

Unleashing the Power of Supplier Performance Metrics

- Easier said than done



Abstract

The purpose of this thesis is to investigate how companies evaluate their suppliers' performance when purchasing goods. We look into which metrics are used and how they are defined, as well as into the whole process for gathering, storing and working with supplier metrics data. This topic is of interest since the purchasing function of companies has become increasingly important. The thesis highlights and analyzes some difficulties that arise when measuring suppliers: both internally in the purchasing organization and externally with regards to the suppliers. The thesis also includes some best practices.

Data has been gathered from interviews with 25 persons working in purchasing and quality management. In total 13 companies from 4 different countries were visited, the majority being industrial companies. The theoretical foundation is comprised of theories about measurement in general, performance measurement in organizations, and measurement within supply management. An interesting find from the theory studies is that there is a lack of theory explaining in detail *how* to measure. This theory gap correlates with the difficulties encountered by companies in their practical work.

The main findings of this work are that there are large differences among companies in their purchasing maturity. Further, suitable IT tools are almost a necessity for efficient usage of metrics. Using a large amount of metrics without the right IT support seems to be a very difficult undertaking. It is also vital that metrics are understood and used by the entire organization, especially since different departments can have different interests. This goes especially for subjective metrics, which are more difficult to use than objective metrics. A final proposition is that the complexity of the field of supplier performance metrics is larger than what is hinted by most academic literature.

Stockholm School of Economics
Master Thesis
Department of Marketing and Strategy
Authors: Lisa Levinson and Robin Sunström
Tutor: Björn Axelsson

Thank you

Claes-Erik & Håkan
8th, 9th & 10th floor
Respondents
Björn Axelsson

...for your support and knowledge

Table of Contents

Introduction	6
Background	6
Purpose	6
Contribution	7
Delimitations	7
Theory	8
Measurement in General	8
Basic types of measurements	9
Measuring in Purchasing	9
Why measure supplier performance?	10
Four levels of supplier performance measurement	11
Four types of supplier performance measurement	11
Three levels of analysis for supplier performance measurement systems	12
What metrics are used?	12
Measuring and IT	13
Why suppliers are not measured	14
Design of supplier performance measurement systems	14
The complexity of how	15
Method	16
The Study	16
Scope of the study	16
A qualitative study	17
How the Data was Collected	17
The interviews	17
Other types of data gathering	17
Documenting the interviews	17
Validity and Reliability	18
All thirteen companies are treated as independent companies	18
Data gathering and reliability	18
Selection of companies and validity	18
Empirical Findings	20
COMPANY A	20
COMPANY B	21
COMPANY C	23
COMPANY D	24
COMPANY E	26
COMPANY F	27
COMPANY G	28
COMPANY H	29
COMPANY I	30
COMPANY J	32
COMPANY K	33
COMPANY L	35
COMPANY M	36
Matrix for Comparison	38
Analysis	40
Large Differences in Purchasing Maturity	40
Which Suppliers are Evaluated	40
Direct suppliers	40
Indirect suppliers	40
Segmentation	41
How often Evaluation Takes Place	41

Many Metrics do not Equal Good Performance	41
Subjective Metrics	42
Advantages of subjective metrics	42
Disadvantages of subjective metrics	42
Difficulties with Defining the Metrics	43
PPM	43
Delivery precision	44
Different goals inside the company	44
Difficulties with Internal Compliance	45
The Value of IT for Supplier Performance Measurement Systems	46
Technology use in company Alfa	46
Technology use in company Beta	47
Implications for the Host Organization	47
Implications for Practitioners	47
Make clear definitions of metrics	47
Do careful follow-up of metrics, internally	47
Do careful follow-up of metrics, in regards to suppliers	48
Strive for simplicity in a system	48
Use technology to improve measurement	48
Implications for Future Research	48
References	50
Appendix	52
The Question Template	52

Word List

Here we define some terms used in the thesis:

Supplier performance: how a supplier meets a buyer's running business demands. Supplier performance is normally assessed through different metrics.

SPM system: Supplier Performance Measurement system, a company's central process for gathering, storing and using data about supplier performance.

Direct/indirect material: direct material is goods purchased by an organization that goes directly into production. In a furniture factory an example would be wood. Indirect material is purchased goods not used in the production. For the furniture factory two examples would be insurance and printers to the office. Indirect material is seldom as important for a company as direct material is.

Metrics: what a company measures a supplier on over time. In many companies they are also known as key performance indicators (KPIs). PPM is one common metric.

PPM: Parts Per Million. A common metric for measuring the quality of goods delivered by suppliers. $PPM = [\text{(the number of faulty parts received)} / \text{(the quantity of delivered parts)}] * 1\,000\,000$

ERP: Enterprise Resource Planning systems integrate all data and processes of an organization into a unified system. A typical ERP system will use multiple components of computer software and hardware to achieve the integration. www.wikipedia.org

BS6001: a standard for sampling goods, mainly used in the UK. Established by the National Standards Body of the UK, www.bsi-global.com.

ISO 9000: a family of standards for quality management systems, often used in manufacturing companies. The ISO standards are maintained by the International Organization for Standardization, www.iso.org.

ISO TS: a quality management standard common in the automotive industry. It is based on, and is more advanced than, the ISO 9000 standard.

Purchasing Maturity: is a construct that illustrates the level of purchasing professionalism at the business unit level. Purchasing maturity derives from purchasing stage models.¹

¹ Rozemeijer, Van Weele & Weggeman, 2003

Introduction

Background

The Supply Management function in large companies has been receiving an increasing amount of attention the last years. One explanation is that several trends affect supply management, for example globalization, rising competition, increasing focus on customer orientation, and improvement of information technology tools such as databases and electronic communication.²

Supply chain management has been acknowledged as the new source for competitive advantage, raising the interest of CEOs. At the beginning of this decade, purchased goods and services typically represented from 50 to 70 percent of turnover.³ In some industries it could even amount to as much as 85 percent.⁴

Purchasing was for long a routine function where orders were made in a straightforward fashion. Now it is a strategic issue. In many cases relationships between buyers and sellers have shifted from being arm's length relationships to being a form of close collaboration.⁵ Specialization among companies is increasing. The modern industrial company has become more dependent on its environment and part of a larger network.⁶ This has lead many of them to move their production, outsource or just purchase from suppliers in geographically distant locations. This leads to an increase in the amount of purchased goods, as a ratio of the turnover.⁷ Another interrelated development is an increased competence level and professionalism within the Supply Management function.⁸

In conclusion, it is of great importance for companies to have a well run Supply Management function.

The main concern for Purchasing is to make sure that one buys from the best suppliers and to improve the current suppliers. This brings us to the concept of supplier performance measurement, an integral part of the purchasing function. It has been said that, *"the design of supplier performance measurement systems is possibly one of the final structural areas of business in which significant savings can be made. It is becoming an increasingly important strategic tool as trade becomes global."*⁹

One of the companies affected by these developments is a global Swedish manufacturing corporation (which will remain anonymous) that has commissioned this thesis. The corporation consists of several recently acquired companies with thousands of employees, has a presence in over 40 countries, and is one of the largest actors on its market. Our assignment was to analyze its current situation and propose a way for them to work with supplier performance metrics. This resulted in a status report with recommendations, which we presented for the corporation, as well as this academic thesis.

Purpose

Initially, our purpose was to find practical answers to the question of how the different companies within our host organization and the other benchmarking companies work with supplier performance metrics. The following questions were therefore studied:

² Shepherd & Günter, 2006

³ Spekman, Kamauff & Spear, 1999

⁴ Van Weele, 2005

⁵ Schary & Skjott-Larsen, 2007

⁶ Axelsson (edited by Samuelsson), 2004

⁷ Axelsson, 2006

⁸ Ibid

⁹ Morgan, 2004

1. *What* are the primary metrics used by companies? How are they defined? For which suppliers are they used?
2. *How* does the process of gathering, storing and working with metrics data work, in practice?

However, while studying the purchasing processes in companies we were astounded by the fact that it was so difficult to work with supplier performance measurement. It was nowhere as straight-forward and clear-cut as it is mainly described within normative supplier performance measurement literature. Once a decision is made to measure supplier performance, literally all functions are affected (including R&D, IT and accounting).

This brought us to explore an additional issue:

3. Explore root causes behind the complexity in the seemingly simple act of measuring suppliers' performance.

The aim is to gain an understanding of supplier evaluation through how metrics works *in practice* when evaluating ongoing supplier performance. The aim is not to dig deep into conceptual or strategic issues. From these insights we give some advice to practitioners.

Contribution

This thesis gives an insight into how different companies work with supplier performance metrics. It is a practical status report of the current state in the field. This is valuable since most academic studies in the field of measurement systems focus on what to measure and not on how the measurement actually should be done.¹⁰ Moreover, this thesis highlights specific difficulties in the practice of measuring supplier performance. The contribution is an increased understanding of the complexity in employing supplier performance measurement, as well as practical advice both for practitioners within Supply Management and for general managers.

Delimitations

There are several levels of measurement.¹¹ Our unit of measurement in focus is the suppliers' performance. Supplier performance can be evaluated in many ways. We have chosen to separate supplier **assessment** from supplier **evaluation**. In this thesis we only study the evaluation part. Supplier assessment is about estimating a supplier's probable performance before the buying commences, for example by visiting the supplier's production plant to see how advanced it is. Or it can concern CSR issues. Supplier evaluation on the other hand takes place after the deliveries have started and evaluates the actual performance.

Moreover, we do not study *price*. Price is the most obvious and common ways to both assess and evaluate suppliers. However, price is a special component. It is established in contracts and is thereby under the control of the purchasing organization. The focal point is here on the occurrences in the supplier's deliveries, after a contractual agreement has been made.

Furthermore, all studied companies purchase services. Measuring of service suppliers' performance differs greatly from measuring performance of direct material suppliers.¹² Therefore service suppliers are not included in the scope of this thesis. Finally, the thesis is focused on companies that provide a product with a physical component, so we did not study service-oriented buyers.

¹⁰ Paranjape, Rossiter & Pantano, 2006

¹¹ Axelsson, Laage-Hellman & Nilsson, 1991

¹² Axelsson & Wynstra, 2002

Theory

Within the next section we will present the theoretical foundation for supplier performance measurement. Firstly, we will introduce the concept of measurement in general and its four different levels. Secondly, the theoretical scope will be narrowed down to purchasing and more specifically; suppliers' performance. The existing theories within this niche will be presented, as well as an area wherein there is inadequate research coverage, which is how to measure supplier performance.

Measurement in General

"What gets measured gets done"

The above quote summarizes why things are measured. Measuring is an instrument that reduces uncertainty. It is also used to ensure a desired outcome. Measuring performance is something all companies do.¹³

Measuring is a control activity and by measuring one can control the outcome of past events. Through measuring it is possible to assess the value of different activities and efforts. Performance measurement determines whether the increased value of an activity exceeds the cost of it. The information that is generated from measuring performance can support and improve decision making. It can also be used to map where the starting point of something is and how it could be improved. With this common picture of the situation it is easier to communicate. The visibility of the measured activities is increased which makes it easier to notice early indicators of changes.¹⁴

Further, performance measurements can work as a tool to influence future behavior. By connecting targets to certain measurements, the motivation to improve is increased.¹⁵ To further reinforce motivation the targets could be linked to incentives (or punishment).¹⁶ By linking it to consequences, a measurement becomes a mean to influence behavior.

The extra attention a measured activity receives hopefully increases the efficiency and/or the effectiveness of that action. This in turn is supposed to contribute to the welfare of an organization. However, it is not certain that measurement brings advantages: that depends on whether the right activity is measured in the right way. Measuring could be counterproductive. This could happen either if energy is consumed at measuring unimportant things, or if measurements are given so much emphasis that people focus only on the measurements and not on the entire situation, thereby neglecting important things that are not measured. Moreover, performance measurement is a way to control the action of people, and to be controlled is not always desired. In conclusion, the choice to measure activities within an organization requires change.

Performance measurement can be defined as the process of quantifying the efficiency and effectiveness of an action. A *performance measure* can be defined as a metric used to quantify the efficiency and/or effectiveness of an action.

A *performance measurement system* can be defined as the set of metrics used to quantify both the efficiency and effectiveness of actions.¹⁶

¹³ Parker, 2000

¹⁴ Axelsson, 2006.

¹⁵ Ibid

¹⁶ Barnett, 2006

Basic types of measurements

The activity of measuring requires a standardized scale. The definition of physical science measurement demands a strict ratio scale: *"the determination or estimation of ratios of quantities [where] quantitative attributes are those which it is possible to measure, at least in principle"*. The scale must therefore have a constant zero value and uniform units. Whereas, there are more liberal definitions that allow measuring to take place without numerical quantities where *"numbers need only be assigned according to a rule."*¹⁷ With this liberal definition it is possible to measure almost anything including individuals' perception of various events. This implies that even subjective opinions can be measured.

The four different levels of measurement are:

- (1) *The ratio level*, which is characterized by having an absolute zero. For example age, mass and number of participants.
- (2) *The interval measurement*, where there are equal differences between each entity, but no absolute zero. Examples of interval scales are the temperature and calendar dates.
- (3) *The ordinal measurement*, which ranks entities without regard to the differences between each entity. This measurement is used in racing where each entity is ranked as 1st, 2nd, 3rd etc.
- (4) *The nominal measurement*, the fourth level, is the most liberal definition of measurement. Numbers label a mode, and the numerical value of the number is irrelevant. Examples of measurable entities are for example gender, home town and marital status.¹⁸

Measuring in Purchasing

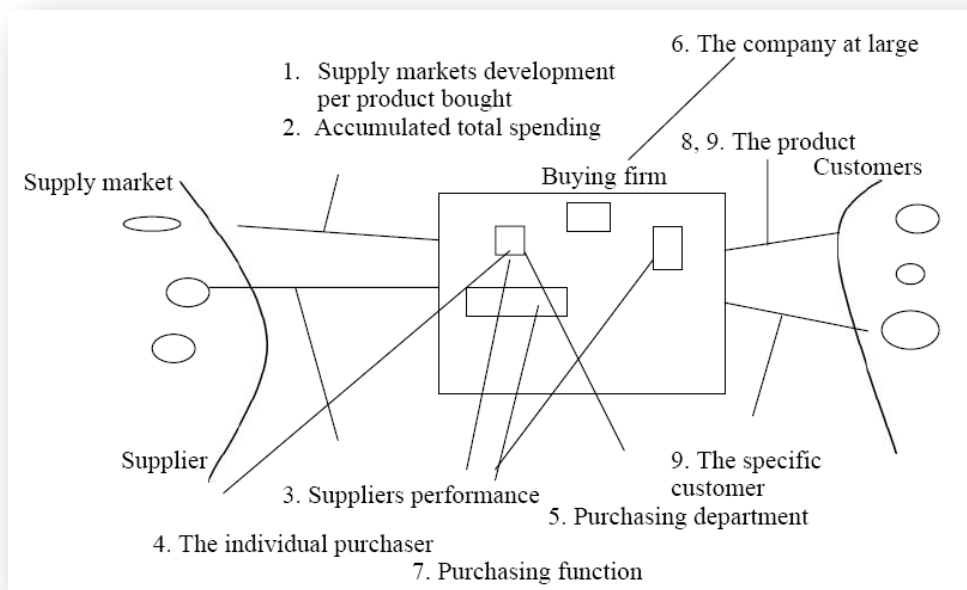


Figure 1: Areas subject to measuring in purchasing ¹⁹

In Supply Management there are several subjects to measure, the above figure gives a view of the entire field. It is possible to measure the *company's* total spend, or the efficiency of each *purchaser*, or perhaps the size of the *supply base*. In this thesis the focus is on the individual *supplier* and its performance.

¹⁷ Stevens, 1946

¹⁸ Ibid

¹⁹ Axelsson, Laage-Hellman & Nilsson, 2002

Why measure supplier performance?

"Garbage in = garbage out."

Suppliers influence companies' quality performance more directly than is often acknowledged.²⁰ Because of globalization and an increased proportion of purchased goods, more is purchased from geographically distant places. It is not unusual that a month's supply of components is out on the sea, just leaving China, when the production is forced to shut down because of faulty products. This is a monetary cost, and it is up to Purchasing to measure and regulate the damage, and to ensure that this does not happen again. If the companies are not measuring their suppliers' performance they cannot demand compensation for shortcomings. It is hard to evaluate, and penalize or reward, suppliers' performance if it is not measured.

A study from the research company Aberdeen Group²¹ shows that companies that have a supplier performance measurement program (SPM) improve their suppliers' performance markedly, regarding price competitiveness, on-time delivery, quality and service. The average improvement of supplier performance is over 20 percent according to this study.²² These are great differences; the quality improvement with a SPM outperforms those without SPM with more than 15 percentage points. One reflection that arises though is that companies that have undergone the process of implementing a SPM might be prone to rationalize that decision and exaggerate the supplier improvement.

	On-time delivery	Quality	Service
Supplier performance program	23 %	21 %	21 %
No supplier performance program	11 %	5 %	17%

Figure 2: Average supplier performance improvement ²³

The above table indicates that a supplier performance measurement system does improve supplier performance. According to the same study, this is exactly what companies mainly (see figure below) use the collected supplier performance data for, which implies that a supplier performance measurement system delivers to expectations. Companies also use supplier data to improve supplier selection, to ensure that they collaborate with the best performing supplier. This is closely related to the use of supplier data to consolidate the supply base where data is used to discard underperformers. Supplier performance data can also point out risk elements, since low performance can be an indicator of bad management, old machinery or low competence. Moreover, companies make use of the data to ease prioritization in supplier development. With supplier performance data it is easier to find the suppliers that are most in need of support. Along the lines with the saying *"the best predictor of future behavior is relevant past behavior"*²⁴ supplier performance data is used to predict supplier performance. The last mentioned use of supplier performance data is to segment the supply base. This could imply that suppliers receive different treatments based on performance.

²⁰ Chen & Parajuna, 2004

²¹ The study is based on an online survey of 197 mainly North American enterprises in 2005

²² Aberdeen Group, 2005

²³ Ibid

²⁴ University of Illinois at Urbana-Champaign, 2007

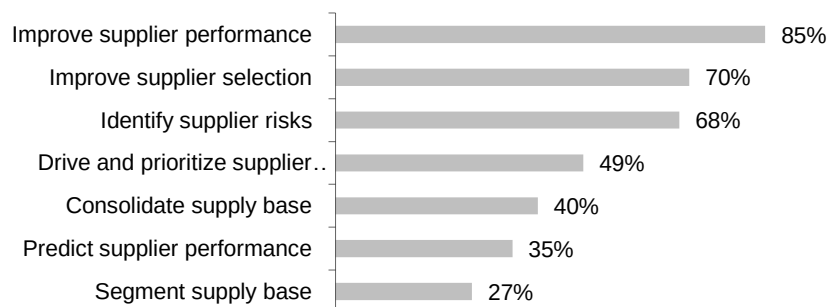


Figure 3: Use of supplier performance data ²⁵

Four levels of supplier performance measurement

Supplier performance can be measured at several levels. The level that the supplier is being evaluated on influences *what* is measured and *how* it is measured. One framework for how companies evaluate suppliers is the four-level framework proposed by Arjan van Weele:²⁶

- (1) *Product level* is when the incoming deliveries of goods from suppliers are measured and controlled against specifications.
- (2) *Process level* focuses on the supplier's production process. The supplier's ability to generate value is measured.
- (3) *Quality Assurance level* is when the supplier's Quality Management system is measured to ensure that the supplier has an ongoing process to improve quality.
- (4) *Corporate level* is when the supplier's financial status and management is measured. This is to estimate whether a supplier has the ability to continue its operations.

Usually only the two first methods, which are very close to production, are employed.²⁷ It is also more common to attribute the *product level* measurement to *supplier evaluation* whilst the three other levels are more closely associated to *supplier assessment* or less frequent, e.g. annual, *supplier evaluation*.²⁸

Four types of supplier performance measurement

When measuring supplier performance, the measurement can be classified into four different types depending on the measurement's focal point:

- (1) *Outcome measures* look at what is produced. Measuring a finished output indicates the quality of the performance. However it says nothing about the preceding events.
- (2) *Action measures*, in contrast, measure the events that create the finished product. This is a leading indicator that measures the source of performance.
- (3) *Input measures* are lagging indicators that are most useful when they are compared to the outcome. Thereby it is possible to compute for example the resource efficiency of a certain production process.
- (4) *Diagnostic measures* measure the measures. In order to understand what causes the results of outcome and action measuring, the diagnostic measure is used.²⁹

These four types of measurement relate to the above four levels of measurements where e.g. *product level* measurement often is an *outcome measure*. When looking at *supplier evaluation* the most often occurring type of

²⁵ Aberdeen Group, 2005

²⁶ Van Weele, 2005

²⁷ Ibid

²⁸ see clarification of *supplier evaluation* and *supplier assessment* on page 7 - delimitation

²⁹ Parker, 2000

measurement is the *outcome measure* since supplier performance evaluation is often performed at the purchaser's site where the other type of measures are harder to perform. The same line of thought is valid for *product level* measurement.

Three levels of analysis for supplier performance measurement systems

Supplier performance measurement systems can be analyzed at three levels:

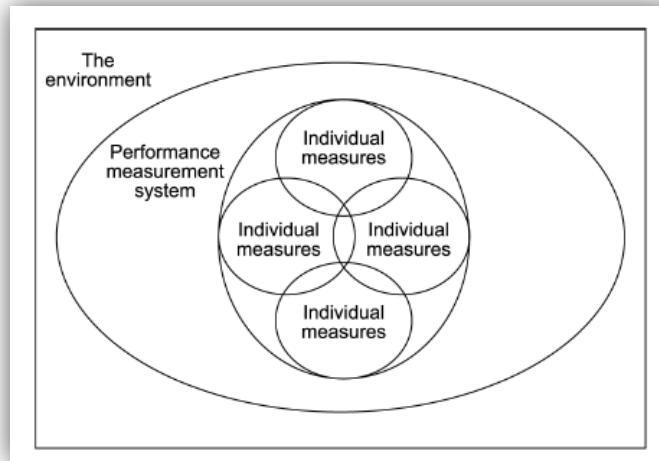


Figure 4: A framework for performance measurement system design ³⁰

- (1) *The individual metrics* where it is possible to analyze why they are used, what benefits they generate and at what cost.
- (2) *The set of measures* shows if there is consistency between the separate metrics. Here it is possible to analyze if any metrics contradict each other, whether the focus is long or short term, if the metrics are thoroughly integrated in the company, and if improvements are measured.
- (3) *The performance measurement system* as an entity in a context: at this level it is possible to see if the performance measurement is attuned to the corporate strategy.³¹

What metrics are used?

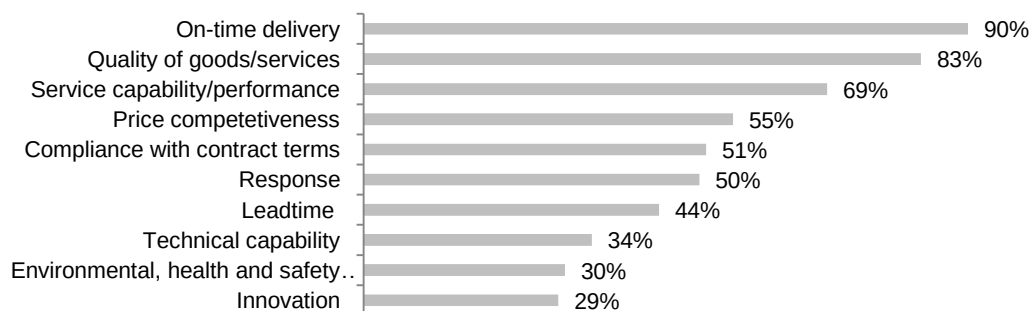


Figure 5: What supplier performance metrics do companies use? ³²

There is a massive fauna of supplier performance metrics. According to the Aberdeen study, the most commonly used metric is delivery performance, which compares the factual delivery time with the requested, to see if the delivery has arrived on time. The second most used metric is quality performance. This metric establishes whether the supplier has met the demands stated in the order specification. The third most common measurement is service performance, which can be anything from a

³⁰ Neely, Gregory & Platts, 2005

³¹ Ibid

³² Aberdeen Group, 2005

purchaser's perception of the supplier to average service response time. These three metrics are considered to be traditional and relevant *"to support the less sophisticated way of operating..."*³³ The fourth measure most used, according to that study, is price competitiveness, which however is outside of the scope of this thesis.

All these measures can be defined in different ways. It is possible to measure supplier performance with different scales, levels and types. E.g. quality performance can be measured with a ratio scale, on a product level with an outcome measure. In practice this would imply that each incoming product is examined and where every single flaw, regardless of severity, is recorded as a faulty unit. Quality performance can also be measured on a quality assurance level, as an action measure and with a nominal scale. Here the purchaser *estimates* how well the supplier handles quality matters.

The fifth most commonly used metric is compliance with contract terms, which is a rather unspecific measure that could contain just about anything, such as quality performance, as long as it has been stipulated in the contract. The sixth most used metric is also very broad, response time, and could relate to just about any communication or coordination matter. Seventh, lead-time metrics measure how long it takes for a supplier to deliver a product once an order has been made. This is of great importance for companies with a lean production philosophy. Lead-time is often contractually specified and companies measure how well the supplier comply with the stipulated lead-time.

Three other commonly used supplier performance metrics are technical capability; environmental, health and safety compliance; and innovation. The definitions of these are plentiful since the overarching categories are so broad and unspecific. These types of metrics are often on a *process-, quality assurance- or corporate level*³⁴, which is more likely to be a part of *supplier assessment* which is outside of this thesis scope.

In general, Beamon states that *"qualitative evaluations such as "good", "fair", "adequate", and "poor" are vague and difficult to utilize in any meaningful way. As a result, quantitative performance measures are often preferred to such qualitative evaluations."*³⁵

Measuring and IT

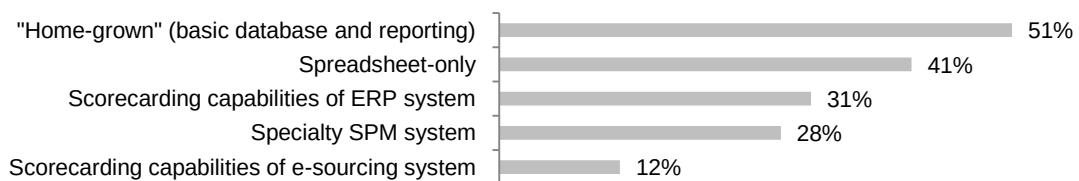


Figure 6: Types of supplier performance measurement technologies used³⁶

IT has revolutionized performance measurement.³⁷ The tradeoff between reach and richness of information has changed to the better. Every second type of technology used for measuring supplier performance is, according to the study by Aberdeen, self-developed databases. The quality and functionalities of self-developed systems are very diverse. They can be customized to the idiosyncratic needs of a company or they can be a mediocre consequence of adapting to old legacy systems. The quality of the second most common technology, Excel spreadsheets, is in contrast not very diverse. It is instead most certainly limited and old fashion. The third most employed type of

³³ Axelsson, Rozemeijer & Wynstra, 2005

³⁴ Van Weele, 2005

³⁵ Beamon, 1999

³⁶ Aberdeen Group, 2005

³⁷ Axelsson, Rozemeijer & Wynstra, 2005

technology is the scorecarding capabilities of Enterprise Resource Planning (ERP) systems.³⁸ *“ERP systems appear to be a dream come true. These commercial software packages promise the seamless integration of all the information flowing through a company...”* These systems hold great promises, but sometimes turn out to be costly and difficult to implement.³⁹ 28 percent of the surveyed companies use systems especially designed for the purpose of measuring supplier performance and finally 12 percent use their e-sourcing system to evaluate supplier performance.

Why suppliers are not measured

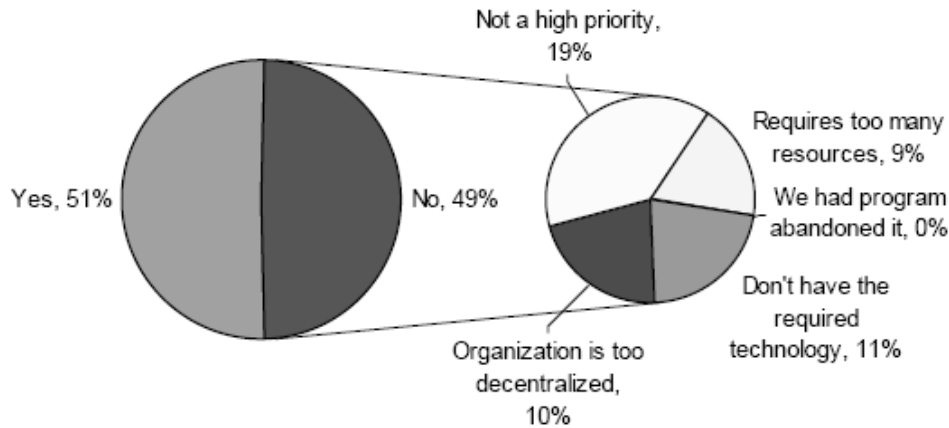


Figure 7: Supplier performance measurement system adoption levels and reasons for non-adoption⁴⁰

According to Shepherd and Günter, few academics that have explored *“whether the benefits of supply chain performance measurement systems are outweighed by the cost of implementing and maintaining them in increasingly dynamic business environments.”*⁴¹

As seen in previous paragraphs, there are plenty of benefits to gain from supplier performance measurement. Therefore it is surprising that a research group has found that *“nearly half of companies still have no formal program in place for measuring supplier performance.”*⁴²

The reasons that half of the researched companies do not have a supplier performance measurement system in place are diverse. It could be a question of priority, where it is not labeled important enough. Other companies conclude that it requires too many resources or that there is a lack of appropriate technology. The companies' organizations could also be too centralized which for example could imply that there is no gathered or standardized data, low technical integration or poor coordination.⁴³ The possibility of performance measurement not being worthwhile increases for small companies, which have a shortage of the resources required to optimize performance measurement activities.^{44 45}

Design of supplier performance measurement systems

*“Few studies have developed performance measurement systems, delineated metrics, or benchmarked supply chain practices.”*⁴⁶

However, there is a best practice in supplier performance measurement that advocates that the practice of measurement should not only align with

³⁸ Aberdeen group, 2005

³⁹ Davenport, 1998

⁴⁰ Aberdeen Group, 2005

⁴¹ Shepherd & Günter, 2006

⁴² Aberdeen Group, 2005

⁴³ Hervani, Helms & Sarkis, 2005

⁴⁴ Shepherd & Günter, 2006

⁴⁵ Morgan, 2004

⁴⁶ Shepherd & Günter, 2006

corporate strategy, but also reinforce specific strategically important parameters.⁴⁷ To further reinforce the importance of the measured activity the supplier's performance should be connected to an objective, which in turn should be connected to an incentive for e.g. the purchaser or supplier.⁴⁸ Neely describes this practice as the highest level of analysis where the corporate strategy is linked to the whole practice of measuring suppliers within a company.⁴⁹

There must be a clear causality between the measurement and the overarching goal. The scope of measurement can be holistic, profitability of the company; or more partial, the number of rejected supplied goods. When measuring at a higher level it is easier to establish if the measure is relevant, but it is harder to prove direct causality between the outcome of a metric and the fulfillment of the corporate strategy, since the strategic goals are dependent upon many other factors.⁵⁰ Neely states that the link between performance measurement and strategy is thoroughly researched in academic literature, even if it "*does not always appear to happen in reality.*"⁵¹

This approach to connect corporate strategy to precise measurements is made operative by Kaplan & Norton in the balanced scorecard. The scorecard emphasizes the relevance of having a clear causal, cause-effect, relation between the specific measurement and the corporate strategy.⁵²

However, the complexity of evaluating supplier performance gets in the way of creating a supplier performance framework that suits all, "*there will never be a one size fits all solution.*"⁵³ Instead the commonalities should lie within the design process that creates the performance system, not the framework.⁵⁴

There are however certain issues that have to be considered and resolved when designing a performance measurement system.⁵⁵ Blenkinsop and Davis have brought forward the specific considerations needed to reflect upon when designing a supplier performance measurement system. They find that it is important to use metrics whose improvement do not conflict with other goals or policies. The metrics also have to be valid indicators; i.e. to measure the number of phone calls a supplier makes probably has very little correlation with its lead-time. There also has to be an understanding of the measurement system within the whole organization, including management consensus on how the chosen measures increase the possibility to attain business goals. The measures have to have clear ownership to avoid the "*it is not my responsibility*" syndrome. Moreover, the measures have to balance fixations with monthly/quarterly numbers to ensure sustainability.⁵⁶

The complexity of how

The focus in academic literature is often to "*specify 'what' to measure and not 'how' to measure, which is an equally important part of a performance measurement system.*"⁵⁷ There is inadequate coverage in the area of *how* to measure supplier performance.

There is very little academic literature on *how* to measure supplier performance

The center of attention in supplier performance measurement literature is to decide *what* metrics should be selected or discarded.⁵⁸ Many articles contemplate on what constitutes a well-designed performance metric⁵⁹ and the

⁴⁷ Skinner, 2000

⁴⁸ Stonich, 1981

⁴⁹ Neely, Gregory & Platts, 2005

⁵⁰ Axelsson, Laage-Hellman & Nilsson, 1991

⁵¹ Neely et al, 1994

⁵² Kaplan & Norton, 1996

⁵³ Morgan, 2004

⁵⁴ Neely, Gregory & Platts, 2005

⁵⁵ Morgan, 2004

⁵⁶ Bleckinsop (in Neely, Gregory & Platts, 2005)

⁵⁷ Paranjape, Rossiter & Pantano, 2006

⁵⁸ Ibid

⁵⁹ Neely et al, 1997

factors that hinders and ruins the design.⁶⁰ Paranjape et al illustrate this focus on *what* with an excerpt from an interview with Robert Kaplan.⁶¹ The creator of the balanced scorecard says that “*It is interesting to notice that neither David Norton nor I is an expert in measurement techniques. David says: We are experts in what to measure, not in how to measure.*”⁶²

There is an attempt made by Tangen to approach the *how* in the process of designing performance measurement systems.⁶³ Paranjape et al’s comment upon Tangen’s concluding formula, is that the practice of *how* to measure is too complicated to “*be addressed by a simple formula.*”⁶⁴

The difficulty of *how* to measure resurfaces again in the implementation stage where abstract concepts are to be concretized. Companies are prone to exaggerate the number of included metrics in order to be thorough which is operationally expensive and very hard to implement and utilize.⁶⁵ This implies a focus on *what* to measure rather than *how*. The choice to include as many metrics as possible could be made worse by the rise of more efficient information technology systems which makes it possible, but still not desirable, to measure everything.⁶⁶

When examining the theory summarized above, perhaps the most interesting find is the last part: how complex measuring is in practice. Measurement theory is often discussed on a conceptual level, and does not analyze many of the practical problems that appear “on the factory floor”. This theory gap is an opportunity to make a contribution, by unearthing and discussing some of these complexities through a detail-focused study. The type of methods used to gather these insights are outlined in the section below.

Method

In this section we present how the study has been set up and how the empirical material has been gathered.

The Study

Scope of the study

To get information on how companies work with supplier performance measurement we interviewed in total **25 persons** from **13 companies**. Seven companies were part of the host corporation and six were other companies with no relation to each other. These are called the benchmark companies. Of the companies from our host corporation, four were based in other European countries. The benchmark companies were based primarily in Sweden.

Seven of the companies were in the manufacturing business. This is a broad area, which includes mechanical and electronic products. Some of these companies manufactured rather simple goods, while some had more advanced production, involving many parts as well as electronics. Three of the companies came from the automotive industry, one was in the construction industry, one in the chemicals industry and one in the retail business.

The interviewed persons worked mainly as purchasing managers and quality managers. We also interviewed several persons at our host corporation’s head office to get background information about how supplier performance measurement worked on a broader level. All companies and persons have

⁶⁰ Dalton, 2004

⁶¹ Paranjape, Rossiter & Pantano, 2006

⁶² De Waal, 2003

⁶³ Tangen, 2005

⁶⁴ Paranjape, Rossiter & Pantano, 2006

⁶⁵ Ibid

⁶⁶ Axelsson , Rozemeijer & Wynstra, 2005

been made anonymous. However our tutor has been given much more detailed information so he could assess the quality of the data.

A qualitative study

When gathering the data we chose to use a **qualitative approach** since it fits the purpose. Among other things, it gave us the ability to capture a wide picture of complex issues and to ask follow-up questions when something was unclear. It allowed us to go deep into the complicated questions that arise in this type of study,⁶⁷ especially with reference to the purpose. A quantitative approach would have required large amounts of standardized data and been impractical in this case.

The qualitative approach brought us to use an **inductive method** for analysis of the data. This means that we make observations and give a relatively broad description of reality, and thereafter amalgamating the knowledge into general conclusions and recommendations. In the Discussion section we use the conclusions drawn to be normative and give some recommendations on how to construct a good supplier performance system. The results are naturally not analyzed statistically, but an **analytical generalization** is made based on the theories. This means that generalizations can be made for cases where the background theory is similar to that of the cases described in this study.⁶⁸

How the Data was Collected

The interviews

We tried to set up interviews with more than one person from each company in order to view the issues from several perspectives, and increase the reliability. There might for example be a difference in what a quality manager and a purchasing manager emphasize. The goal of interviewing several persons was reached for 10 of the 13 of the companies we visited. All interviews took place person-to-person at the site of the respective company, no telephone interviews were conducted.

The interviews were conducted with the help of a question template (see Appendix). The interviews were semi-structured, meaning that we covered a set of main questions from the template but allowed people to branch out and talk about related matters. We specifically tried to ask many follow-up questions, as the issue of how an organization works with supplier performance measurements is a complicated one.

Other types of data gathering

After the interviews had been made, we would ask if we could see documents that described the supplier performance routines. Much of that information was confidential, but from a couple of companies we did receive some printouts with supplier performance data as well as documents outlining some procedures.

In some companies we were also allowed to take a look at their supplier database via a computer. This was of great help for us to understand how the systems worked. It also increased the reliability of our findings since that information could confirm the oral information given by the interviewed persons. In almost all cases we also took a look at the factory floor and the actual production process, which gave us a broader understanding of how the organization worked. At several of the companies we were showed exactly where and how the supplier performance measurement took place.

Documenting the interviews

Everything that was said during the interviews was recorded with a digital voice recorder. With the exception of one occasion, all people approved of being recorded. Afterwards we listened carefully to each interview and wrote

⁶⁷ Merriam, 1994

⁶⁸ Yin, 1994

down everything that was said. This means that all quotes in this text are word-by-word. Each interview was in the magnitude of 7 to 11 A4 pages. This text was then sent by email to the interviewed person and approved by him or her.

Validity and Reliability

All thirteen companies are treated as independent companies

The seven companies from our host organization have been treated in the same way as the six benchmarking companies, that is, as independent companies with their own ways of working. We did this even though the seven companies are part of the same legal entity. The reasons for this are many: Our host company is made up of over 150 rather independent companies. Most of them have been acquired during the last decade and centralization in the host company is low. As a result of the low centralization, these companies have almost no history of cooperation within the supplier performance measurement area. Moreover, the companies are both more and less advanced with regards to their purchasing function, four of them were based in other countries than Sweden, and finally, they manufacture different types of products.

Data gathering and reliability

Reliability is a concept that describes the difference between the measured results from a study from how the reality looks. The difference emerges due to random errors in the measuring.⁶⁹ One way of looking at this is: would one get the same results if conducting another study of the same format? If the answer is yes, the reliability of the first study is high.⁷⁰ As for the reliability of this study, interviews of course depend on the individual respondents. Asking the same questions to different persons from the same company or asking the questions with a different wording or attitude might yield slightly different answers, i.e. random errors in the measuring. One way of ensuring a higher reliability is to measure in several ways.⁷¹ We tried to improve the reliability by, in most cases, interviewing several people from the same company, asking many follow-up questions and asking for clarification, as well as receiving documentation in the form of electronic documents and factory walks that could verify what we had been told.

Selection of companies and validity

Validity is a concept that describes how well one measures what one says that one is measuring. The aim is to give a good view of the reality.⁷² High validity means that there is a lack of both random and systematic errors and bias in the measuring.⁷³ The six companies we interviewed outside our host corporation were selected in two ways: Firstly we wanted to talk with companies from different industries to get a wider perspective. This goal was achieved as six different industries are represented in this study. All of these companies are from Sweden and they are not affiliated with each other.

We selected the companies based on lists of big actors in the respective industries. The sample was not random, it was more like a convenience sample: We contacted the biggest and/or most successful companies in a number of relevant industries. Many of the contacted companies did not wish to participate in the study, but the chosen companies are still some of the biggest in their respective industry in Sweden. This might introduce a selection bias, if companies that accept such an interview proposal have distinct properties from those that decline.

⁶⁹ Söderlund, 2005

⁷⁰ Merriam, 1994

⁷¹ Söderlund, 2005

⁷² Merriam, 1994

⁷³ Söderlund, 2005

With regards to the validity of the study, it is important to remember that we interviewed companies mostly in Sweden, and mostly within the manufacturing industry. In essence, the scope of this essay did not permit interviewing a statistically significant large number of companies. Another factor that could decrease the validity is that we conducted a study which, for seven of the companies, was commissioned by their corporate head office. Because of this it is possible that the respondents' answers could be affected by hidden agendas, that is, that they want to give biased answers due to hidden interests. That would introduce a systematic error in the result of the study. We addressed this issue by asking for evidence to substantiate the answers, and by downplaying our affiliation with the host company: we were only temporary employed at the company and our goal ultimately was to pursue an academic purpose.

Empirical Findings

In this part we present the information gathered from the interviews, in summarized form. It begins with a list of the companies interviewed. After that follows descriptions from each of the companies. In the cases where interviews with several persons were made at one company, they were put together into a single text for the company. In the end we present a matrix that compares the practices in each of the 13 companies.

List of companies

	Company A	Company B	Company C	Company D	Company E	Company F
Part of Host organization or Benchmark company	benchmark	host	benchmark	host	host	benchmark
Industry	automotive	manufacturing	chemicals	manufacturing	manufacturing	retail
Purchase of direct or indirect material	direct	direct	indirect	direct	direct	direct
Number of persons interviewed	2	3	2	2	3	1

	Company G	Company H	Company I	Company J	Company K	Company L	Company M
Part of Host organization or Benchmark company	host	host	benchmark	host	host	benchmark	benchmark
Industry	manufacturing	automotive	manufacturing	manufacturing	manufacturing	construction	automotive
Purchase of direct or indirect material	direct	direct	direct	direct	direct	direct	direct
Number of persons interviewed	2	2	1	2	3	1	2

COMPANY A

Multinational company in automotive industry - Interviews with Supplier Development Purchasing Manager and Plant Quality Manager. This was a benchmark company.

Which suppliers are evaluated?

All suppliers are evaluated. The Company has 1300 European suppliers and the requirements to become an approved supplier are very high. If a division has identified a potential supplier, central purchasing must verify the supplier's suitability before any purchases can be made. There is currently an ambitious plan to decrease the supplier base with 15 percent a year until merely the 200 most competitive suppliers are left.

What metrics are used?

PPM

When an error is encountered either in the reception control, production or at the customer, it is reported in a web-based system. A complaint includes the number of faulty items and this number is transferred to a supplier performance system. PPM is automatically calculated by taking the number of errors compared to the number of received goods, from the logistics system.

Every complaint also includes a problem description which triggers a problem-solving activity. The supplier is probed by an automatic email to respond to the situation with an action plan. The Company's PPM target is 25.

Delivery precision

Once a week the Company issues its delivery plan to all suppliers. The supplier is given a delivery time slot. The acceptance of lateness is extremely low: "If a supplier is five minutes late, it will cause a remark". Upon arrival all goods are checked and measured for full delivery and correct labeling. If there are remarks a complaint is filed that triggers the same kind of problem solving activity as the quality-related complaints do.

Number of written complaints

All complaints, regardless of their nature, are graded after severity. If the supplier's error causes a customer incident it is considered grave. An error that does not affect the production is minor.

Number of customer satisfaction complaints

If a supplier fails to respond to a PPM or Delivery precision complaint, the Company files a customer satisfaction complaint. This kind of complaint is

also a response to defective or tardy documentation, which is a part of the production part approval process.

Commercial competitiveness

Central purchasing is responsible for measuring a supplier's competitiveness annually. Examples of measurements used are annual price reduction and number of price reduction suggestions from the supplier.

How is the supplier data stored?

The whole process is computerized. The Company has one system where complaints are filed. *"It is mainly the quality and logistics departments that provide input to this system"*. The information from this system is automatically transferred into the supplier performance system. In order to calculate e.g. PPM the system also receives information from a material planning system.

"Previously we started a diverse range of supplier evaluations that were all stored in excel files. After a while the files became too cumbersome to handle [...] now we are happy that we invested in a good system [...] it is much more transparent". Moreover, the transfer to a common database has meant less administration. *"Even the suppliers like it since it is an easy standardized procedure used throughout the Company"*.

All suppliers can go online and access information about their performance. The Company even demands, and measure, that the suppliers do it. *"Suppliers should review their performance information at least monthly and then submit action plans to improve the parameters."*

How is the supplier data used?

From the supplier performance system reports can be obtained: *You just choose whatever you want to look at. You can look at specific suppliers, time periods, production units, etc. It is easy to find out which suppliers perform well and which do not. With this information it is easier to drive suppliers toward improvement."* Each month and quarter the suppliers are reviewed and the five worst performing suppliers are identified and invited to an improvement meeting. This process takes place at three levels: production level, division level and central level. *"We have worked with the five worst performing suppliers for many years. If we solve those five worst, we perhaps solve 75 percent of the total problems. If we solve those problems, we take the next five."* Moreover, there is champions program with meetings for suppliers of *"strategic importance"*.

This year there will be changes in the sourcing procedure. Requests for quotations will take place within the system in order to bind the purchasers to the already approved suppliers. When asked if there were anything they wanted to change or any disadvantages with the system the answer was: *"We do not want to change anything. [...] All industries must ensure good quality and I think all companies can gain from a similar system to measure supplier performance and set goals."*

COMPANY B

Manufacturing company - Interviews with Purchasing Manager, Category Manager and Production Manager. This was a company in the host organization.

Which suppliers are evaluated?

The Company has about 200 suppliers of direct material. It has segmented the suppliers according to the 20-80 rule, so the 20 percent of the suppliers that account for 80 percent of the purchases, that is 39 suppliers, are evaluated.

What metrics are used?

Delivery Precision

When goods arrive, the delivery date is compared to the confirmed delivery date. The percentage of correct deliveries is calculated and the supplier is given a percentage as grade. The supplier has different levels of responsibility. In "free carrier" contracts the responsibility to fulfill the delivery lies with the transportation company, and not the supplier, so in these cases the Company adds three days on the delivery time for transports within Sweden and one week for transports within Europe. These deviations in dates require the purchasers to work actively in the computer system. There is also another similar problem: *"Principally we can make an impossible order, which is making an order today for delivery we want yesterday. Our computer system does not critically review the data so that would mean the supplier is measured as late before he has even received the order."* The reliability of the delivery performance data is still deemed as giving *"a fair enough picture"*.

Faulty parts are naturally rejected, but the Company does not measure and gather data about quality. One of the interviewed people commented: *"There is too little focus on quality, but it has its reasons. Previously we produced everything with an impact on quality in-house and as long as our internal control was good, our quality was good. But now we purchase much more."*

Another function the company would like to have is the severity of the lateness for late deliveries: *"If a supplier has a delivery precision of 90 percent, and the delayed 10 percent is one day late, then we will hardly notice the delay. Another supplier could be better on paper by having 95 percent goods on time. However if the other 5 percent are 10 weeks delayed, we will have to stop production."* Conclusively, the Company sees a need to make the measuring of delivery precision automatic and to work more with, and measure, supplier quality.

Financials

For the biggest suppliers, the Company checks their financial data to ensure that it sources from stable suppliers.

How is the supplier data stored?

The delivery precision information is manually extracted from the Company's ERP system and gathered in excel files each quarter. Only purchasing can access the information. All other supplier information, e.g. annual reports and protocols from meeting, is stored in physical files, in paper format.

How is the supplier data used?

When a problem appears it is first dealt with by the material planners. If the problem is occurring repeatedly, the responsibility is transferred to the purchasers who set up meetings and create an action plan, or re-source. The Company was honest about its situation, saying that it believes that *"most of our suppliers have customers that are more demanding than we are"*. There was some good results recently though when a new evaluation/assessment tool was tried out at a supplier. The Company pushed to see improvements and the result was very good.

Each quarter the suppliers are sent a copy of their performance. *"Some of the suppliers are very active and send us their own delivery precision data when they believe our data is wrong."* The quarterly report identifies suppliers with bad performance. *"The problem is that suppliers can deliver to different business units in the Company, which are all run as separate companies with low awareness of what is happening in the other units. This means that we have a lag in detecting problematic suppliers. An error can appear on the first day of the quarter and we will not notice it before the end of the quarter."* Within the Company there is a wish to become more proactive.

COMPANY C

Multinational manufacturer of chemicals - *Interviews with Supply Management Director and Category Leader for Equipment. This was a benchmark company.*

Note that Company C is different since we discuss also purchasing of indirect material. (The "Equipment category") The Company is however still relevant for this study, as many of its processes, such as supplier segmentation and data management, are the same for both its direct and indirect purchasing, i.e. we learn about its direct material processes.

Which suppliers are evaluated?

The Company

The Company has recently segmented its direct and indirect suppliers into three groups, after doing a benchmark study of how other companies worked with suppliers. The groups are: 1. Opportunistic suppliers: non-strategic purchasing with automated processes and focus on price. 2. Partnership suppliers: a diverse group of suppliers where some are very important. 3. Strategic suppliers: a very limited number of high-risk suppliers. Each supplier has an 'owner' in the Company but there is a cross-functional team working with important ones. Direct material suppliers are audited, but very seldom those supplying indirect material.

The Equipment category - indirect material

"We focus on the suppliers that stands for the majority of the monetary spend. Then there is a long tail of small specialized suppliers we do not pay much attention. There are no strategic suppliers in this category."

What metrics are used?

The Company

The purchased goods are divided into many categories, each led by a Category leader. The leader and his team decide on a number of supplier performance metrics to be used for the category. Most categories use around 5 metrics. Delivery precision and Quality is used by most. Company-wide, there is data on the total spend on different suppliers.

The Equipment category - indirect material

- Delivery precision: For this category, Delivery precision is not critical, instead they rely on the supplier's delivery data, while making random checks by asking key internal users about problems.
- Lead times for service of the equipment: The Company relies on supplier data, there is no data gathered about the lead time or the service occasions. *"We go on perception here, but if there is a large problem our internal users usually communicate that clearly to us."*
- Internal survey: For the 8 largest suppliers to this category, the Company sends out a survey to key users in the organization, who then rate the suppliers on service, competence, willingness to help, and other subjective metrics.
- Quality of the equipment: This data is stored on local systems or in local excel files and measured in different ways. There is no category-wide data, but a standardization process is underway. Sub-metrics that will be included are e.g. Mean time between failure and Maintenance cost.
- Financial stability of supplier: Only for the largest 5-8 suppliers in the category. A future goal is to standardize this metric for the largest suppliers in all the Company's categories.
- The supplier's technological development and suggestions for improvement of the products.

For the future, the person we interviewed would like to have a clearer opinion from a representative of the internal users, what the users think about a supplier and its products. Today there are too many internal opinions and it is difficult for the interviewed person to get a clear picture. *"In this category, spend, Delivery precision and such things are less important than how the users work with and perceive the equipment. We want to be more proactive and*

communicate our needs better to suppliers and create stronger relationships. There is so much they can help us with."

How is supplier data stored?

The Company

The metrics data gathered in the categories, as well as data from audits, is not stored in a company-wide database. Each category has an own system. However, *"it is not certain that building coordination would pay off"*, since the categories are quite dissimilar. There are also diverse ways of working with purchasing in different global regions in the company.

How is supplier data used?

The Company

The data gathered is communicated with suppliers through meetings. Suppliers cannot log in to any online system and see their own data. The meetings are held 1 to 4 times a year with Strategic and some Partnership suppliers.

The main discussion points at the meetings are: 1. Current issues and future plans. 2. Contract, terms and price. 3. How the supplier has performed according to the metrics. 4. Corporate Responsibility issues. What was said on a meeting is summarized and uploaded in each category's supplier information system.

The spend data stored in the ERP is also used to evaluate purchasers internally. For example, purchasers should give more business to suppliers that previously have been deemed to perform well.

The Equipment category - indirect material

"On the meetings with suppliers we naturally have different perspectives about problems, but we will still pretty quickly reach a consensus."

If there is a problem with the metric 'Service of equipment', there is an escalation process where a technician takes the problem to higher and higher levels in the organization. For the other metrics there is no formal escalation process. *"We haven't felt a need for it because dissatisfied users make their voices heard."*

COMPANY D

Manufacturing company - Interviews with Purchasing Strategy Manager and Purchasing Manager. This was a company in the host organization.

Which suppliers are evaluated?

The company has roughly 170 suppliers of direct material, from which purchases are made regularly. All purchases are ranked into three segments after their impact on the quality of the end product. The segment of greatest importance, segment 1, includes 11 suppliers. In segment 2, the components have a moderate influence on the final product. Components in segment 3 do not influence the final product, for example packaging, labels and paint.

What metrics are used?

There are up to three groups of metrics used in the supplier evaluation depending on the segment. All suppliers are however evaluated every quarter. The grades given for the three different groups of metrics are A, B and C. *"A is perfect, B is when something is not quite right but we are quite ok with it, and C is when there is something wrong. We tolerate B grades."*

Purchasers' opinions

This group of metrics is used for suppliers in all three segments. *"The first part of the supplier evaluation is the grades given by our purchasers. I prepare the list of suppliers and give it to our purchasers who give the suppliers a grade. The first line is how suppliers keep delivery and quantities. The next line*

is how willing suppliers are to communicate and be flexible. The third line regards the quality of packaging. Then it is price, if it is low or high or middle, invoice maturity, and the last one is terms of the supplier."

Each purchaser has books of deliveries for each supplier, but mostly the information that the grading is based on is in his head. *"Not everything is possible to have on paper, some things are up only to people."*

"We have tools in our ERP system where you for example can receive a report for each product: when we ordered it and when we received the delivery, and we can make statistics from it, but sometimes it can be information that is not absolutely correct, because sometimes we order something but don't confirm with customers, since sometimes it is ok with our customer in case we postpone it, sometimes it is even welcomed. You have to find an agreement."

Quality control

This metric is used for suppliers in segment 1 and 2. It is the incoming goods quality control. *"The incoming quality control gives us a list of the claims. If the number of claims is 0, it is an A. If the number of claims is below 5 percent it is a B. If it is above 5 percent it is a C."*

"We control quite much. When a delivery comes, the incoming quality person has information about it in our ERP system and there is a plan for the control of the components. For example if the delivery is 20 000 pieces, we have to take 10 sample pieces and measure them and write down the measurements in a computer and see if the tolerances are correct."

Certification

This part of the evaluation is used for suppliers in segment 1. It is to ensure that they have a valid ISO certification. The majority of the suppliers gets A:s, some get B:s, and a very small number get C:s.

How is the supplier data stored?

There is an ERP system where all quality and delivery data is gathered. *"Everything is going through our ERP system so most information is possible to extract from it. We know how many components are in stock, in production, are machined, are waiting for machining, etc."*

The evaluations for all suppliers each quarter is made on Excel sheets. When these are finished the grades are submitted to the ERP system *"and when a purchaser is to purchase, he will see the grade that each supplier has, in the system. The system recommends him suppliers with the best grades."*

How is the supplier data used?

Suppliers are contacted directly by people from the concerned departments when a problem arises. When the quarterly evaluations are done the strategic purchasing manager *"inform all the important suppliers in segment 1 about their evaluation results. As for the other suppliers, I have to inform them only if they have the grades B or C."* The suppliers are then asked to submit an action plan with possible improvements. If a supplier has grade C it initiates a process where the Company decides whether to keep the supplier and ask from improvements, or to resource.

The purchasing manager also writes an action plan: *"The quality department reports about the problems and I report what we are doing with each supplier to improve quality."* This evaluation system has been used since 2001. Since then it has been used every quarter with some adaptations. There is a general agreement that the system works fine: *"I think the system is ok and the results of the evaluation fit to what we need. I would not want to change anything now."*

COMPANY E

Manufacturing company - Interviews with Purchasing Manager, Quality Manager and Purchase & Logistics Manager. This was a company in the host organization.

Which suppliers are evaluated?

There are around 150 suppliers of direct material and those are evaluated. For a long time the Company produced most things in-house but has started outsourcing the last years.

What metrics are used?

Quality

There is a reception control that tests sample goods that are critical in the production, and goods from suppliers with previous problems. The most common test method is that the physical properties of the goods measured and compared to the drawing. The result is PPM data for the biggest supplier and Number of rejections data for the rest of the suppliers. *"When a fault is found it is entered into the ERP system. A report is written with details about the problem. The report is sent to the supplier and the person at the Company who responsible for the product. He decides if the goods can be modified and used, or must be sent back to the supplier. Our quality manager classifies the report as completed when the supplier has responded what actions it takes to prevent the problem from happening again."* These rejection reports are updated continually.

When faulty parts are discovered in the production, the Company charges the supplier an administrative fee for the cost of writing the rejection report. Sometimes the Company also demands compensation for the extra time it takes to modify faulty parts on the spot. But normally faulty parts are not repaired since the supplier in those cases can make arguments like *"we could have fixed those parts ourselves for half the cost"*.

Delivery precision

This is a comparison of the supplier's promised delivery date with the actual date. Deliveries that arrive early count as correct deliveries.

How is supplier data stored?

Delivery precision data is automatically extracted from the ERP system and stored in Excel files, one for each supplier. The rejection reports are also in Excel format. All employees have access to this information on the server. There are no subjective/soft metrics, but it was mentioned that much of that kind of information about suppliers (specialties, communication, technical ability, etc) is stored particularly in one person's head: a purchaser that has worked in the Company for many years.

How is supplier data used?

Once every month the quality performance and delivery performance of suppliers are reviewed, except for one very big supplier where these metrics are evaluated every week. Once every year a more thorough evaluation takes place.

There are no reports going out to suppliers that tell them how they are performing. *"We would like to have monthly reports that say something like: 'on quality and delivery you have this performance, and you are ranked as the 20th best supplier on quality and 10th best on deliveries.'"* Moreover, supplier of the year contests is something that could be developed, according to the same interviewed person. *"We were part of a supplier of the year contest arranged by one of our customers. It was quite inspirational that someone gave us feedback on specific performance in relation to other companies."*

COMPANY F

Multinational retail company - Interview with Quality Manager. This was a benchmark company.

Note that this is a retail company and do not have its own production. This gives us a slightly different perspective on the company's purchasing.

Which suppliers are evaluated?

The Company sells many types of goods to consumers and uses thousands of suppliers. *"It is not practical to control every supplier or product."* Goods from large suppliers are purchased centrally and amount to about 70-75 percent of the total spend. The rest is purchased by the Company's local stores from small suppliers.

Many suppliers are audited centrally, and focus is on quality management systems, but there are different procedures for different kinds of suppliers. Re-audits take place every 12-18 months, by the Company or by third party certifiers. Big suppliers in the industry get many visits from similar buyers so they instead prefer a third party company to make a single audit. Well-known trusted suppliers are not audited at all.

Each category is led by a quality manager and a team. *"Purchasers and Product developers must work tightly together with us at Quality. Sending them reports is hopeless, they must see a quality problem with their own eyes to get a good understanding."*

What metrics are used?

There is a reception control on most goods where delivery precision and some types of quality parameters are checked, but the interviewed person did not work in logistics.

Quality

Samples of products are taken from the stock and routine chemical and functional tests are made, about 1-4 times a year per product. If the quality is consistent the frequency of tests decreases, but if customers complain about a product the frequency increases. Some products are simple and are only tested in case of customer complaints. The tests are different for different kinds of products, but generally the purpose is to see that the goods meet the required standard. Actual old samples are stored so that the quality over time is well documented.

In addition, for sensitive products a group of experts and a group of laymen make a sensorial and subjective evaluation of the quality. If there is no conclusive result from these tests or if a product is very important, a group of consumers can be brought forward to test the goods, through a market research company.

Faulty goods discovered in the stores are sent back to the central warehouses. If the company detects one damaged good, often the whole batch is returned. A process starts and the supplier must correct what is wrong. There is not an automated routine for this; instead different people from the Company email the supplier. *"The follow up on a supplier depends much on how competent our Company's contact person is."* In case of repeated problems the supplier might come to the Company for a meeting. Or sometimes the Company sends out one person from Quality and one from Product development to take a look at the supplier's problem. If a very serious quality problem should be discovered on a product that is out in the stores, the Company can via its IT system stop that product from actually passing through any cashiers.

Suppliers pay the cost for recalled products and for visits by the Company. *"But suppliers pay only the administrative cost, we have no penalty fees. We want to be able to have a pure negotiation on price. That simplifies the negotiation and we get better prices when the supplier does not have to take a lot of other costs into account."*

There is no total grade for a supplier's quality. Instead specific problems are followed up, as described above. As long as a supplier produces according to the requirements it is fine. *"We could work more on continually improving a supplier/product."*

How is supplier data stored?

Quality data is stored in a database. There is no common supplier evaluation system with all types of data like deliveries and prices. The suppliers do not see their own data in any of the sub-systems in the Company. *"A common system for supplier data in the Company would be good, but so far it has been difficult to build. We want a system where suppliers can log in and see their own performance and their recalls, but we have not implemented it. It is quite complicated since we have made a lot of acquisitions, we have many suppliers delivering the same types of goods, and suppliers are coming and going."*

"Generally we have a very good quality control but we could improve in sharing information internally and in tracking goods through the entire supply chain. As an example: "When recalls are made, we who work at Quality do not see everything, we request reports from Logistics on the exact numbers of goods that have come back from the stores."

How is supplier data used?

"Each month we have a meeting with people from our different departments to discuss the product, recall, tests, future development and other issues."

COMPANY G

Manufacturing company - Interviews with Purchasing Manager and Supplier Quality Manager. This was a company in the host organization.

Which suppliers are evaluated?

There are around 170 suppliers of direct material and all are monitored. There is extra focus on the top 20 percent suppliers, who stand for more than 80 percent of purchases. *"I think the number should be reduced since a supplier that sells 1-2 percent of its goods to us is not going to be very responsive."*

What metrics are used?

Cost

This is the most important metric. Price development of the direct suppliers is followed monthly. The definition is how much the price has increased/decreased from the date 12 months ago.

Delivery precision

Goods are correctly delivered if they are 'on time in full'. A delivery that is one day early is allowed. The target for delivery precision is 95 percent, but it is not there yet. The data is evaluated each month by the purchasing department. Raw material is difficult to deliver in exact quantities so a 10 percent variation in volume is allowed from these suppliers. Previously the Company measured the exact number of days all late deliveries were delayed. However, things became too complex with that amount of data, so those measurements were abandoned.

Quality

PPM is the metric, and it is evaluated each month. The target is a PPM of 2000. The data is not entirely correct as there are some difficulties in measuring PPM. For example if faults are found in a sample test and the supplier quickly comes and fixes everything, should the faults be counted in the PPM? Generally the Company does not do that. *"We need clearer rules, but even perfectly clear rules will be bent. Basically, if parts can be fixed and used they will be that since the Production department is very concerned with meeting delivery times."*

Corrective actions

These are reports sent to the supplier when goods are faulty. The Company keeps track of how many Corrective actions have been sent to each supplier. It is also noted how many of these are 'open'. A corrective action is open until the supplier has presented how it will deal with problem. *"That a supplier acknowledges a problem or makes extra inspections is not enough, we need a robust root cause analysis and action plan that solves the problem once and for all."* There is a target for a total number of open corrective actions of 20 in total, but there is some way to go. *"The point of corrective actions is to identify those suppliers that have the same problem many times. PPM does not show that. We must look at how good a supplier is at fixing problems."*

How is the supplier data stored?

The quality and delivery data is stored in a customized database and evaluated each month. There is data for a rolling 12-month period so the trend is visible. There have been problems with getting employees to gather all data for the system, and as a result new resources have been put in and the routines are being reworked.

How is the supplier data used?

The five suppliers with the highest PPM each month are visited by the supplier quality manager and action plans for improvement are implemented.

COMPANY H

Company in automotive value chain - Interviews with Purchasing Manager and Quality Manager. This was a company in the host organization.

Which suppliers are evaluated?

There are around 100 direct suppliers, of which all are evaluated. *"If there is a concern with a supplier of indirect material it will be pursued, but we do not monitor any metrics for indirect suppliers."*

What metrics are used?

There are two metrics for supplier performance, Quality and Delivery performance. *"As long as a supplier delivers what you ask for on time and in good quality, there are no problems."*

Quality

Each supplier is measured on its PPM. *"Because we reject whole batches when we find errors and not just the faulty parts, the PPM can vary a lot from month to month."* Depending on the severity of a problem, the supplier comes to the Company and sorts the parts, or the parts are sent back to the supplier. The total target for the whole supplier base is about 6000 PPM and the Company is currently below that. The goal is continuous improvement.

The routines for quality control: *"When we have a supplier for the first time we have a reception control for a probationary period of 10 deliveries. If there are no concerns, we then accept the goods without checking the quality. If a problem occurs in the future, we reinstate the goods inward inspection."* There are also some sample tests, based partly on the BS6001⁷⁴ standard sampling procedure. If a batch is large, it might be accepted even if there are one or two faulty parts.

On a question about how exact the PPM figure needs to be, the quality manager responded: *"It does not need to be absolutely accurate, but accurate enough to detect changes and give us information on what to improve. Whether you improved by 400 or 500 PPM is not vital, the important thing is you know you have improved. If you try to be totally accurate and want to know if you have rejected exactly 179 parts, it will not work because there will always be one or two parts that slip through the net and go unnoticed."*

⁷⁴ A standard for sampling goods, used in the UK.

Delivery performance

Data about deliveries is stored in real-time in the ERP system. At the end of each month, the quantity of the delayed deliveries is subtracted from the total quantity that has been, and should have been, delivered during the month. So the Delivery performance figure is a snapshot at that day. *"The average delivery performance is about 96 percent. We are very lean and need to rely on correct deliveries."*

How is the supplier data stored?

The PPM and Delivery performance statistics and the Corrective actions (see below), are all stored in the same Access database. Charts of the development of these metrics, as well as internal performance metrics, are printed and put on the factory walls.

How is the supplier data used?

When there is a quality problem, a Corrective action report is emailed to the supplier. It details the problem and demands that the supplier answers with: 1. an immediate action. 2. what the root cause of the error was. 3. a corrective action to prevent a recurrence. All Corrective action reports are stored in the database and they are classified as either open or closed. They are closed when the supplier has solved the problem, which can take a varying amount of time depending on the problem. However, a supplier needs to acknowledge the problem, i.e. send a response to the report, within 24 hours.

Regarding the periodicity of the follow-up of suppliers, this comment came up: *"You can generate a lot of data, but it is only worth it if someone is actually sitting down and digesting it. I would think generally speaking that one month is a good interval for following up suppliers. One month gives you plenty of time to capture a lot of data, and it does not mean that you must jump on problems immediately. You also have a natural break throughout the year."*

COMPANY I

Multinational manufacturing company of advanced automated products - Interview with Operational Supply Manager. This was a benchmark company.

Which suppliers are evaluated?

Suppliers deliver very different types of goods, but they are all measured in the same way. There are around 250 suppliers for direct material. There are strategic and non-strategic suppliers, segmented on volume and complexity of goods. The strategic suppliers get more attention from the Company's strategic purchasers. How the suppliers are evaluated is described in a Supplier requirements manual.

What metrics are used?

Delivery precision

A delivery that is not full or on time is an incorrect delivery. The target is a Delivery precision of 98 percent, that are complete and in time. But the tolerance is slightly bigger in practice for complicated products. Suppliers have a 2-day margin, early and late, for meeting the correct delivery date. At the delivery, the goods are registered in the ERP and the supplier data system and that date is compared with the date previously agreed upon. This matching of dates is not fully automatic because sometimes the Company needs goods with a shorter lead time than the contracted lead time, and the supplier is not late if it cannot meet this extra short lead time. This needs to be entered manually in the system.

One way in which the Delivery precision metric could be improved, is if there was a number in the system for how many days in average a supplier's goods were late. This would give more information than just late/not late.

PPM

The PPM is 500, but as with Delivery precision, the tolerance is higher in practice for complicated products. There is no reception control at the Company, except in special cases, so errors are found by the Production department. Every production line has a quality coordinator with reporting as one of his tasks. An error report with a description of the error is entered into the system, and a call is also made to the supplier. A copy of the report is sent by email through the system, and in paper, to the supplier. The supplier should answer within 5 days what it does to solve the problem in the short term and in the long term. Almost all suppliers respond electronically through the system, and some email their answer.

When an error is found the batch is controlled and the faulty parts are returned to the supplier. If there are many parts to sort, the supplier has to come and sort out the faulty parts. If there are many small parts sorting would take too much time: consequently the whole batch is returned.

If it is urgent or many parts are concerned, the error checking and reporting are done immediately. If not, the faulty parts are put on a shelf and dealt with later. *"That can take a few days if we have other urgent situations to deal with."* If a supplier does not send the reports the Quality department will try to keep track of that manually by checking the report status of each supplier and call that supplier. *"We would like to have a function that sends out automatic reminder emails to suppliers."*

Reporting errors on goods *"is a somewhat of a grey area"*. Sometimes Production needs parts urgently and the number of faulty parts is not too big. Then, the faulty parts might not be shipped back to the supplier, but instead fixed by the Company. *"We can choose either to stop our production or we can make some extra work to fix the parts."* This practice can happen for customized products, but not for the high volume production. The monetary costs of fixing the parts are normally not forwarded to the supplier, since that is an administrative hassle. *"But we might make the supplier pay if there are recurring errors and we want to send a signal."* Another option which also saves time is that the supplier pays for an external firm, closely located to the Company, to come and fix the parts. Such fixed parts are normally included in the PPM figure, but not always. Generally, it is a matter of working together with the supplier for many years: there are rules for most things, but not for all situations. *"We give and we take, we build a trust. It is difficult to explain."*

How is supplier data stored?

All supplier data is online and suppliers can log in to the system and see what orders they need to deliver. Each month a report is made automatically by the system and suppliers can view their result for Delivery precision and PPM for the month in diagram format.

For the Company itself, when one clicks on these diagrams there is more data about each specific delivery: number of days late, number of parts delivered, and about each error and error report sent. The system has been used for many years and it is about to be standardized company-wide, through all divisions.

How is supplier data used?

Small problems are handled by the tactical buyers on an ongoing basis. *"There is no standardized process, we have an ongoing dialogue."* If a supplier performs badly in Delivery precision and/or PPM for a longer time the strategic buyers will move in and arrange a meeting with the supplier and discuss the problem and an action plan. Then there will be follow-up meetings. It is not a standardized process where e.g. a meeting would be held if the Delivery precision drops below a certain point.

The involvement from the suppliers differ greatly, some are more willing to improve than others. *"Some suppliers seem to view us more as a problem than a customer."* The high level of detail on the data means that in that *"there are no protests"* from the supplier when the Company informs it about bad

performance. I.e. the supplier trust the Company's judgment. *"What we can improve is the dialogue with suppliers so they understand our needs."* There are commodity groups within the Company that reviews some suppliers periodically. In those groups a supplier's communication and commitment might be discussed, but there is no system for such subjective evaluations. Finally, the PPM and Delivery precision data is also used for internal reporting. The strategic purchasers are evaluated on how the suppliers they are responsible for improve.

COMPANY J

Manufacturing company - Interview with Supply Chain Manager and Quality Manager. This was a company in the host organization.

Which suppliers are evaluated?

There are around 250 active suppliers of direct material, which are classified as A, B or C. Those in category A are critical to the production, B are moderately important and C are less important.

What metrics are used?

The metrics below are used twice a year to evaluate the 10 A-suppliers, which stand for 80 percent of purchases. However the metrics Percentage of rejected goods and Delivery precision are used for all suppliers since *"these are the most important metrics."*

Delivery precision

A delivery date is compared to the previously confirmed date from the ERP. Deliveries are then classified as either a) In time (5 days early to 1 day late). b) Late1 (2-5 days late). c) Late2 (6-10 days late). d) Late3 (over 11 days late), e) Early1 (6-10 days early, or f) Early2 (over 11 days early). The statistics can be broken down to specific suppliers. Late deliveries are followed up continually by calling or emailing the suppliers.

Percentage of rejected goods

A and B classified goods are normally sample checked by the quality department when they arrive. *"We have three different levels of incoming control. A new supplier is checked more carefully. An old reliable supplier is not checked as much, or maybe not at all."* If goods are faulty the quality department writes a rejection report which is sent to the strategic purchasing department. The report describes the problem with words and a photo. The purchasers consider the amount of goods in stock and decide whether there is enough time to send the goods back, or if the goods should be sorted in the factory.

After the purchasers have thought the problem through they shall within 24 hours contact the responsible supplier, who should return the rejection report to the Company within a certain time. The supplier should answer how it views the problem *"because it might be that we have different perspectives"*, what it does to solve it in the short run, and what it does in the long run. The reports are closed when they are answered.

Quality inspection

This is a subjective grade from 1 to 5, of the overall perceived quality of the supplier. This is set by asking people working in the quality department what they think. *"It is for things too small to write a rejection report for: broken packaging, documents that are incorrect or hard to locate, etc."*

Collaboration

A subjective grade set by the purchasing department, evaluating the customer service of suppliers.

Price development

Change in price. This is usually renegotiated each year, except for raw material where prices are fluctuating.

Technological level

If the supplier is on a high technological level is measured. *"In that case we do not need to teach them."*

How is the supplier data stored?

Data is stored in different Excel sheets. The ERP system is old which makes it difficult to consolidate data in one database.

How is the supplier data used?

There is a list of suppliers' answered rejection reports, late deliveries and order confirmations, counting the number of times the Company has had to remind a supplier to respond. *"But the communication is via email, not through a system, so exact numbers are difficult to set."*

Suppliers with many problems are visited by two persons from the Company, one from Quality and one from R&D, working part-time with helping suppliers with production process improvement. The suppliers do not pay for this. *"The rejection reports sometimes do not help. Chinese suppliers can have difficulties understanding our emails and drawings."*

The suppliers are not shown their own performance data. *"Following up suppliers takes a lot of time; we can improve in that area. Currently, we work mainly reactively with those suppliers who give us the most problems. We want a process that gives us a helicopter view, for strategic decisions about the total cost of quality."* Therefore the Company will start doing larger quarterly reviews of suppliers of A and B products. The suppliers will be given both positive and negative feedback from those working in quality and purchasing about perceived (subjective) performance, and asked how they will improve. Another intention is also to start benchmarking suppliers on how fast they answer their rejection reports, since all suppliers do not answer their reports.

COMPANY K

Multinational manufacturing company - Interviews with Purchasing Manager, Quality Assurance Manager and Material Supply Manager. This was a company in the host organization.

Which suppliers are evaluated?

There are about 250 suppliers, of which 20 percent stand for 80 percent of the purchased volume. These are called A-suppliers and are evaluated more carefully once a year on a variety of metrics.

What metrics are used?

All suppliers are evaluated on these three metrics:

Deliveries on time

When goods arrive, the ERP compares the delivery date with the previously confirmed date. The allowed margin is 1 day late or 3 days early. Late deliveries are followed up by the logistics department.

Quality

This is the percentage parts that are rejected. There is a reception control on high-risk items and goods that have been faulty as of lately. If a sample check finds a certain number of errors among the samples, the whole batch of goods is returned. *"With our statistical models, we accept the delivery if about 1 of 100 parts is faulty."* In case of faulty parts the quality department sends a report to the supplier, which should answer within 10 days what they have done to fix the problem. *"But one must look at each case, sometimes we might have made the mistake of not specifying clearly enough."* The report includes pictures of the error. *"Quality manager: A picture tells more than a thousand words in this situation."* There are weekly meetings where the quality department looks over if these reports have been answered, if not they send

reminders. If a fault is not discovered and makes a complete product break down at the end customer, the Company seeks monetary compensation from the supplier.

Adherence to the lead time

"This is an important metric since we build many things to order and we also work with keeping the stock level low."

The A-suppliers are evaluated on three additional metrics, which are:

Service

This is a purely subjective metric. It includes Response times, Flexibility to the Company's needs, and Communication. The purchasers ask people in their organization for their opinion about suppliers. These metrics were said to be important, but also difficult to measure reliably since different people have different perceptions.

Financials

The solvency, liquidity, cash flow and profit margin of the supplier. The supplier is asked to provide its financial statement and the Company also has a credit information service from the Swedish Upplysningscentralen (UC). Swedish suppliers are easy to find financial information about, but non-Swedish are not. The Company's share of the supplier's turnover is also measured. *"A few years ago around 10 percent of turnover was optimal, since we did not want suppliers to be too dependent on us, because that increased risk. But now when business activity is so high we want suppliers to give us more attention, so around 15 percent is better."*

Purchasing

This includes Price development, Lead time, Terms of delivery, Terms of agreement and Terms of payment.

The end result is a weighted score. If the score is too low there will be actions taken. *"Initially we measured these things for all suppliers but it was difficult to find time, so we scaled it down to the A-suppliers. For one supplier, it takes about 4-8 hours of effective work time to gather the data for the evaluation."* This work is divided among the purchasers. On a question if it is worth the time, this response came up: *"Perhaps the ambition level is too high, it takes a lot of time to contact all the people about the Financial and the Service metrics. Also, the score on Service is often close to what one expected, since we work daily with the suppliers anyway. The most important metrics are clearly quality, delivery performance, adherence to lead time and price."*

How is supplier data stored?

The data about Delivery performance and Adherence to lead time metrics data is stored in the ERP. Quality data is stored in an Access database and rejection reports are entered electronically, and monthly Performance reports are extracted. Most employees have access to the database. There is no compound list that compares all suppliers. *"We need to develop such a list, with approved suppliers."* The other metrics, Service, Financials, and Purchasing, are stored in Excel sheets, one for each supplier.

How is supplier data used?

Each month the data for Quality, Delivery performance and Adherence to lead time for all suppliers are discussed at an internal meeting. Badly performing suppliers are contacted. *"If a supplier does not improve after several problems, we start looking for new suppliers."* Follow-up is partly carried out by an engineer who only works with supplier quality. He visits suppliers to help them improve their production and work routines. *"The supplier quality engineer has produced good results in improving suppliers' performance."* Regarding the cost of this service: *"If we decide that a supplier needs a new tool to improve quality, it will bear the cost, because it has an obligation to provide us with a certain quality."*

"Overall, we need to develop more structured procedures. We have pretty good supplier follow-up tools, but we do not use them all the time. Everyone at the Company does not see advantages with all the metrics." There are also proactive quarterly meetings with representatives from the 5 biggest suppliers. Quality, Delivery performance and Adherence to lead time is discussed, but also other issues like plans for the future, and how to cooperate to find smarter solutions. *"My experience as Purchasing Manager is that one needs continuity in the relationship-building to achieve results."*

COMPANY L

Multinational construction company - Interview with Purchasing Manager. This was a benchmark company.

Which suppliers are evaluated?

The Company has a list of around 26 000 suppliers, since data about all transactions are stored. But only about 5000 suppliers are used frequently. They account for a very large part of all the purchases and *"are important to control"*. Evaluating the other suppliers *"would be too much"*. The suppliers are segmented into three groups: consultants, direct suppliers and subcontractors.

What metrics are used?

Many metrics are used, like call off time, quality, delivery precision, service and flexibility. The most important metric is considered to be delivery precision: *"The market activity is very high so being able to deliver on time is very important."* Ranked second is price, which however is handled as more of a commercial factor. Third is the supplier's ability to cooperate. By working with subcontractors that supply both material and services the Company must rely heavily on the subcontractors' ability to cooperate.

There is a commentary function included in the supplier evaluation where subjective comments on the supplier's performance can be submitted. This *"soft"* parameter is included since evaluating suppliers *"is not black or white"*. These subjective evaluations are elicited with questions like *"how has the supplier's management performed?"*

To some extent the supplier performance data is collected automatically from around 50 suppliers that are included in an electronic call off system. When the goods arrive their bar code is scanned. Whenever there is a deviation this should be entered into the system, but *"we have not been able to make that work"*.

How is the supplier data stored?

The Company's purchasing system has a supplier evaluation tool where the supplier data is stored. It is not a standard ERP system, it is instead developed within the Company. This system was built in the late 1990s and was equipped with very advanced solutions like the possibility to scan bar codes in the call off system. The initial plan was to have a system that matched invoices to orders and was fully integrated in the operations; however *"it became too advanced"*. The system was built before the Internet was widely available and it was cumbersome to connect suppliers to the system. Currently 5-10 suppliers are connected to the integrated solution. The information is accessible to all employees within the Company. The suppliers do not have access to the system, but they should be informed about the results of evaluations. Today this is done sporadically.

This autumn the Company will roll out a new system, and it is looking into standard solutions, *"like all companies are"*. By finding a standard solution *"most of the suppliers could be connected to it"*. This should be compared with the number of suppliers (5-10) currently connected to the integrated solution. Since the 1990s the Company has matured and learned from the previous endeavors, so there is a strong belief that the new system will be a success. As an example, Internet facilitates easier access to the system. There is also a

realization that the system must be adaptable to different needs. Moreover, based on previous experience with manually putting in much data about suppliers, it is concluded that *"the systems must generate the metrics automatically"*. Focus is on connecting as many suppliers as possible and making the system user friendly, *"simplicity is in focus"*. In addition, all metrics will be reviewed and the focus will be on those metrics that directly translate to production costs.

How is the supplier data used?

It is difficult for the Company to draw as many conclusions as it wishes from its current system. *"If everyone used the system we would get statistics for our metric, but too few people really use the system"*. When encountering an error, workers often *"call the supplier to complain"* and fail to submit the data in the system. The same goes for the other metrics. The persons within the Company interacting with suppliers are requested to answer questions like *"how has the supplier's management performed"*, but this is often overlooked. The purchasing manager would like more data since *"statistics is facts"* and that would be useful when renewing a contract with the supplier.

A main tool for following up the relationships with suppliers is quarterly meetings with the largest suppliers. Here three representatives from each company meet and review the statistics and plan ahead. Moreover, each year all important suppliers are invited to a 'supplier's day' where ways to optimize the relationship are discussed. *"It can take almost a year to build a trust with a supplier and to match our needs with their capabilities."*

There are also internal meetings twice a year with the Company CEO and the regional managers, where internal goals for purchasing are set. As an example there is a goal for 2007 to purchase a certain amount from low-cost countries. The regional managers are measured against each other in a scorecard to create a competitive factor.

COMPANY M

Multinational company in automotive industry - Interview with two Supplier Quality Managers. This was a benchmark company.

Which suppliers are evaluated?

All suppliers of direct material are evaluated. It is thousands of suppliers.

What metrics are used?

Quality

PPM is the metric. Data about faulty goods are entered into the quality system and the supplier is automatically sent an email with the information. *"When there is a quality problem it is vital that the supplier is notified immediately. The description of the problem that is emailed to the supplier always includes a reference to the original drawing and a photo of the faulty part. This eliminates unnecessary discussions."* Each supplier's total PPM is calculated by the system every night so that deviations can be noticed quickly. Different categories and products have different target PPM depending on the complexity. *"Some have principally 0 PPM as target"*.

If a supplier delivers many faulty goods they are put on Control shipping. The first stage of this is that the supplier must sort its goods each day and send a report of the numbers of errors found, before they make each delivery. The second stage is that the Company tells a certified third-party company to visit the supplier and do the sorting at the supplier's expense.

The data is put into the supplier performance system by the quality department in the Production. There is almost no quality reception control. *"I as a Quality manager do not modify the data, I work with following up issues with suppliers. I also educate our suppliers on how to work with quality systems. This education is voluntary for suppliers."*

Logistics

The Logistics department also gathers a lot of data, and store it in the same system. 'Schedule adherence' is the metric that evaluates how well the suppliers keep their deliveries. There are also different sub-metrics. A very important sub-metric is 'Stock-outs', which is when the supplier is so late that the company has no goods in stock. The target is zero stock-outs. Another sub-metric is if the goods are correctly labeled or not.

Responsiveness

This is a subjective metric that grades the supplier on how responsive, reliable and flexible it is. *"This is very important, for example if we need 150 articles today because we are in a hurry, we must feel secure that the supplier can do it."* The grade is set by the responsible person working in logistics and it is the only subjective metric used in the ongoing performance measurements.

Financials

Recently the Company has started to look at the financial status of suppliers, to assess the risk of the supplier going out of business. Dun & Bradstreet's reports are used. *"Judging financial risk is very difficult. When we get unexpected large problems with a supplier, it is more often financial instabilities than quality problems."*

Price

The supplier's price reduction over time is evaluated. This is done by the Purchasing department.

How is supplier data stored?

The system is global and online. There are lists in the system of how all suppliers in all categories perform regarding all metrics, over time and currently. Few people in the Company can see all the data, due to its sensitivity. Quality people and Logistics people can not see the price data, for example. Suppliers can log in and see their own performance data. Moreover, the Company measures its own factories in the same way as it measures suppliers.

"The system is very easy to use since it is so automatic. One advantage with the daily updates is being able to quickly get a snapshot of a supplier and see what the current problems are. That the system is transparent is great because the supplier can never say that he did not know about a problem. It is also a service to the supplier that helps it improve. The system is very good, I see no disadvantages, except perhaps that it requires much technology. Before we had the system we did the same types of tasks in following up suppliers, only we did more on paper."

How is supplier data used?

"We only have global suppliers that supply all of our factories. If a supplier delivers faulty items to one of our factories this is visible immediately for everyone with access to the system. We have the possibility to escalate and put a supplier on new business hold. That gives us power." New business hold is used in case of big problems and it means that no new sourcing is done with the supplier for all the Company's factories. It is technically impossible to source from the supplier through the system. This is maintained for a couple of months while the supplier goes through a mandatory improvement program.

However, the Company go to great lengths to solve problems before considering the last measure, to abandon a relationship. *"We choose suppliers extremely carefully and invest very much time and money in our supplier relationships. If we throw out a supplier we have to do that process all over again. Problems have to be very big for us to do that. Normally we solve the problems. However, indirect material, like supplies, is another story, there we can change suppliers easily."*

In the system it is possible to break down a supplier's deliveries to each of the Company's different factories. *"If one region gets better performance from a*

supplier, why is that? Do we have different demands on the supplier? Does the supplier send us different goods? Those are interesting questions."

All suppliers are required to use the ISO/TS⁷⁵ standard. There are regular automatic notifications in the system when suppliers' ISO/TS certificates expire, so that the Quality manager can contact them about it.

Matrix for Comparison

The matrix

The matrix below is a graphical summary of the interviewed companies, in order to make comparisons easier.

The first rows in the matrix (from *industry* to *vertical integration in the value chain*) describe the companies' structure, the industries they are in, and the conditions in which they exist. This information is based on our general knowledge about the companies and the industries. It is not based directly on what was said during the interviews. The reason for filling in this data was in order to see if a company's supplier evaluation processes could be explained by its structure or by the market conditions.

The lower rows in the matrix (from *segmentation of suppliers* to *plans for improvement*) describe some key facts about how the companies work with supplier performance metrics. This information is directly taken from the interviews.

Finally, at the bottom of the matrix is the *purchasing maturity*, which is our concluding and subjective grade of how advanced a company's purchasing is, based on all information. (See the Word List)

Conclusions drawn from the matrix

We hypothesized that a number of background factors (*market and company background* in the matrix) could have some influence on the purchasing maturity in companies. We sorted the companies after their graded purchasing maturity and took a look at this. However these relationships were generally weak:

There was no significant correlation between purchasing maturity and the following factors: *Geographical coverage*: It does not seem that buying and selling many products internationally gives higher purchasing maturity. *Complexity of production*: The data does not support that manufacture of complex products leads to more mature purchasing, as one would perhaps initially suspect. *Variability of purchased products*: The data does not support that buying many different types of products implies more or less mature purchasing. *Place in value chain* and *Vertical integration in value chain*. One could hypothesize that a higher place in the value chain, and less vertical integration, increase the importance and maturity of purchasing in a company, but the data does not support this. *Number of suppliers*: That this relation is so weak is somewhat surprising, since many interviewed wanted to reduce the number of suppliers to increase their influence over them. The relationship to a high *Demand (legal or safety concerns) for fault-free products*, was slightly to the positive side. Such demands increase the cost of poor quality and emphasize the need for quality control of suppliers.

The more strong conclusions we could draw from the matrix, as well as the entire interview material, will be discussed in the next section: Analysis.

⁷⁵ A quality management standard common in the automotive industry.

Market and company background - estimations														
	Company A	Company B	Company C	Company D	Company E	Company F	Company G	Company H	Company I	Company J	Company K	Company L	Company M	
	Note: indirect material			Note: retail, no production										
Benchmark company?	benchmark	manufacturing	benchmark	manufacturing	manufacturing	benchmark	manufacturing	automotive	benchmark	manufacturing	manufacturing	benchmark	benchmark	
Industry	automotive	chemicals	chemicals	large	medium	retail	large	medium	large	medium	medium	construction	automotive	
Size	very large	large	very large	low	low	very large	low	medium	high	high	high	large	very large	
Geographical coverage	high	low	high	high	low	low	medium	high	medium	medium	medium	low	high	
Degree of purchasing centralization	very high	medium	medium-low	low	medium	medium	medium	low	high	medium	medium	high	very high	
Variability of products	medium	low	high	low	low	high	medium	low	medium	medium	medium-low	medium	medium	
Complexity of production	---	low	very high	low	low	high	medium	low	medium	medium	medium	medium	medium	
Variety of purchased products	very high	low	low	low	low	high	low	low	low	low	medium	medium	very high	
Demand (legal, safety) for fault-free products	middle	middle-low	high	middle	middle-low	high	middle	middle	middle	middle	middle	middle-high	high	
Place in value chain	low	low	high	low	low	medium	low	low	low	low	low	high	medium	
Vertical integration in value chain														
Usage of supplier data - Information from interviews														
Segmentation of suppliers	all are evaluated	suppliers accounting for 80% of purchases are evaluated, the rest are not	three groups of suppliers, the biggest are evaluated	three groups, different levels of evaluation	all are evaluated	majority are evaluated	all are evaluated	all are evaluated	strategic and non-strategic, all are evaluated	three groups, suppliers accounting for 80% of purchases are deeper evaluated	all are evaluated, suppliers accounting for 80% of purchases are deeper evaluated	three groups, not all suppliers evaluated	all are evaluated	
	very high, online integrated system	low, old ERP and Excel files	low, many different databases	medium, ERP and Excel files	low, old ERP and Excel files	medium, different databases	medium-high, online database but some data lacking	medium-high, Access database	high, online integrated system	low, old ERP and Excel files	medium, ERP and Access database	low, database but few suppliers connected	very high, online integrated system	
	1300	200	---	170	150	thousands	170	100	250	250	250	5000	thousands	
Number of direct material suppliers	PPM	no	many different	percentage rejects or PPM, for most	percentage rejects or PPM	many different	PPM	PPM	PPM	percentage rejects	percentage rejects	yes	PPM	
Quality metric	yes	yes, for most	partly	yes, perception	yes	yes	yes	yes	yes	yes	yes	yes	yes	
Delivery precision metric	Number of written complaints, Number of customer satisfaction for large suppliers, Commercial competitiveness	Some financial data for large suppliers	Lead time for service, Financial stability, Technological development				Number of corrective actions	no	no	in the deeper evaluation: Price development	Adherence to leadtime in the deeper evaluation: Price development, Lead time, Terms of delivery, agreement and payment	Call off time	Logistics, Financials, Price development	
Other objective metrics														
Subjective metrics	no	no	Competence, Willingness to help, Service	Communication, Quality, Price, Delivery precision	no	no	no	no	no	in deeper evaluation: Perceived quality, Collaboration, Technical level	in deeper evaluation: Service	Service and flexibility	Responsiveness	
Persistence of follow-ups on suppliers	very high	low	low	medium	low	high	medium	high	high	low	medium	low	very high	
Satisfaction with current system	very high	low	medium	high	---	medium	medium	high	high	low	medium-low	low	very high	
Plans for improvement	some new functions will be added	an ambition to improve	large recent changes, more improvements underway	no, satisfied	an ambition to improve	ideas, but no fixed plans	more functions will be added	no, due to migration of production	will standardize some more	new ERP and new routines to be implemented	perhaps simplify the system	will implement a whole new system soon	no, satisfied	
Overall purchasing maturity														
	very high	low	medium-low	medium	low	medium	medium	high	high	low	medium	low	very high	

Analysis

In this part of the report we discuss our main findings. We reflect on how important parts of the supplier performance measurement are carried out, what some main difficulties are, and how they ultimately affect the purchasing maturity of an organization.

Large Differences in Purchasing Maturity

One conclusion is that there are very large differences in purchasing maturity and ways of working between different industries, as well as between different companies in the manufacturing industry. A few of the interviewed companies had quite little experience in supplier performance measurement, gathering data about most suppliers somewhat unstructured in excel files. Some companies only recently had increase of the ratio of purchasing expenditure to turnover, and/or started to source in Low Cost Countries. In contrast, other companies had very clear opinions about how they wanted to work with suppliers, and used big dedicated databases for gathering supplier data.

The most advanced industry of those we visited is clearly the automotive industry. It was often mentioned by the interviewed persons as a good example. However, although being an example, reaching the same level of purchasing maturity and complexity as the automotive industry is not always attainable for smaller and/or more flexible industries or companies. This is because the automotive industry is special. It is centralized and highly controlled, plans its production far in advance, produces quite few product variants, and manufactures in large scale in large factories. This gives it a very high degree of exactness and standardization, something few companies in other industries can match. This relates to the concept that purchasing maturity is connected to the market, technology and competitive environment.⁷⁶

Another reason for the automotive firms' high purchasing maturity is that the level of competition is very high.⁷⁷ In other kinds of manufacturing firms the pressure has not always been so high, meaning that those firms can afford to be inefficient in their purchasing and still make a healthy profit. That naturally decreases the motive to work harder with purchasing.

Which Suppliers are Evaluated

Direct suppliers

To keep good track of one's suppliers, it is clear that the performance of all suppliers of direct material should be measured. The reason is that manufacturing companies strive to be very lean; consequently there are few extra parts in stock. This means that even relatively insignificant parts which are late or defect could stop the production. This is rather straightforward, but there were still a few companies in this study that did not evaluate all their suppliers, due to having a primitive data system for supplier evaluation or having too many suppliers.

Indirect suppliers

Indirect suppliers were not evaluated by any company, except for Company C, which was a special case. The consensus among the interviewed was that indirect material purchases were too small and too diverse to keep a meaningful track on. It would be difficult to measure quality on all the different categories like electricity, printers, cleaning services, and coffee

⁷⁶ Rozemeijer, van Weele & Weggeman, 2003

⁷⁷ Rozemeijer, van Weele & Weggeman, 2003

machines. In addition indirect material is not very complex, like paper and electricity, so normally not much can go wrong. If there was a problem with some indirect material it was dealt with directly, without going through a routine.

Segmentation

Moreover, almost everyone interviewed agreed that some suppliers require more energy than others, due to their size of their deliveries or complexity or strategic importance of the goods. This resulted in segmentation of suppliers so that important suppliers were supervised more carefully. For example, some companies had recurrent meetings with representatives from important suppliers, and some employed additional metrics for the important suppliers.

How often Evaluation Takes Place

Almost all of the companies interviewed keep a daily watch over deliveries and production through the people working with incoming deliveries and in the production. If there is a late delivery or damaged products it is normally detected right away and dealt with that day or a few days later. This is how things are dealt with on a daily basis in the running operations.

However the word *evaluation* here means a reflection at separate points in time on how the suppliers perform. For this, the normal time period used by the companies is one month. This means that supplier performance data is gathered continuously in the systems, but evaluated 12 months per year. The exceptions are the two most advanced companies (Company A and M), which both gather and evaluate supplier performance in real time in their large systems.

A quality manager commented on this common time frame for evaluation: *"You can generate a lot of data, but it is only worth it if someone is actually sitting down and digesting it. I would think generally speaking that one month is a good interval for following up suppliers. One month gives you plenty of time to capture a lot of data, and it does not mean that you must jump on problems immediately. You also have a natural break throughout the year."*

Many Metrics do not Equal Good Performance

An important conclusion from this study is that there is no correlation between using many metrics and having a well-functioning supplier performance system. From the companies surveyed it seems that to utilize a few metrics well and make useful analyses from them is more important than having many metrics that one does not fully utilize. This is supported by previous studies that describe how easy it is to fall for the temptation to introduce so many metrics that one cannot handle all of them.⁷⁸

In this case, Company K and Company J used many metrics but did not get the system to work perfectly. Company K had so many metrics that some of them were not actually used in practice. In contrast Company H and Company I used just a few metrics and made their supplier performance system work very well. The importance of thoroughly integrating metrics in a company is also emphasized by Neely et al.⁷⁹

Having many metrics means that there is more data to gather and to analyze. Since time is limited, using that data in a good way can be difficult. This difficulty can to some extent be alleviated by having a much automated system for gathering, however analyzing the data will still be done by humans. Almost everyone that was interviewed said that the two most important metrics were Quality and Delivery precision, so it seems that most companies are satisfied with evaluating their suppliers at the *product level*⁸⁰

⁷⁸ Paranjape, Rossiter & Pantano, 2006

⁷⁹ Neely, Gregory & Platts, 2005

⁸⁰ Van Weele, 2005

and working with *outcome measures*.⁸¹ Perhaps these two metrics are so much more important than other metrics that using just the two in a thorough way gives a company a good foundation. A company with low purchasing maturity had a similar view for developing their system, feeling that starting in a small scale and establishing routines for a few metrics was the right way to go. *"One must learn to count before starting to use a calculator"*, as the purchasing manager expressed it.

Subjective Metrics

Subjective metrics are those based on people's perceptions about a supplier, and not on numbers. According to Stevens, it is possible to measure such perceptions.⁸² Two examples of subjective metrics are 'how good the communication with a supplier is' and 'how responsible a supplier is when a fault is found among the goods'.

Six of the 13 companies use some kind of subjective metrics. This is not surprising as these kinds of metrics are common. For example, in the Aberdeen study, the metrics *service capability* and *compliance with contract terms* were commonly used.⁸³ However, it is somewhat surprising that in this study, using subjective metrics is negatively related to having a high purchasing maturity: these metrics are not common among the most mature companies, and many companies had problems with their subjective metrics. This relates back to the proposition that having many metrics do not equal good performance, but it is also a result of the special complications with subjective metrics.

Advantages of subjective metrics

Proponents of subjective metrics say that they capture a lot of interesting information which cannot be captured with numbers, that is, objective metrics. Objective metrics like the PPM metric can be considered imprecise since it only says 'correct' or 'incorrect'. There is no middle ground and no room for interpretation.

Disadvantages of subjective metrics

Many of those interviewed said that subjective metrics are difficult to define and compare. What is 'good' for one person might be 'acceptable' by another. One quality manager remarked: *"I can talk with ten different purchasers and get ten different opinions about how good a supplier is."* A purchasing manager commented that *"the end result of our subjective grades is often what one had expected anyway"*.

A category manager said that when one asks a company why it buys much from a certain supplier, it is easy for the company to use subjective metrics to motivate that the supplier fits the company's needs well, is flexible, and so on. *"If you ask people, the best supplier is the one a company buys most from. If we instead evaluate on objective metrics, such discussions will be much more transparent and the company has to show concrete evidence of that a supplier performs well."* Moreover, it takes time to collect data for subjective metrics since people have to take time to write or say what they think about suppliers. That information cannot be as automatically gathered as statistics, as for example PPM can be. Conclusively, it seems that the negative sides of subjective metrics overshadow the positive ones.

⁸¹ Parker, 2000

⁸² Stevens, 1946

⁸³ Aberdeen Group, 2005

Difficulties with Defining the Metrics

Metrics might seem rather straightforward, especially when studying literature in purchasing and management control. It is only a matter of making some definitions. However, metrics that seem simple are in fact quite complicated.⁸⁴ Below is a look into definition issues for the two most used metrics by the companies surveyed: Quality (PPM) and Delivery precision.

PPM

The definition of PPM is the number of rejected parts from a supplier divided by the total delivered parts from the supplier, times 1 000 000. The resulting figure will be the number of faulty parts per million parts. Faulty parts can be found both in incoming quality controls and in the production. Parts are generally faulty due to supplier errors, but it could also be a sign that one's own purchasing organization has sent a supplier incorrect or difficult-to-understand product specifications.

PPM	
$\frac{\text{Quantity of faulty parts}}{\text{Quantity of delivered parts}} \times 1\,000\,000$	

Our interviews made clear that it is surprisingly difficult to define which parts are faulty. Some companies had very unclear rules for PPM. As an example, when 1 faulty part is found in a batch of 10 items, should one check every part and return and count 1 item as faulty, or instead immediately return the entire batch of 10 and count 10 as faulty? And does that choice change if the batch instead contains 10 000 parts? (Sorting 10 000 parts is almost impossible.) Moreover, if we take a sample of 20 items from a batch of 100 000 items and 2 are found to be faulty, how should the PPM number be affected? One alternative is that all 100 000 parts are faulty. Another alternative is that the 2 found parts are faulty, and a third alternative is that $2/20 = 10\% = 10\,000$ parts are counted as faulty. Additionally, what if a supplier that has delivered faulty parts responds immediately to this by sending a person to the factory to sort the parts at their expense? Will those faulty parts be counted fully, despite the active work by the supplier?

Furthermore, if faulty parts are urgently needed in the Production and the faults are easy to fix, companies normally repair them to be able to meet production deadlines. How much modifications to these parts should be allowed, and should only these parts or the entire batch be counted as faulty in the statistics? One purchasing manager commented: *"Basically, if parts can be fixed, Production will use it, since they are concerned with completing their deliveries and will rather fix things and save time, than wait for new parts. So a policy to reject a whole batch when there is one error will be difficult to implement."* Since it is difficult to set rules for all these situations, most companies use common sense and traditions instead of fixed rules. Often it is up to the responsible person in quality control to make a decision as he sees fit.

If rules for these PPM calculations still are set, the interviewed persons said it is very difficult to uphold them in the midst of a hurried production process. Most people take the approach that achieving perfect adherence to rules is unrealistic: A quality manager from one of the companies with high purchasing maturity commented: *"Our PPM is not absolutely accurate, but it is accurate enough to detect any changes and give us information on what to improve. If we reject 900 parts or 1000 parts, it does not matter, if you get it down to 500 you know you have improved. As the data is accurate enough to do that, it can be relied upon. If you try to be totally accurate and want to know if you have rejected 179 parts, it will not work because there will always be one or two parts that slip through the net and go unnoticed."*

⁸⁴ Paranjape, Rossiter & Pantano, 2006

Delivery precision

The normal definition of this metric among the interviewed companies is the number of deliveries, in time and in full, by a supplier, divided by the total number of orders by the supplier, times 100. The resulting figure will be the percentage of correct deliveries by a supplier. This definition of delivery precision is sometimes called OTIF (on time in full). A delivery must be complete and in time to be correct.

Delivery precision	
$\frac{\text{\# of in time and in full deliveries}}{\text{\# of ordered deliveries}} \times 100$	

According to this definition the delivery is viewed as one item: it is either correct or it is not. It does not take into account how late a delivery is in number of days, or how many items that were missing in the delivery. A few of the surveyed companies did track the degree of lateness/severity, since they saw this as an important development of the delivery precision metric. One person commented: *"If a supplier has a delivery precision of 90 percent, and the delayed 10 percent is one day late, then we will hardly notice the delay. Another supplier could be better on paper by having 95 percent goods on time. However if the other 5 percent are 10 weeks delayed, we will have to stop the production."* Tracking these things is however much more difficult than just tracking if a delivery is late or not. One interviewed company had previously implemented a system for measuring the degrees of lateness, but abandoned it after a short time since it was too complicated.

Like with PPM, there are some definition problems with delivery precision. One disparity is each contractual agreement for the transportation. Some contracts state the delivery date as when it reaches the carrier and not the production site. For example, is the delivery OTIF if someone informs us that it has arrived on the right date to the carrier in China? No, since we do not know if it is in full yet. But what if it arrives on the right day, according to a schedule or an ERP, but still misses the transport boat which left early in the morning? What if the supplier delivers correctly, but the carrier delays the arrival to the factory? Moreover, raw material can be difficult to deliver in exact quantities. Therefore one company accepted errors of plus/minus 10 percent in the volume of for example steel.

A lot of this type of information is difficult to obtain since the logistics chains can be very long. It is easier to tell what date the goods arrive to the company's factory, and therefore some companies choose not to measure the delivery date to the carrier, and instead generally add an extra week for transport within Europe and two weeks for Asia. This reduces the precision of the delivery metric.

Another issue, although less significant, is when orders are made or changed in haste by the company so that the time frame to the delivery date is more narrow than the previously contracted lead time. Then the supplier is not given enough time and cannot be held responsible if the delivery does not take place on this 'extra early' date. Such deliveries are therefore counted differently.

A similar situation is when the supplier changes delivery date. If the new date is still within the contracted lead time it would normally be ok. However if the date is outside the lead time the delivery normally is counted as late, even though the supplier gave a heads up notice. This heads up notice is however good communication and could be awarded in other ways.

Different goals inside the company

One complication with introducing a strict definition of a metric is that different people have different goals. Managers want clear and strict rules. They would like to return every faulty product in order to get more easily interpreted statistics and to keep suppliers honest. On the other hand, people working in the Production process want to have flexibility to act after their

own judgment. There are large disparities between how urgent different situations are and how important different parts are, which also corresponds with the findings of Hervani & Helms of the difficulties of agreeing upon metrics.⁸⁵

Difficulties with Internal Compliance

Introducing or changing routines for supplier performance measurement is a big organizational change that involves many people. The soft issues are very important in purchasing initiatives.⁸⁶ This means that lack of internal cooperation and compliance is one of the biggest problems one will face as a purchasing manager wanting to drive a change process.

One interviewed person said that in his company there was a big difference in attitude between the management level and those on the 'factory floor'. He said that *"there was a lot of resistance"* when he previously participated in introducing a supplier performance measurement tool. Today four years later there is still internal resistance which results in faulty data. A purchasing manager in one of the largest companies surveyed echoed this sentiment, telling how employees would not put in information about faulty and missing products into the company's supplier database, even after many years of using the system. *"We are decentralized and there are so many different people involved. [...] A result is that we lack a lot of statistics for our metrics."*

A manager in yet another company said that the culture in that company was *"super conservative"* since many of the workers had been there for around 20 years. According to a fourth person one common problem with implementing supplier performance systems was that the Purchasing function did not have the highest status in the organization. People in the lower levels of an organization often resist reporting to and giving up control to centrally placed person like a purchasing manager. They want to keep their autonomy.

Furthermore, it seems that is not only the employees working in Production that resist: one person said that purchasers sometimes choose local suppliers out of convenience, and that the suppliers they choose is very much influenced by how they perceive that they fit together. This means that corporate directives about which suppliers are best to use are not always adhered to. An example from that company was that purchasers located in some countries always seemed to choose suppliers that were located close geographically even though, according to the interviewed person, there were also many good suppliers countries further away. Conclusively, it seems that due to these types of internal resistance, implementing a supplier performance system is considerably more difficult than designing one.

Another point is that the standards should be the same for those suppliers that are companies in the same organization (internal suppliers) as the standards for normal (external) suppliers. We were told that this could be a problem in a large organization. One person commented: *"They [the other Group company] do not consider us a 'real customer' since we are a sister company. Our delivery precision from our internal suppliers is much lower than for our external suppliers."* This type of different standard is obviously not optimal for efficiency, and that the supplier is an internal one severely decreases the ability to use sanctions to stop the problems.

However, not everyone recognized these kinds of internal compliance problems: a manager from one the most advanced companies, in the automotive industry, was of the opinion that once the company had made decisions to introduce more supplier performance measurements, everything happened smoothly and the employees conformed.

⁸⁵ Hervani & Helms, 2005

⁸⁶ Compare with Rozemeijer, van Weele & Weggeman, 2003

The Value of IT for Supplier Performance Measurement Systems

In order to collect and analyze data from the supplier metrics, a company needs a data system. This could be a collection of excel files with data about suppliers, it could be a stand-alone supplier performance system/database, or a system integrated into the company's ERP system.

A few of the studied companies had very old ERPs and therefore had difficulties getting all the data they wanted as well as aggregating the data into useful information. For example, one of these ERPs was from the 1980s. This resulted in the use of many excel files for gathering supplier data. With time these files accumulate in numbers which result in difficulties when retrieving and comparing information. In contrast to the old ERPs, some companies had modern systems that were continuously updated with information straight from the purchasing, quality and logistics departments as well as from the ERP system. With a click of a button, supplier information could be retrieved quickly and automatically on a computer screen. It did not seem to matter much whether the supplier performance measurement system was stand-alone or integrated with the ERP. What mattered was how automated it was in its data-gathering and how easy it was to use.

Another example of badly used IT tools was the experience of Company L. Their supplier evaluation system was too complicated to use: thus, despite the big investment only 5-10 suppliers decided to connect to the system. Company L's new system is being designed to be much more easy to use and standardized.

IT is both expensive and cumbersome to implement, therefore advanced supplier performance systems require large investments. This entails a large initial fixed cost, which might not be economically sound or feasible for small companies since future benefits to the purchasing function are variable, i.e. roughly based on the total purchased volume. One purchasing manager in a mid-sized company acknowledged his company's difficulties with its old ERP, but said that they had no intention of upgrading since that would cost a fortune. This is in line with Morgan's findings of about lacking resources in small enterprises.⁸⁷

However, given that a company has the necessary resources and scale, it became clear from our studies that a well-functioning data system was absolutely crucial for an efficient usage of supplier metrics. A Purchasing Manager at one of the most advanced (and large) companies commented: *"Previously we started a diverse range of supplier evaluations that were all stored in excel files. After a while the files became too cumbersome to handle [...] now we are happy that we invested in a good system. [...] You just choose whatever you want to look at. You can look at specific suppliers, time periods, production units, etc. It is easy to find out if which suppliers perform well and which do not. With this information it is easier to drive suppliers toward improvement. My conclusion is that a company that wants to have a good supplier follow up process should invest in good systems early on."*

As an illustration of the difference between an advanced and a basic IT system, compare the following two cases of a company that encounters a faulty product. Although slightly simplified, they are based on reality.

Technology use in company Alfa

Someone discovers the fault and enters information about it into an electronic database. This instantaneously affects the PPM levels for the supplier, which is in the same system. When the report is saved a copy is automatically sent to the supplier via email and the supplier is required to answer the report within a certain time, describing why the fault appeared and what actions it is taking to prevent it from happening again. The supplier's response to this report is

⁸⁷ Morgan, 2004

also tracked, and automatic reminders with warnings are sent out if the supplier is late in answering.

Technology use in company Beta

Someone discovers the fault and calls the supplier to tell them something is wrong. Information about the fault is entered into an excel sheet that stores quality data about only this supplier, making no note of how it performs compared to other suppliers. After a couple of days, the person who discovered the fault realizes that the supplier has not answered. The person must then look through a pile of different files stored on a local computer to remember exactly now again what the fault was and when the first report was sent.

Implications for the Host Organization

From the lessons learned throughout this work, many recommendations were given to the host organization, in a separate rapport. In extreme briefness, some recommendations are:

Since the host organization is composed of many different types of companies that work within somewhat different niches, each company should be allowed to gather its own types of supplier data. Some metrics might be extra essential for one company. Daily follow-up on supplier shortcomings is also handled by the separate companies. However, the organization should implement the use of a few simple and standardized metrics, common for everyone. This data is then aggregated to give a broad picture of the best suppliers, globally, in order to leverage the organization's size. The best way to gather such data would be through a common ERP, but since different ERPs are used, another, simpler, IT system must be built. To get the crucial user adoption, much internal informing and marketing as well as a thorough implementation process is needed.

Implications for Practitioners

During the work with this thesis we got some insights into supplier performance measurements, which invites to an opportunity to use the gathered knowledge to outline a few best practices for using supplier metrics:

Make clear definitions of metrics

As previously discussed, metrics are more complicated than one might think. For example, is it in some case ok to make a minor adjustment to a good that is slightly off the specification, and then use it and count it as fault-free? There is no easy answer and very little academic guidance. It is more important that a company actually makes a statement and decides upon a definition, than exactly what the definition looks like. When a common ground is set, statistics gathered by different people and business units can be compared. Clear definitions will facilitate the best practice suggestions that follow below.

Do careful follow-up of metrics, internally

One of the most important things with supplier performance metrics is to utilize them in their fullest extent. The data gathered must continuously be analyzed, otherwise it serves no purpose. In the words of one quality manager: *"data in itself is just bits of paper, which do not do anything."* This observation is similar to Rozemeijer et al's discussion about the uselessness of 'indigestible mountains of paper that the businesses are unwilling or unable to wade through'. Purchasing must be developed so that there is a competence to handle all information.^{88 89} Relating back to the conflicting interests between

⁸⁸ Rozemeijer, van Weele & Weggeman, 2003

different internal departments: a company can decide on using advanced metrics that give very interesting information. However, if this information is not properly understood, gathered or used by the employees, nothing will be gained. User adoption of the system is critical for success. As previously discussed, a problem for some of the interviewed companies was that they had many metrics but were not using them well. Using a wide variety of metrics should be reserved for companies that have both the necessary technological, organizational and managerial competence.

Do careful follow-up of metrics, in regards to suppliers

In addition, a company needs solid routines for following up suppliers according to the data. For example, suppliers can be told how they perform in relation to suppliers in the same category. One could for example say to a supplier: "With your PPM of 1000 you are ranked 3rd of 8 among the suppliers of cold-rolled steel." One purchasing manager told of an experience of being measured: *"We were part of a 'supplier of the year' contest arranged by one big automotive customer and our performance was compared to other suppliers anonymously. That had a pretty big mental effect, that someone gave us attention and measured us in that way."* A company also needs routines for strategic evaluations: for example monthly reviews of all suppliers' performance, or regular meetings with strategic suppliers.

Strive for simplicity in a system

For companies lacking advanced supplier performance measurement systems, what is the starting point when constructing one? According to our findings, perhaps the most important point is to make a system easy to use. User friendliness, ease of implementation, and automated routines were recurring themes. One person commented that *"if people need to put in extra resources to work with the system, they will resist it."* The main reason why employees sometimes choose not to support a system is because they perceive the disadvantages to be larger than the advantages. People need to see the advantages for them personally. In response to this, one way to increase usage of a system is to clearly communicate the benefits to everyone.

Use technology to improve measurement

It was clear that there was a relationship between companies with a high purchasing maturity and companies with a modern and technologically advanced system for gathering supplier performance data. Having a good system is crucial. Having information in excel files or on paper do not suffice. A data system that is not automated also increases resistance among employees toward using the system, since no one likes to manually fill in data.

Implications for Future Research

This master thesis has given a brief overview of supplier performance metrics. A theoretical angle we would like to emphasize is the complexity of *how*. I.e.: how are metrics utilized in practice when a firm encounters the problems of everyday operations? Setting up a metric is easy, the important part is making sure the metric is utilized and valuable for the purchasing department's decision making. This is a quite unexplored area within the field of measurement in purchasing, and there is ample room for further research. One fruitful approach for the future theorist might be to probe deeper into other academic streams that might have some connection with this area, for example production management, quality management, accounting internal control, or organizational internal control. Here one might for example find connections between supplier performance and departmental and organizational performance.

⁸⁹ Hughes, 2005

As for making a study, it would be interesting to further explore the differences between different industries. In that case, larger samples of respondents would be necessary. It could also be interesting to study different kinds of industries than those studied here. To get an even broader picture of this field, it would also be beneficial to interview the suppliers themselves. They are affected by metrics to a high degree.

Another approach, instead of going broader, would be to go deeper. This study is qualitative and its aim is to understand the process behind metrics. Still however, a few interviews cannot give a full understanding. This means it would be interesting to see if a very deep case study at a purchasing organization, performed under a longer time period, would yield other conclusions. Finally, it would be very interesting with a study of how different kinds of supplier performance metrics processes contribute financially to companies' purchasing departments. That is to really establish how much a company's success relies on its capability to measure its supplier's performance.

References

Aberdeen Group, *The Supplier Performance Measurement Benchmark Report*, September 2005, Retrieved through www.verticalnet.com.

Axelsson, B., Rozemeijer, F., Wynstra, F. (2005) *Developing Sourcing Capabilities*, Wiley

Axelsson, B. *Inköp en komplex funktion*, Edited by Samuelson, L. (2004), *Controllerhandboken*, Teknikföretagen.

Axelsson, B., Wynstra, F. (2002) *Buying Business services*, Wiley.

Axelsson, B., Laage-Hellman, J., Nilsson, U. (2002) *Modern management accounting for modern purchasing*, European journal of purchasing and supply management, vol 8, no 1, p 53-62.

Axelsson, B., Laage-Hellman, J., Nilsson, U. (1991) *Inköp, en ledningsfråga*, Stockholm.

Axelsson, B. Course material: *Purchasing Management 2109*. Text retrieved from Stockholm School of Economics Courseweb, October 21, 2007.

Barnett Y. (2006) *Reinforcement Theory*. Retrieved from: <http://www.referenceforbusiness.com/management/Pr-Sa/Reinforcement-Theory.html>, on October 21, 2007.

Beamon, B. (1999) *Measuring supply chain performance*, International journal of operations and production management, vol 19, no 3, p 275-292.

Blenkinsop, S., Davis, L. (1991) *The road to continuous improvement*, Insight, vol 4, no 3, p 23-26. (cited in Neely, A., Gergory, M., Platts, K. (2005) *Performance Measurement System Design: a literature review and research agenda*, International journal of operations and production management, vol 25, no 12, p 1228-1263.)

Chen, I., Parajuna, A. (2004) *Toward a theory of supply chain management: the constructs and measurements*, Journal of operations management 22, p 119-150.

Dalton, F. (2004) *Evidence-based performance measures*, Business Credit, vol 106, no 10, p 12.

Davenport, T. (1998) *Putting the enterprise into the enterprise system*, Harvard business review, Jul/Aug 1998, vol 76, issue 4, p 121-131.

De Waal, A. (2003) *The future of the balanced scorecard: an interview with professor dr Robert Kaplan*, Measuring business excellence, p 30.

Hervani, A., Helms, M., Sarkis, J. (2005) *Performance measurement for green supply chain management*, Benchmarking, 2005 vol 12, no 4, p 330.

Hughes, J. (2005) *Supplier metrics that matter*, CPO agenda autumn 2005.

Kaplan, R., Norton, D. (1996) *Using the Balanced Scorecard as a Strategic Management System*, Harvard Business Review, July-August 2007. Reprint from January-February 1996.

Merriam, S.B. (1994) *Fallstudien som forskningsmetod*, Studentlitteratur.

Morgan, C. (2004) *Structure, speed and salience: performance measurement in the supply chain*, Business process management journal, vol 10, no 5, p 522-536.

Neely, A., Gregory, M., Platts, K. (2005) *Performance Measurement System Design: a literature review and research agenda*, International journal of operations and production management, vol 25, no 12, p 1228-1263.

- Neely, A., Richards, H., Mills, J., Platts, K., Bourne, M. (1997) *Designing performance measures: a structured approach*, International journal of operations and production management, vol 17, no 11, p 1131-1152.
- Neely, A., Mills, J., Platts, K., Gregory, M., Richards, H. (1994) *Realizing strategy through measurement*, International journal of operations and production management, vol 14, no 3, p 140-152.
- Paranjape, B., Rossiter, M., Pantano, V. (2006) *Insights from the Balanced Scorecard Performance measurement systems: successes, failures and future - a review*, Measuring business excellence, vol 10, no 3, p 4-14.
- Parker, C. (2000) *Performance measurement*, Work study, vol 49, no 2, p 63-66.
- Rozemeijer, F., van Weele, A., Weggeman, M. (2003) *Creating corporate advantage through purchasing: toward a contingency model*, Journal of supply chain management, winter 2003.
- Schary, P., Skjøtt-Larsen, T. (2001) *Managing the global supply chain*, Copenhagen business school press.
- Shepherd, C., Günter, H. (2006) *Measuring supply chain performance: current research and future directions*, International journal of productivity and performance management, vol 55, no 3/4, p 242-258.
- Skinner, W. (2000) *Manufacturing - missing link in corporate strategy*, Harvard business school.
- Spekman, R., Kamauff, J., Spear, J. (1999) *Towards more effective sourcing and supplier management*, European journal of purchasing and supply management 5, p 103-116
- Stevens, S.S. (1946) *On the theory of scales of measurement*, Science, vol 103, no 2684, p 677-680.
- Stonich, P. (1981) *Using rewards in implementing strategy*, Strategic management journal, vol 2, p 345-352.
- Söderlund, M. (2005) *Mätningar och mått i marknadsundersökarens värld*, Liber.
- Tangen, S. (2005), *Improving the performance of a performance measure*, Measuring Business Excellence, vol 9 no 2, p 4-11.
- Trost, J. (2002) *Att skriva uppsats med akribi*, Studentlitteratur.
- University of Illinois at Urbana-Champaign, course material: *Internal Recruitment selection*, Text retrived from https://netfiles.uiuc.edu/zhang27/www/course/562slides/562_12.ppt, on November 24, 2007.
- van Weele, A. (2005) *Purchasing and Supply Chain Management*, Thomson.
- Yin, R.K. (1994) *Case Study and Research: Design and Methods*, Sage Publications: Thousand Oaks.

Appendix

The Question Template

Below are the questions we used in the interviews. The template was also modified over time and with regards to whether the interviewed company was a part of our host corporation or not.

Introduction

- Your title here is X, what does that mean?
 - How long have you worked here?
 - What did you work with before?
- How many at your company are involved in evaluating your suppliers?
- Which suppliers are your biggest?
 - Where are they located?

Supplier performance

- How do you evaluate your current suppliers?
 - What metrics do you use?
 - How is the measuring done, in practice, and by whom?
 - Is this way of working common for your entire corporation?
 - Are all suppliers evaluated in the same way, or are they divided into different groups?
 - Do the suppliers get feedback on what you measure? How do you follow up their performance?

Gathering data

- How is the supplier information gathered and then stored electronically?
 - How is it accessed, and by whom?
 - How is the data used and distributed?
 - How much time does that data gathering consume?

Results

- Does this system for supplier evaluation work well?
- What are the results? (reduction of number of suppliers, comments from the employees, usage by the employees, comments from suppliers, etc)
- How did you implement this system and how did you market it internally?

An optimal supplier performance system

- If you could design a new, optimal, system for gathering data and evaluating your suppliers, how would that look like?
 - How often should the suppliers be measured?
 - What would be the most important metrics to use and how would they be defined and measured in practice? (For example: quality, delivery precision, corporate culture, price, geographical position, innovativeness, certifications such as ISO, research capability, flexibility, etc, etc)
 - What would be the best way, in terms of systems and IT, to gather the data?
 - What is the best way to motivate suppliers to perform well? (carrot, stick, setting goals, etc)
 - If you would implement this kind of better system in your company, what would be the obstacles?