F., & Luzzini, D. (2018c). Supplier performance measurement systems: Communication and reaction modes. *Industrial Marketing Management*, 74, 298-308.
- Maestrini, V., Luzzini, D., Maccarrone, P., & Caniato, F. (2017). Supply chain performance measurement systems: A systematic review and research agenda. *International Journal of Production Economics*, 183, 299-315.
- Magnani, G., & Gioia, D. (2023). Using the Gioia Methodology in international business and entrepreneurship research. *International Business Review*, 32(2), 102097.
- Moretto, A., Grassi, L., Caniato, F., Giorgino, M., & Ronchi, S. (2019). Supply chain finance: From traditional to supply chain credit rating. *Journal of Purchasing and Supply Management*, 25(2), 197-217.
- Morgan, C., & Dewhurst, A. (2007). Using SPC to measure a national supermarket chain's suppliers' performance. *International Journal of Operations & Production Management*, 27(8), 874-900.
- Mortensen, M. H., Freytag, P. V., & Arlbjørn, J. S. (2008). Attractiveness in supply chains: a process and maturity perspective. *International Journal of Physical Distribution & Logistics Management*, 38(10), 799-815.
- Muralidharan, C., Anantharaman, N., & Deshmukh, S. G. (2001). Vendor rating in purchasing scenario: a confidence interval approach. *International Journal of Operations & Production Management*, 21(10), 1305-1326.
- Nair, A., Jayaram, J., & Das, A. (2015). Strategic purchasing participation, supplier selection, supplier evaluation and purchasing performance. *International Journal of Production Research*, 53(20), 6263–6278.

- 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.
- Niazmand, K., Mirzazadeh, A., & Rezaie, K. (2014). A Fuzzy SQFE approach in supplier's performance monitoring. *International Journal of Production Research*, 52(22), 6841-6862.
- Pagell, M., Wu, Z., & Wasserman, M. D. (2010). Thinking differently about purchasing portfolios: an assessment of sustainable sourcing. *Journal of Supply Chain Management*, 46(1), 57-73.
- Pal, P., & Kumar, B. (2008). "16T": toward a dynamic vendor evaluation model in integrated SCM processes. *Supply Chain Management: An International Journal*, 13(6), 391-397.
- Park, J., Shin, K., Chang, T. W., & Park, J. (2010). An integrative framework for supplier relationship management. *Industrial Management & Data Systems*, 110(4), 495-515.
- Parthiban, P., Zubar, H. A., & Katarar, P. (2013). Vendor selection problem: a multi-criteria approach based on strategic decisions. *International Journal of Production Research*, 51(5), 1535-1548.
- Patrucco, A., Frattini, F., & Di Benedetto, A. (2022). Characteristics of supplier performance measurement systems in collaborative innovation projects: the role of the purchasing department. *Supply Chain Management: An International Journal*, 27(2), 207-231.
- Prahinski, C., & Benton, W. C. (2004). Supplier evaluations: communication strategies to improve supplier performance. *Journal of operations management*, 22(1), 39-62.
- Prahinski, C., & Fan, Y. (2007). Supplier evaluations: the role of communication quality. *Journal of Supply Chain Management*, 43(3), 16-28.
- Prajogo, D., Chowdhury, M., Yeung, A. C., & Cheng, T. C. E. (2012). The relationship between supplier management and firm's operational performance: A multi-dimensional perspective. *International journal of production economics*, 136(1), 123-130.
- Purdy, L., & Safayeni, F. (2000). Strategies for supplier evaluation: a framework for potential advantages and limitations. *IEEE transactions on Engineering Management*, 47(4), 435-443.
- PwC (2023). *Digital operations*. <https://www.pwc.com/gx/en/services/consulting/digital-operations.html>
- Rezaei, J., Kadziński, M., Tsimopoulou, V., & Tavasszy, L. (2022). Embedding carbon impact assessment in multi-criteria supplier segmentation using ELECTRE TRI-rC. *Annals of Operations Research*, 312(2), 1445-1467.
- Roodhooft, F., & Konings, J. (1997). Vendor selection and evaluation an activity based costing approach. *European journal of operational research*, 96(1), 97-102.
- Ross, A., & Buffa, F. P. (2009). Supplier post performance evaluation: the effects of buyer preference weight variance. *International Journal of Production Research*, 47(16), 4351-4371.
- Ross, A. D., Kuzu, K., & Li, W. (2016). Exploring supplier performance risk and the buyer's role using chance-constrained data envelopment analysis. *European Journal of Operational Research*, 250(3), 966-978.

- Ross, A. D., Buffa, F. P., Droge, C., & Carrington, D. (2009). Using buyer–supplier performance frontiers to manage relationship performance. *Decision sciences*, 40(1), 37-64.
- Saldaña, J. (2013). *The Coding Manual for Qualitative Researchers*, (2<sup>nd</sup> ed.). SAGE Publications, 1-303.
- Sarkis, & Talluri, S. (2002). A Model for Strategic Supplier Selection. *The Journal of Supply Chain Management*, 38(4), 18–28.
- Saunders, Lewis, P., & Thornhill, A. (2016). *Research methods for business students* (7<sup>th</sup> ed.). Pearson Education.
- Schiele, H. (2019). Purchasing and supply management. In Zijm, H., Klumpp, M., Regattieri, A., & Heragu, S. (Eds.), *Operations, Logistics and Supply Chain Management*, (pp. 45-73). Springer.
- Shen, F., Xu, J., & Xu, Z. (2015). An automatic ranking approach for multi-criteria group decision making under intuitionistic fuzzy environment. *Fuzzy Optimization and Decision Making*, 14, 311-334.
- Simpson, P.M., Siguaw, J. A., & White, S. C. (2002). Measuring the Performance of Suppliers: An Analysis of Evaluation Processes. *The Journal of Supply Chain Management*, 38(4), 29–41.
- Steward, M. D., Morgan, F. N., Crosby, L. A., & Kumar, A. (2010). Exploring cross-national differences in organizational buyers' normative expectations of supplier performance. *Journal of International Marketing*, 18(1), 23-40.
- Sundtoft Hald, K., & Ellegaard, C. (2011). Supplier evaluation processes: the shaping and re-shaping of supplier performance. *International Journal of Operations & Production Management*, 31(8), 888-910.
- Talluri, S., DeCampos, H. A., & Hult, G. T. M. (2013). Supplier rationalization: A sourcing decision model. *Decision Sciences*, 44(1), 57-86.
- Talluri, S., & Narasimhan, R. (2004). A methodology for strategic sourcing. *European journal of operational research*, 154(1), 236-250.
- Talluri, S., & Sarkis, J. (2002). A model for performance monitoring of suppliers. *International Journal of Production Research*, 40(16), 4257-4269.
- Tian, X., & Sarkis, J. (2020). Expanding green supply chain performance measurement through emergy accounting and analysis. *International Journal of Production Economics*, 225, 107576.
- Thunberg, M., & Persson, F. (2014). Using the SCOR model's performance measurements to improve construction logistics. *Production planning & control*, 25(13-14), 1065-1078.
- Tsai, W. C. (2009). Fuzzy measures of supplier evaluation under lean concepts. *Journal of the Operational Research Society*, 60, 1005-1011.
- Ulutas, A., Shukla, N., Kiridena, S., & Gibson, P. (2016). A utility-driven approach to supplier evaluation and selection: empirical validation of an integrated solution framework. *International Journal of Production Research*, 54(5), 1554-1567.

- Van der Sande (2021). *The Rise of the Data-Driven Enterprise*. Accenture. <https://www.accenture.com/nl-en/blogs/insights/data-driven-enterprise>
- Vokurka, R. J., Choobineh, J., & Vadi, L. (1996). A prototype expert system for the evaluation and selection of potential suppliers. *International Journal of Operations & Production Management*, 16(12), 106–127.
- Wagner, S. M., Grosse-Ruyken, P. T., & Erhun, F. (2018). Determinants of sourcing flexibility and its impact on performance. *International Journal of Production Economics*, 205, 329-341.
- Wu, T., & Blackhurst, J. (2009). Supplier evaluation and selection: an augmented DEA approach. *International Journal of Production Research*, 47(16), 4593-4608.
- Ye, Y., Peng, X., Fan, R. L., & Narayanan, A. (2022). An empirical investigation of governance mechanism choices in service outsourcing. *International Journal of Operations & Production Management*, 42(9), 1467-1496.
- Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). Sage.
- Zeydan, M., Çolpan, C., & Çobanoğlu, C. (2011). A combined methodology for supplier selection and performance evaluation. *Expert Systems with Applications*, 38(3), 2741–2751.
- Zou, W., Brax, S. A., Vuori, M., & Rajala, R. (2019). The influences of contract structure, contracting process, and service complexity on supplier performance. *International Journal of Operations & Production Management*, 39(4), 525-549.



## Appendices

### Appendix 1. Interview questions

#### Background:

1. What is your current position in the company and what kind of tasks does your daily work contain? How long have you been working for the organization?

---

To gain understanding how interviewees experienced supplier performance evaluation, it was asked if their work contained supplier performance evaluation. Also, it was used as an introductory question to the topic. Third question is from practical point of view interest of the case company, but also it has been suggested by theory that choices for metrics should come from group of experts (Genovese et al., 2014). Fourth question emerged while interviewing the first interviewee as the power situation and ease of working with suppliers seemed to influence supplier performance measurement system design.

---

#### Current state:

2. Does your job include supplier performance measuring/evaluation? If it does, how do you assess it?
3. What aspects/areas of supplier performance should be considered/evaluated?
4. What kind of suppliers do you have (in terms of size, delivered products etc.) and how well the cooperation works with them?

---

Fifth question was obtained from the study of Hald and Ellegaard (2011). It was pointed out by Purdy and Safayeni (2000) that different sources for performance information exists and therefore, the fifth question contains a question about who is the evaluator. Seventh question is about the updating practices of SPMS, it was included to see what kind of frequency they would like to see from the results and also what is included in the model.

---

#### Measurements and practical application:

5. Should strategic suppliers be measured in other ways than quantitatively? If so, how? (Hald & Ellegaard, 2011).
6. Who should be in charge of evaluating the supplier? (e.g., supplier self-evaluates, certain people, group of people)
7. How often should the results be updated? (Weekly, once a month, quarterly, etc.) How often should the measurements that are included in the model be updated?

---

To understand the interviewees actual use for supplier performance measurement system, their personal needs were tried to assess through the interviews. As it has been suggested by literature, the supplier performance evaluation criteria should be chosen based on the intended use (Lima-Junior & Carpinetti, 2016; Maestrini et al., 2021), therefore it is vital to see, how an interviewee would like to use the system and what kind of measurements they see as relevant. Also, visualisations matter as the outcome from the evaluation can be for example a grade (Bai et al., 2019), a graph (Morgan & Dewhurst, 2007) or a dashboard (Pal & Kumar, 2008). A list is usually used in supplier selection situations where the option would be to compare suppliers (Bai et al., 2019).

---

Personal needs:

8. For your specific job **which criteria** would be most beneficial in supplier evaluation? What would be essential to know about supply performance?
9. What kind of special requirements does your sourcing group have for measuring? (Only for category managers)
10. What kind of supplier performance measurement tool would be beneficial for your job? If you visualize such a tool, what kind of information would you like to see?
11. Would it be beneficial to be able to compare the data to other suppliers? How?

---

As the SPMS full potential related to supplier performance management is reached by communication a question about it is added (Cousins et al., 2008; Maestrini et al., 2018c). As Hald and Ellegaard (2011) identified that employees might take their liberties while communicating the results, a question related to communication plan is added.

---

Use/communication:

12. How would you communicate the results to the supplier? Should there be a company-level communication plan for these results? Should all measurements be communicated to the supplier?
13. Do you have anything else to add?

---

Sustainability issues are nowadays part of companies strategy and therefore, sustainability should be a part of SPMS. As the sustainability issues and related measuring are gaining maturity it is useful to have a question about them. However, it was pointed out by Genovese et al. (2014) that sustainability numbers are rarely comparable. Similar questions were used for safety as well, which is a part of social sustainability.

---

Sustainability (Only for persons responsible for sustainability):

14. How should a supplier be evaluated from the perspective of sustainability?
  - a. What is already being measured and how should measuring be implemented? How far do you consider the measuring of [case company] to be from the ideal, so to speak.
15. What kind of future development prospects are there for sustainability in terms of measuring?
16. What criteria/measurements can be used to monitor/track supplier responsibility?
  - a. How about supplier's development in terms of sustainability? In other words, on which measurements can have goals to be set in terms of sustainability?
17. How reliable is the data related to sustainability?

Safety (Only for persons responsible for safety):

18. How is the supplier evaluated from the HSE point of view and what kind of evaluation should there be / how should they be evaluated?
19. What criteria/measurements can be used to monitor/track supplier safety?
  - a. How about supplier's development in terms of safety? In other words, on which measurements can have goals to be set in terms of safety?

Appendix 2. Quote examples of 1<sup>st</sup> order concepts

Quote	1 <sup>st</sup> order concept
<p>“... filtering the results, a bit, because they are always subjective assessments and background [of the assessment] should also be investigated a little. For example, if someone only gives negative feedback and so on, what are the background reasons there.” [Interviewee 13]</p>	<p>Feedback should be filtered as it is subjective opinion</p>
<p>“... I would at least gently ask ‘how do your on-time delivery numbers to us look like?’ or ‘what is your supply security for [the case company]?’ ...” [Interviewee 10]</p>	<p>It is important that the supplier has a possibility to comment performance evaluation results, for example, if there are errors in the results</p>
<p>“A more neutral measurement would be to get a percentage of the deliveries that have caused a reclamation.” [Interviewee 20]</p>	<p>Measurements should aim for neutrality, e.g., number of reclamations divided by number of deliveries</p>
<p>“Perhaps here, we must also begin by going through the positives, so that we should not start with ‘this is what is amiss with you’. Let’s leave the bad part out and state ‘here is where you should improve’.” [Interviewee 5]</p>	<p>No need to present feedback in a negative way</p>
<p>“... [we previously] collected a subjective evaluation of the suppliers. Of course, there are two sides to a subjective assessment, that if there is such a small group of people who do it, then a person easily sets the grade based on recent experiences. If something has gone wrong with supplier, then they can easily be graded with a really bad grade. And then if there is such a small group of respondents, then the final result can fluctuate.” [Interviewee 20]</p>	<p>Recent problems with suppliers might be highlighted in evaluations</p>
<p>“... also to the suppliers, that it will not be tolerated that things go as they go and then we just live with it, but that if we notice any problem areas, then the problems will be dealt with and they will be brought up to the suppliers in the form of data and in the form of a report” [Interviewee 15]</p>	<p>Measuring should provide evidence to support the performance discussions with supplier</p>
<p>“Thinking from the supplier’s point of view, if I just got some results and did not have the ability to go through them or I was not told what they were based on, it could be confusing ... As in, they would get evaluated, but in reality, only the evaluator would know what they are based on and why this is being done or where the numbers come from?” [Interviewee 2]</p>	<p>Suppliers should be informed about how they are measured and how the measuring is implemented</p>
<p>“... so that we are able to say that we do it this way and we want you to do it like that as well.” [Interviewee 3]</p>	<p>Two-sided measuring could show that the company demands the same things from themselves that they demand from their supplier</p>
<p>“... we have tried a little more to keep it specifically for them to follow their own performance and how it develops, or if it develops at all.” [Interviewee 15]</p>	<p>When the supplier provides the data, they have to assess their own performance</p>
<p>“... they are likely to base their views on what others have told them.” [Interviewee 19]</p>	<p>If evaluator does not work closely with the supplier, they just evaluate what they have heard from elsewhere</p>
<p>“Once a year is, perhaps to some extent, a bad practice, because we are already a bit reactive as we are talking about a period that is already in the past, and we might not necessarily have the regularity to influence it.” [Interviewee 21]</p>	<p>If supplier performance discussions are held once a year, the effect might be more or less reactive</p>
<p>“[the employee thinks:] ‘I will just call that supplier now, leaving no trace of it anywhere’, even though our supplier manager does not know at all that the supplier performed poorly or made a mistake or something else [as it is not marked anywhere]. And the next time they go on to negotiate with the supplier, everything will be fine and there is no problem, even if there really was one. And then of course, it might be that you only take care of those individual cases, even</p>	<p>Properly reported reclamations might influence how suppliers see their performance as well</p>

*though the same supplier has received complaints from several parties.” [Interviewee 14]*

*“Regarding the supplier evaluation, the fact is that if we do it really objectively and not just by feeling, then it requires a fairly large-scale and systematic collection of data” [Interviewee 3]*

*“... that I would limit it, or so that everyone would be measured by those clearly measurable things” [Interviewee 1]*

*“But we are a very important and big customer for many [suppliers], and in that sense, we are an interesting target, that we get answers to offers quite well ... Then we’re in a growth business, so it interests [suppliers] and megatrends support this business, so many [suppliers] want to be involved with us.” [Interviewee 13]*

*“In my opinion, it is worth it to simply separate the performance evaluation into those suppliers that have a direct impact on our process and production, and those that do not.” [Interviewee 9]*

*“... however, there is always a limited amount of capacity, the number of drivers and cars cannot suddenly increase.” [Interviewee 20]*

*“Well, ideally, and in general, we have the same metrics in use that the supplier also uses for self-evaluation and development. We should not have our own metrics that we do not share with the supplier. Rather, these are our common metrics, where we define what is the target level in our operational life, and then we both work towards this common goal.” [Interviewee 16]*

*“I’ve learned over the years that the customer, who shouts the loudest, gets the most things done. Of course, the number of purchases has an effect in the background, what our spending is or how much we buy, how important we are. But if we are roughly as big a customer as, for example other buyers, then if we are completely silent and the other company nevertheless points out or reports things to the supplier, then [the supplier] does those improvements internally. The one who speaks gets something in return.” [Interviewee 9]*

*“We should be able to define what we want, so that they can be measured.” [Interviewee 12]*

*“... there was always some, let me say contrary targets. You know, we choose suppliers delivering from really deep in the South of Europe and South of Italy because they are cheaper, and they are innovative, but at the end we have this long transportation from the southern Europe to the middle or to the north of the Europe. And that has an effect on the environmental rating for the supplier.” [Interviewee 11]*

*“[performance goals] can be fixed in frame agreements with a supplier. It is much easier to bring it in frame agreements, because then it is signed by the senior management of the supplier. Then it is agreed on a high level.” [Interviewee 11]*

*“... the suppliers usually want to know how we view them, and then they typically go back and try to fix those things [that we notify them about]” [Interviewee 6]*

*“It could be good to have frequent follow-up meetings, where you also present what has been working well, what can be improved, how do we get more of these win-win situations? How can we help each other to grow and so on.” [Interviewee 7]*

*“A follow up is always required. I don’t know if once per quarter, once per half a year, or so that there is an exchange of information. What were the achievements*

To reach objectivity in supplier evaluations, extensive and systematic data collection is required

All suppliers should be measured based on the metrics available from the systems

The case company's size benefits the sourcing function through the willingness of suppliers to cooperate

The suppliers that offer critical services or products should be identified, and they should be evaluated differently

There is a limited capacity available in operation region

For measuring, there should be a consensus between the buyer and the supplier about important metrics

If the supplier is not informed about the case company’s desires, it is obvious that they cannot improve as buyer wants

It should be well defined what we want from the supplier to be able to make measurements for it

Some measurements are contradictory

Supplier performance measuring should be agreed with the supplier in the frame agreements to ensure support from the senior management

Comparing can indicate to the supplier what they have to improve to ‘compete’

Performance measuring and communication should revolve around what works and what does not, identifying improvement areas

Regular meetings with suppliers would ensure the exchange of information

*from the suppliers, concerning innovation or concerning sustainability, cost reduction.” [Interviewee 11]*

*“What happens often is that we start collecting all kinds of data with great enthusiasm and sending it to the supplier, but then the kind of genuine improvement actually comes through discussion” [Interviewee 1]*

*“It’s good to have instructions that then we act as a unified company and not different depending on who does what.” [Interviewee 16]*

*“... of course, we have these governance models and so on, so maybe it could be linked there.” [Interviewee 8]*

*“And it develops it towards systematization, so that everyone does not have to re-invent the wheel and do things in their own way and carry those operating methods with them [from previous workplaces] all the time. And this way, it is done in a standardized way and more time is left for the essentials, when learning the common work practices.” [Interviewee 3]*

*“... they are also the two first indications that will let you know if a supplier starts struggle, if the quality or reliability of the delivery start to suffer ...” [Interviewee 3]*

*“That a supplier would be able to bring their own views forward, telling us that they have this kind of a thing, and they would be ready to develop these types of things. Those are the things that, in my opinion, should also be measured to determine how capable a supplier is. Not only in that, what they normally do from Monday to Friday, but also in what additional value they can bring to our process” [Interviewee 10]*

*“... but there might be quite rigid working styles. That if you say that ‘this matter could be developed a little and you could tweak it a little’, [Supplier] might not necessarily be so good at it. On the other hand, there, in another field of suppliers, might be even a little more desire for strategic cooperation and development.” [Interviewee 3]*

*“... those business units are like the number one priority, where feedback needs to be asked from, so that in a way, you have to get them to commit and then collect the feedback that tells you which suppliers work, and which do not.” [Interviewee 13]*

*“If you think, for example, that the responsibilities suddenly changed and someone else came over to do my work and I wasn’t around to share my own experiences, that new person would then, in a way, have to learn the same things all over again.” [Interviewee 4]*

*“In an ideal situation, [powerful reclamations] could trigger an alert for these other factories as well, where the same supplier operates. ‘Is it possible for something like this to happen to us as well?’” [Interviewee 21]*

*“[In a previously conducted survey], it was clear that the same supplier could receive varying feedback. For example, one BU commented that this is an absolutely brilliant supplier, and another said that this is an absolutely terrible supplier.” [Interviewee 14]*

*“Communicating within [case company] to help others, because you learn a lot from these cases. When everything works fine, you know everybody’s happy of course. But when you have these problem cases, then it’s when you are really learning things.” [Interviewee 7]*

*“Is there anything that we could perhaps do better. For example, is there something to improve at the level of communication or otherwise, so that we could*

The results themselves should not be the most important thing in the measuring, but the areas where buyer or supplier could improve themselves

A company level communication plan would be beneficial, enabling the company to appear the same during all the discussions

Communication could be linked to governance models

There is a risk that performance measuring results are not communicated if there is no communication plan

Measuring should capture indications about the supplier’s future performance and how they impact the buyer

Suppliers’ innovativeness, improvement propositions and product suggestions could also be measured

Suppliers which are willing to help and improve should be recognized

An opportunity to hear performance feedback from internal customers and business units

Having historical data about the supplier helps to integrate new employees

Impactful reclamations could trigger a warning to other factories that use the same supplier to look if they have same vulnerability

Supplier’s performance can vary across different business units

Communication about these problem cases facilitates learning inside the case company

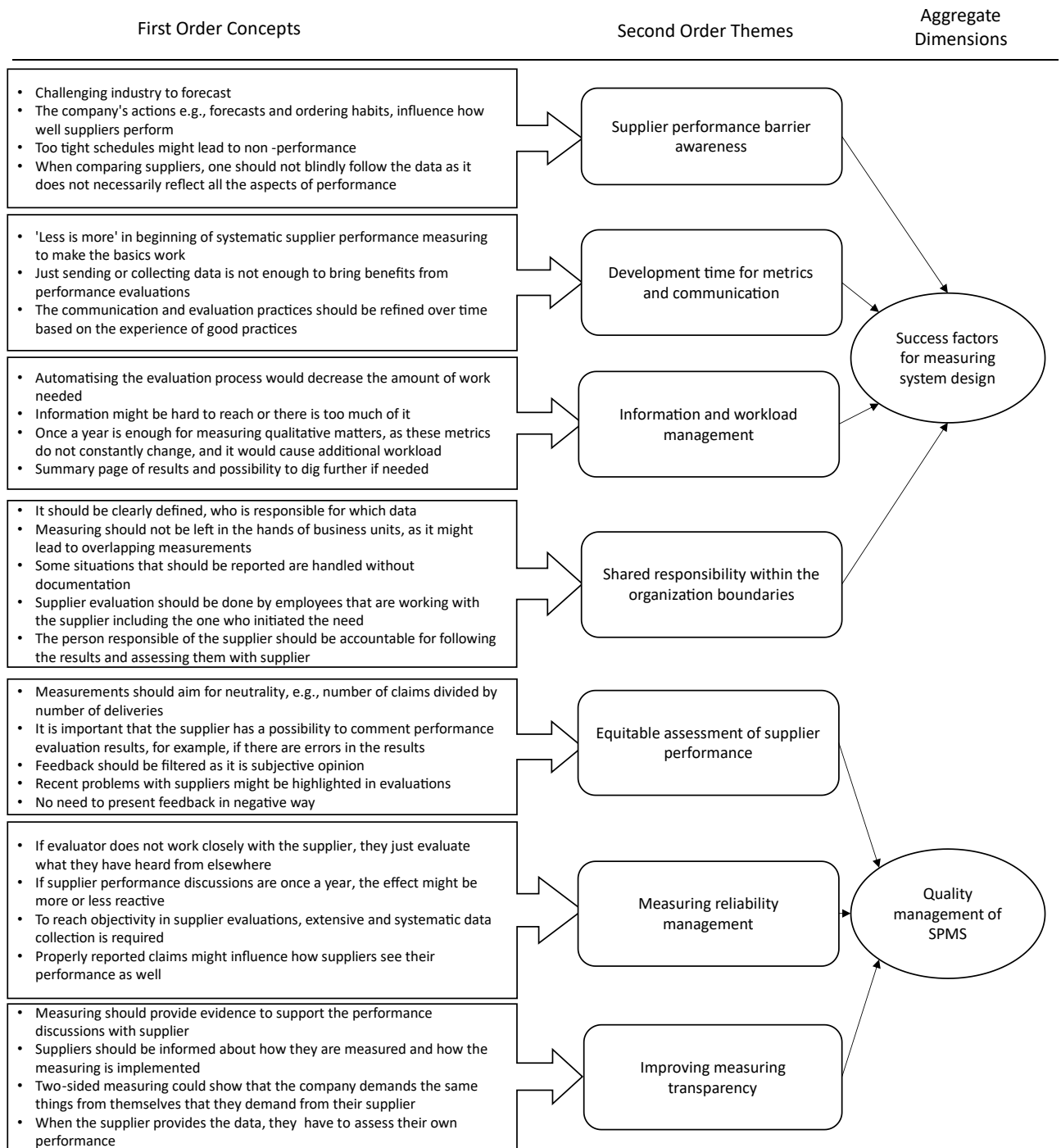
Comparing can be used to identify which suppliers are the best or the

- reach the same results as these other suppliers that are performing better” [Interviewee 15]*
- worst, and to try to identify why that is the case
- “... when the pieces of information, for example from complaints or something like that, so that I can somehow get them out of from the system. Because if we think about making a contract, then of course you could prepare for it in a completely different way and prepare for the contract, e.g. tackle the problems that there have been ...” [Interviewee 8]*
- Performance information could be used to develop better contracts
- “However, as we may not be full professionals in the supplier’s field, then of course we want to trust the supplier and cooperate to build that practice in such a way that it also works in practice, so that we do not just make instructions that exist on paper but are not actually followed” [Interviewee 21]*
- The buyer can learn more about how they can improve their own processes as well when discussing with suppliers
- “Somehow evaluate what was good with this delivery, what was bad, what could be improved for the next time?” [Interviewee 7]*
- What could be improved next time when working with supplier
- ” When the volume changes come in a short cycle, the time span for arranging more capacity with the current suppliers or obtaining additional capacity from new suppliers becomes a bit too short. So perhaps partly through our actions and partly through the nature of our operations, we contribute to these challenges.” [Interviewee 20]*
- Challenging industry to forecast
- “How we plan that. That is, how we plan, for instance, our own production and deliveries and such. If our schedule planning is too strict, the supplier might not be able to perform accordingly” [Interviewee 16]*
- The company’s actions e.g., forecasts and ordering habits, influence how well suppliers perform
- “Is the order information from [the case company] on time? If we have agreed, for example in the contract, that these things are ordered with two weeks’ notice, will the order be placed before or will it slip to ten days or something like that, so that the suppliers always have to act in a hurry” [Interviewee 3]*
- Too tight schedules might lead to non-performance
- “A car or truck is registered as having arrived when it enters through the gate to the factory area, but what greatly affects the comparability of arrival data is the presence of a lunch restaurant in the port area, which means that very often, the driver might take a break there in the port area and have lunch. And then, of course, the departure is registered when they drive out of the factory area. So, if we, for instance, compare that to the port of [a city], where when a vehicle arrives, it unloads and departs. And if, without knowing these facts, we would start comparing these two locations like: ‘[previously mentioned city’s port] has totally miserable vehicle cycle times, so please do something’. [angrily] ... But on the other hand, as we know these things, we will not go on to make the mistake of directly comparing suppliers with each other using some of these particular indicators [that are not comparable].” [Interviewee 17]*
- When comparing suppliers, one should not blindly follow the data as it does not necessarily reflect all the aspects of performance
- See concept:* The results themselves should not be the most important thing in the measuring, but the areas where buyer or supplier could improve themselves
- Just sending or collecting data is not enough to bring benefits from performance evaluations
- “That is, people in different areas have this and that many things that they would like to know and measure. ‘It would be nice to know this and that. And these and those.’ ... Then it forms, what we also have and can be seen, such a huge amount and a huge jungle of different information, different metrics and different reports. And perhaps with that, what is essential and where we really want to put our effort into becomes lost to us. That less is more. ... if it helps us to move forward, we should prefer simpler, even if they are as simple as possible, they often help more than very advanced, theorized ones. Even in basic matters, there is usually quite a lot to do.” [Interviewee 18]*
- ‘Less is more’ in beginning of systematic supplier performance measuring to make the basics work
- “I think that the [communication] practices form over time based on conversations. When it is identified what produces added value and what does not.” [Interviewee 21]*
- The communication and evaluation practices should be refined over time based on the experience of good practices

<p><i>"... if there was an online tool, which would be sent to you kind of automatically, and then you would fill it out and it would provide a summary, and then that summary would go on to the supplier automatically, then it would reduce work by an insane amount ..."</i> [Interviewee 1]</p>	<p>Automatising the evaluation process would decrease the amount of work needed</p>
<p><i>"It would definitely have to be fast and simple, so that if I had a need to focus on a supplier in more detail, I could see it the necessary information at a quick glance. ... But about the first screen, my idea would be that when you open and search for the supplier, you would quickly get a general idea of the situation. So there should not be too much data there, only the deliberate main points that are chosen. As it often feels like we have everything that we can think of, put on display."</i> [Interviewee 9]</p>	<p>Information might be hard to reach or there is too much of it</p>
<p><i>"As I would see evaluation would take place every year as it is quite a massive operation to evaluate a supplier and go through that process. ... If something were to happen in their operation in the time between, we would get alerted in any case elsewhere."</i> [Interviewee 10]</p>	<p>Once a year is enough for measuring qualitative matters, as these metrics do not constantly change, and it would cause additional workload</p>
<p>See concept: Information might be hard to reach or there is too much of it</p>	<p>Summary page of results and possibility to dig further if needed</p>
<p><i>"Then there's the operational procurement, but also the technical organization in that they record the reclamations, etc. The fact is that everyone does their own part, and handles their plots and entries into the systems ... But if that [only one] person was responsible for ensuring that all the reports are functioning, that the numbers are correct, that the data is imported correctly, etc., then no, it is too large of a responsibility to bear. It is not really possible."</i> [Interviewee 3]</p>	<p>It should be clearly defined, who is responsible for which data</p>
<p><i>"Well, it would seem that you would have to do it from the procurement, as the same suppliers serve several different business units, so BU's should not measure independently. If BU's measured independently, it would mean at the same time, [BU 1] would measure [Supplier X] and soon [BU 2] would measure [Supplier X] again, so yes, it would have to be an organization level [performance] report [of the supplier]."</i> [Interviewee 20]</p>	<p>Measuring should not be left in the hands of business units, as it might lead to overlapping measurements</p>
<p><i>"... it is clearly seen that all quality deviations do not come through the system, that they are handled directly [with the supplier] without recording them into the system."</i> [Interviewee 14]</p>	<p>Some situations that should be reported are handled without documentation</p>
<p><i>"In my opinion, I would include everyone who works with the supplier."</i> [Interviewee 1]</p>	<p>Supplier evaluation should be done by employees that are working with the supplier including the one who initiated the need</p>
<p><i>"then the person responsible of the supplier, who collects this information and then is the one who channels it [to the supplier]"</i> [Interviewee 14]</p>	<p>The person responsible of the supplier should be accountable for following the results and assessing them with supplier</p>

---

### Appendix 3. Data structure of managing internal process and design





## Appendix 4. Supplier performance measurement system role and management – Data structure of SPMS as an enabler of supplier performance management

