A meta-analysis of survey research on the relationship between Supply Chain Management and performance

Boyana N. Petkova*, Wout van Wezel\textsuperscript{1}

Department of Operations Management & Operations Research, Faculty of Economics and Business, University of Groningen, Landleven 5, 9747 AD Groningen, The Netherlands

* Corresponding author. Tel +31503636682
E-mail addresses: b.n.petkova@rug.nl (B.N.Petkova), w.m.c.van.wezel@rug.nl (W.M.C. van Wezel)
\textsuperscript{1}Tel: +31503637181

POMS 19th Annual Conference
La Jolla, California, U.S.A.
May 9 to May 12, 2008
Abstract
A substantial amount of surveys investigates the link between Supply Chain Management (SCM) and various performance measures. Although the survey papers add to our understanding, a quantitative generalization of the combined findings is lacking. Applying a meta-analysis using 247 correlations from 29 survey papers, this paper seeks generalizations on how SCM relates to performance. Hypotheses are distilled from literature on the relationship between different Supply Chain practices (e.g. information sharing, joint decision making) and various competitive capabilities (cost, speed, quality, etc.) and business performance measures (sales, market share, etc.). Results show that specific practices are positively linked to particular performance criteria yet that in many cases the benefit of SCM is highly conditional. Under certain circumstances, the outcome of SCM may even be harmful. The findings are discussed, new research opportunities are distilled and finally potential improvements on the meta-analysis are provided.

1. Introduction
Efficiency in the value chain originated as the core motivation for implementing cross-functional, supplier and buyer collaboration. Enforcing practices such as Just-In-Time manufacturing and Total Quality Management has enhanced operational coordination within the firm dramatically, and now companies start to focus on synergies between value chain links. Many practitioners and researchers (Dyer and Singh, 1998; Lockamy III and McCormack, 2004) believe that in the future the competition will shift from a battle between companies to entire supply chains competing against each other. Therefore it will be of importance to get insights into which SCM practices are successful under which circumstances.

Interest in partnering on an inter-firm level is increasing rapidly in both the service sector and product manufacturing across various industries and countries (Doran et al., 2005). Despite the fact that many studies show a positive correlation between SCM and performance (Devaraj et al., 2007; Frohlich and Westbrook, 2001; Frohlich and Westbrook, 2002; Li et al., 2005; Rai et al., 2006; Rosenzweig et al., 2003), its realization is still an issue within companies. Several studies demonstrate that awareness of modern SC planning and control tools is relatively high but that that their utilization level is relatively low (Olhager and Selldin, 2003), or that the SCM system is merely used for order management but not for inventory, distribution, or logistics management (Hsu, 2005; Lange and O'Driscoll, 2006). A logical reason for the slow adoption of SCM might be that we cannot provide practitioners with an all-round understanding of the drivers of successful SC practice yet. While there is a will to manage the supply chain, inability to determine the right form of collaboration leads to suboptimal logistics and sales performance, and negatively impacts outcomes for both the manufacturers and the retailers.

This lack of understanding SCM has serious implications for business. The Food and Beverages sector e.g. suffers routinely from excess inventory in the supply chain and empty store shelves (Cox, 2007). In the Chemical sector, supply chains are unquestionably valuable to performance, yet they are far from achieving best practice (Lange and O'Driscoll, 2006). Fisher (1997, p. 105-106) confirms that the performance of many supply chains has never been worse and costs have risen to unprecedented levels due to adverse relations and dysfunctional industry practices. Many companies stress that
the enormous waste in supply chains causes frustration among customers but they don’t know how to solve their partner coordination problems.

The importance to the business field of coordinating the value chain activities has not eluded researchers. In the last couple of years the amount of publications in Supply Chain Collaboration has grown exponentially (Burgess et al., 2006). Most of the research conducted in SCM is empirical-descriptive (Croom et al., 2000) and does not actively research context. While this research still provides us great insights in what is going on in business, we need to warrant that we don’t get caught up in adjusting our perception of the usefulness of SCM to the current market frenzy. According to Cox (2001) most practitioners fail to improve performance in supply chains because they are benchmarking and copying what others are doing. To overcome this, several researchers propose a contingency approach to SCM (Bagchi and Skjoett-Larsen, 2003; Bask and Juga, 2001; Benton and Maloni, 2005; Fisher, 1997). However, given the current lack of all-round knowledge on the relation between SCM and performance, their standpoint is theoretical or explorative.

Application of research findings is still challenging for practitioners and there is a need to bundle empirical and analytical outcomes into clear guidelines. In order to build on the current literature and facilitate further research, it is important to come to a generalization of existing empirical results. In this paper, we will make first steps towards an empirically founded contingency approach by generalizing the findings of existing surveys that relate SCM to performance.

In order to reduce the sample size to a realistic number, the scope of a survey is often explicitly limited by restricting the subjects (e.g., by country or region, company size, and/or the kind of product) and the variables that are tested. For example, Humphreys et al. (2004) investigate the relation between supplier development and overall competitive capabilities in electronic manufacturing companies in China, and, consequentially, not other kinds of SC mechanisms and performance objectives in a different business context. A quantitative meta-analysis can be used to generalize the findings of multiple surveys into a ‘general truth’. In this article, we describe the results of a meta-analysis on SCM surveys in two steps. First, we establish the basic elements that play a role in the relation between SCM and performance, resulting in a comprehensive model relating the type of partnering in the SC (e.g. Information Sharing) to several kinds of performance (e.g., speed, flexibility, cost). Second, we will investigate if results are contingent, by analyzing the existence and value of moderators on the relation between SCM practices and business performance. Concluding, we will research the following questions:

- Which SCM types are correlated with which performance measures?
- Is the correlation between SCM and performance moderated?

This paper contributes to existing literature by providing conceptual insight in the link between SCM and performance based on the large amount of empirical data already reported by others.

The structure of the paper is as follows. In section two, we will give an overview of SCM, measures for performance and our research framework. The meta-analysis methodology will be explained in section three. Results of the meta-analysis can be found
in section four. Finally, conclusions, limitations and recommendations for further research are displayed in section five.

2. Research model

2.1 Introduction
The massive flow of results that have been produced in SCM over the last decade is in need for a generalization. Replication studies are an acknowledged way to ensure generalizability, validity and reliability of theory development (Hubbard and Vetter, 1996; Hunter, 2001; Singh et al., 2003; Tsang and Kai-Man Kwan, 1999) but very few papers of this kind are published in SCM. Several very interesting narrative overviews of SCM have appeared (Burgess et al., 2006; Giannakis and Croom, 2004; van der Vaart and van Donk, 2007) yet the subjectivity of this qualitative approach may reduce its impact. Meta-analysis (MA) is a tool to combine the findings and to determine conditionality in the main effect in a more objective manner. Furthermore, areas for future survey research which will be particularly useful for our understanding of SCM can be identified. We believe that a meta-analysis of the effect between SCM and performance is in order. First, it is a field with growing interest and it has matured enough and produced a sufficient body of literature to perform a MA. Second, no MA has ever appeared about this topic.

While a lot of meaningful and very interesting additions have been made, the multidisciplinary nature of Supply Chain research has lead to problems in bundling the findings. Depending on the research angle, different constructs and performance measures for SCM are used. Like in Total Quality Management (Nair, 2006), in Supply Chain Management the borders between the organization and its environment may be considered fluid. Because of this lack of boundaries and the explosive growth of publications in the field, the nature of SCM has not been clearly defined by a set of measurement constructs. Instead, many different definitions of what SCM comprehends have been used and a variety of constructs and measures thereof have been utilized to measure them. However, to be able to perform a meta-analysis, a common ground is needed to compare the survey results. In this section, we propose a comprehensive model for SCM based on existing conceptual and empirical literature. We describe respectively the design parameters for SCM and the kinds of business performance used to measure their impact. As the success of SCM can depend on contextual factors, we also depict apparent moderators. The section ends with the hypotheses which we will test in the meta-analysis.

2.2 SCM design parameters
The design parameters which we consider in the meta-analysis need to both lead to a comprehensive model with sufficient detail and need to be sufficiently available in the surveys in the analysis. This latter implies (as we will explain in the next section where the meta-analysis technique is discussed) that parameters must be apparent in at least two surveys with at least three observations in total in order to be able to generalize research findings. Therefore, we are forced to group the many variables of SCM and performance found in surveys in similar categories.
Several generic categorizations of supply chain management can be found in literature. For example, Simatupang and Sridharan (2005) categorize SCM into decision areas (the design of a collaborative performance system, information sharing, decision synchronization, incentive alignment, and integrated supply chain processes). Similarly, Croom et al. (2000) distinguish Assets, Information, Knowledge and Relationships as possible elements of exchange in Supply Chain Management. Sánchez-Rodríguez (2005) combine intensity of integration with the kind of integrative activities performed by distinguishing basic (assessment of suppliers), moderate (tactical integration) and advanced (information exchange and collaboration) integration. Van der Vaart and Van Donk (2004) identify the transparency (low information sharing), commitment/coordination (high information sharing), and integrative planning (centralized decision making) stages as successive integration phases. Finally, Swaminathan et al. (1998) distinguish configuration (network structure of the supply chain), coordination (operational and transaction issues like inventory control and information exchange), and contracts (e.g. supplier selection).

To cover the various elements related to supply chain management in this meta-analysis, we propose to distinguish (a) design from operational aspects (like in Simchi-Levi et al. (2008)), and (b) the intensity of integration in successive phases (showing the incremental nature of SCM which is emphasized in the majority of the classifications mentioned above). Six clearly distinguishable types of SCM can then be placed within this framework (Figure 1).

Literature reports several levels of integration in design aspects. First, there are one-sided Supplier Selection methods which comprise a minimum of interaction between firms. These procedures don’t intend to influence each other’s decision making but merely select the best fit for a company from a static set of external options. Main constructs from literature are Specifying assessment criteria for suppliers (Carr and Pearson, 1999; Flynn and Flynn, 2005; Humphreys et al., 2004) and Reducing the number of suppliers (Chen et al., 2004; Chen and Paulraj, 2004b). Increasing market competition by organizing public auctions for suppliers is proposed by Johnson et al. (2007) as a third one-sided method but this construct has not gained momentum in the empirical research community yet. One-sided mechanisms are different from and might precede a higher level of involvement. Mutual integration of companies purposely alters the decision making process of both parties.

The remaining SCM categories on a design level state mutual intentions about Development, Incentive Alignment, and/or Supply Chain Design. Development consists of the joint creation of knowledge in order to improve each other’s products (Chen and Paulraj, 2004b; Chow et al., 2006; Devaraj et al., 2007; Flynn and Flynn, 2005) or processes (Chen and Paulraj, 2004b; Chow et al., 2006; Humphreys et al., 2004; Johnson et al., 2007). Incentive alignment captures the supply chain coordination through the identification, formulation and enforcement of joint targets (Chow et al., 2006; Devaraj et al., 2007). Supply Chain Design is the most extreme form of collaboration between companies. Companies merge their existing and new assets to build a configuration that best suits demand (Simatupang and Sridharan, 2005; Swaminathan et al., 1998).

Operational coordination in Supply Chains can be achieved by two means. The first way is to provide Information about a firm’s operations to which a partner is expected to react in a manner that is positive for the Information releasing firm. Joint
Decision Making (e.g. through Vendor Managed Inventory, Setting joint production plans etc.) is the second operational Integration method and requires commitment of both parties for mutually creating value. These mechanisms have been elaborately included in empirical (Bagchi et al., 2005; Kulp et al., 2004; Moberg and Speh, 2003; Petkova et al., 2008; Salvador et al., 2001; Van Donk et al., 2007) and theoretical papers (e.g. (Bernstein and Federgruen, 2007; Cachon, 2001; Hall and Potts, 2003; Netessine and Rudi, 2003).

By categorizing the proposed SCM types in a two dimensional matrix which covers the level of decision making (operational versus tactical/strategic) and commitment intensity of parties, we are providing a structured overview of the resulting framework (Figure 1). This classification into six types has its limitations as well in the context of this meta-analysis. First, the amount of empirical work on incentive system design is slim. Furthermore, researchers tend to mix (very) high commitment SCM types in one construct. This is logical for Joint Decision Making and Supply Chain Design because when resources are integrated, it is obvious that firms need to e.g. make joint production plans. The distinction between Incentive alignment and Supply Chain Design might also be a bit blurry as they most probably are part of a package deal. The other SCM practices are also thought to have a logical build-up with increasing commitment of the parties involved, yet they don’t seem as intertwined as the high commitment SCM types. For purposes of our meta-analysis, we have therefore combined the high/very high commitment dimensions Joint decision making, Incentive alignment, Supply Chain design into a category Organizational Integration. Hereby, sufficient observations per category are achieved.

Figure 1: Six forms of partnering in Supply Chain Management

To complete the framework for our meta-analysis by specifying the dependent variables, we will discuss the kinds of performance that are reported in SCM research in the next section.

2.3 SCM and Business performance
As with the SCM parameters, the performance measures that we consider in the meta-analysis need to be comprehensive, detailed enough to be useful, but also sufficiently available in the surveys. Regarding performance, both financial measures and competitive capabilities are popular performance measures. However, there is a debate
about which type should be used in SCM. LaCrosse et al. (2002) demonstrate that Return on Net Worth is a function of net profit, asset turnover and financial leverage which can be applied to logistics performance measurement. Others on the other hand prefer to use competitive capabilities (Li et al., 2006; Morris and Carter, 2005; Paulraj and Chen, 2007; Tan et al., 1999). Tan et al. (1999) argue that financial data measures such as return on assets, return on investment, or internal rate of return are typically based on accounting data and do not reflect the time value of money and opportunity costs. There might furthermore be a time lag between the realization of aggregate performance through competitive capabilities. In addition, the link between competitive capabilities and financial performance is conditional upon business conditions (Fisher, 1997; Zhou and Benton, 2007). While Cost is the most distinguishing competitive capability when the level of environmental uncertainty is low, Quality, Speed, and Flexibility become determinant competitive capabilities when business conditions become complex. Most SCM papers report on the link with both competitive capabilities and financial constructs and discuss the coherence between the performance measures separately. Gunasekaran et al. (2001) argue that there is a lack of balanced approach and a lack of clear distinction between metrics at strategic, tactical and operational levels. Measures for the management of the integrated supply chain need to facilitate strategies and actions plans. Therefore, we need to ensure that performance measures are distinct and balanced over different levels. A meaningful measure of firm performance should therefore include both the competitive capabilities as the financial performance measures. As a result, we decided that there is not one single ‘right’ measure and included the entire pallet of performance indicators: financial performance measures, overall and individual competitive capabilities.

Common measures which are used within Supply Chain Management research to assess the ability of a company to earn a profit are return on assets (Kannan and Tan, 2003; Rosenzweig et al., 2003), return on investment (Byrd and Davidson, 2003; Vickery et al., 2003), return on equity (Byrd and Davidson, 2003), return on sales (Sanchez Martinez and Perez Perez, 2005; Vickery et al., 2003), sales (Kim et al., 2005; Tan et al., 1999), market share (Kannan and Tan, 2003; Kim et al., 2005), and profit margins (Kulp et al., 2004). In this meta-analysis, we have combined all financial performance measures into one construct in order to achieve sufficient measurements.

The competitive capabilities identified in SCM literature match those defined by Slack et al. (2001): Quality (Bagchi et al., 2005; Cousins and Menguc, 2006; Paulraj and Chen, 2007; Petkova et al., 2008; Zailani and Rajagopal, 2005), Cost (Flynn and Flynn, 2005; Gough, 2003; Kim et al., 2005; Kulp et al., 2004; Li et al., 2005; Petkova et al., 2008; Scannell et al., 2000), Flexibility (Chen and Paulraj, 2004a; Gough, 2003; Sanchez Martinez and Perez Perez, 2005), Dependability (Bagchi et al., 2005; Salvador et al., 2001; Van Donk et al., 2007), and Speed (Flynn and Flynn, 2005; Lockamy III and McCormack, 2004; Morris and Carter, 2005). An additional performance dimension related to Customer Service has been added by OM-researchers (Baiman et al., 2001; Cousins and Menguc, 2006; Paulraj and Chen, 2007; Tan et al., 1999; Zailani and Rajagopal, 2005). For this meta-analysis, we decided to merge the competitive capabilities Speed and Flexibility as their operationalizations resembled each other a lot. Thus, we could test for Quality, Cost, Speed/Flexibility, Dependability, and Customer Service.
Some papers (Kannan and Tan, 2002; Kannan and Tan, 2003; Kannan and Tan, 2004; Tan et al., 1999) measure the impact of SCM on several competitive capabilities combined (e.g. Cost and Quality and Speed). This was captured through the concept Overall Competitive Capabilities in our meta-analysis.

All survey papers in SCM use subjective appraisals of performance with exception of Li et al. (2005). Most measurements of performance are obtained by asking opinions of the responsible top level executives or managers of the functional areas. While most performance results are (implicitly) reported on a firm level, sometimes internal performance results of departments are mentioned (Sánchez-Rodríguez et al., 2005). Besides, performance can be measured by benchmarking with prior own firm performance or comparing results against those of competitors. Another difference that can occur in expressing performance is either in terms of growth (e.g. % increase in sales) or in an absolute number (e.g. sales in $). A third kind of relative performance measure is compared to set goals (Sanders and Premus, 2005). This variety in the performance measurement is not an issue in this meta-analysis: “Mixing apples and oranges is not problematic, as long as your interest is fruit” (Glass et al., 1981).

2.4 Conditionality due to context

As mentioned, various researchers claim that more integration with customers and suppliers will lead to better performance. However, practitioners are not as confident about the benefits. Attempts to implement SCM in the 1990s have not been universally successful and have failed to deliver the desired results in many cases (Tan et al., 1999). Scientific evidence is present (Devaraj et al., 2007) that involving both upstream and downstream value chain links does not always lead to better results. Recent publications attribute this to environmental conditionality. Although some research framework propositions have been made (Simatupang and Sridharan, 2005; Soonhong Min and Mentzer, 2004; van der Vaart and van Donk, 2007), there is no consensus yet on which moderating/mediating factors should be included in the SCM-performance framework and how the actual relationships between them are. The identification of moderating factors is both a powerful feature and a necessity in meta-analysis (Cortina, 2003). We will discuss two kinds of factors that might play a role in the success of implementing SCM: factors that are determined by the subjects in the study, and factors that relate to the methodological setup of the survey.

Most recent work focuses on industry related moderation due to market dynamics. Significant differences between industries have been reported and difference in uncertainty is provided as a potential reason (Sanjay and Ravi, 2006). Other substantial factors relate to e.g. culture (Hsu, 2005; Lin et al., 2005; Zailani and Rajagopal, 2005), the type of partner (Droge et al., 2004; Jayaram et al., 2004; Kulp et al., 2004; Narasimhan and Kim, 2002; Sanders and Premus, 2005; Vickery et al., 2003), or business process maturity and structure of the integrating firm (Flynn and Flynn, 2005; Johnson et al., 2007; Lockamy III and McCormack, 2004; Pagell, 2004).

Another category of influential moderators inherent to the way that the research is set up. While many papers are based on a mix of industries, a substantial amount of authors very understandably tend to investigate industries where they expected to find the most advanced examples of SCM (Automobile and FMCG). Findings can therefore not be extended to industries with less partnering. The measurement of SCM and
performance should also be compared on its reliability which depends on whether single or multiple constructs are used for both concepts. In OM, many researchers still prefer to use single measures or they use composite measures but don’t report all relevant statistics. Examples of other research related moderators are respondent type, data source (own data or database), publication source (e.g. Journal or Working paper) and year of data collection.

2.4 Hypotheses
The research goal of this paper is to achieve insights into the relationship between SCM and performance. First, we want to test if the effect is always positive, which is a presumption that has been made often as described earlier. Optimally, we wish to contribute to answering the research question of which form of SCM increases business performance under which conditions. To this goal, the following hypotheses have been formulated:

**H1:** SCM will have a positive effect on performance.

**H2:** The relationship between SCM and performance is influenced by moderating factors.

Figure 2 expresses the research model, in which we wish to find the relation between SCM and performance and by what factors this relation is moderated.

![Figure 2: Research model](image)

In the next section, we describe the procedures we followed to answer these hypotheses.

3 Meta-analytic procedures

3.1 Literature search strategy and coding
To collect a sample of results for the meta-analysis, we performed an Electronic Database search with the Boolean expression “Supply chain” (Subject) and “performance” (Abstract) and “survey” (Abstract) in Ebscohost. All results have been examined carefully and only the papers that provide quantitative data of the link between SCM and performance have been considered. Furthermore, we have added two recent papers from the Euroma conference proceedings that have not yet been published in journals.
Descriptives of the final selection of 29 papers are shown in Table 1. This sample size compares favorably with other meta-analyses in Operations Management (Nair, 2006: 23; Gerwin and Barrowman, 2002: 26). To extract the correlations which were needed as input in the meta-analytic procedures, we developed a coding scheme which was used by one of the authors to classify all papers. To control for researcher bias, the other author of this paper and a scientist that is not involved in SCM research also coded several papers. Their coding results were consistent with that of the first author. The checklist for descriptive data of effect sizes recommended by Rosenthal (1995) has been used to present the final data set.

3.2 Publication bias

Every overview paper carries the risk of only including a certain kind of research findings. This ‘file drawer problem’ (Rosenthal, 1979) occurs due to the tendency of e.g. non-significant or non-main stream results to get published with a lower probability. Problems arise from solely taking into consideration the significant, generally acknowledged results in a meta-analysis (Hunter and Schmidt, 1990b). Context dependency of the effect size may falsely cause a zero-order correlation. A widely used remedy is to include different effect sizes without taking into account their significance levels (Hunter and Schmidt, 1990b). When the same data source is used in several papers (either own data gathered or a database) this might bias the results as well. In our paper selection no multiple uses of the same database for the same SCM-performance pairs occurred so this was not problematic.

In addition, it is a convention in Business and Economics to rely on individual studies which are much more biased than their combined results. In narrative overviews the existence of publication bias of published versus unpublished results is also hardly ever addressed. The final argument for not caring too much about publication bias, is that even if it exists and we accommodate it (e.g. by two-stage Heckman-selection models), the results appear to be highly similar to non-weighted models. Similar to the use of papers based on the same data, the issue publication bias due to reporting several effect sizes per paper arose. Although an argument can be made to include only one effect size per paper (Rosenthal, 1991), Bijmolt and Pieters (2001) show through Monte Carlo Simulation that procedures that use the full set of reported measures outperform those which average within studies. Therefore, we opted to use several reported measures from the same paper.

If sufficient non-published results are available, the best approach to identify a publication bias is to add these proceedings, working papers, or dissertations to the published paper set and test for a moderator effect. Unfortunately in SCM very few sources of this kind are available. Only about two out of 28 papers in our total data set of the MA consist of non-journal papers. Possibly, this is due to the fact that SCM is very popular and empirical results are well publishable. The papers that have not been published yet are being prepared by the authors for journal submission.

The formal test for publication bias regresses the effect size on its precision (Stanley, 2005). Only few publications report the standard errors which makes the test unsuitable for this MA. Two other publication bias tests can be applied to our data. First, the Funnel Plot is a graphical examination of the correlations. The effect size found in each study is plotted against its sample size. If there is symmetry around a certain
correlation value and if studies with higher sample sizes are close to this value while studies with lower sample sizes exhibit more variation, then there is no heterogeneity in the published versus the non-published results (Vevea and Woods, 2005). Based on the funnel plot of all correlations, we have no reason to assume that a publication bias exists as the data points are scattered fairly symmetric and no skew exists. The effect sizes display a large spread but this could very well be caused by substantial moderating factors. As the number of observations is lower for the individual SCM types, Funnel Plots are not suited as an evaluation method. Instead, we have performed the file-drawer test (Rosenthal, 1995) to test for publication bias. The file-drawer test assesses the amount of zero effects \( k_0 \) that are necessary in order to make the findings non-significant. As many Z-values (or p-values) are not reported, we have taken the effect sizes to be just significant at \( \alpha=5\% \) (\( Z=1.96 \)). The necessary number of zero-publications found for the 247 SCM practices largely exceeds the minimal recommended amount of zero effects by Rosenthal (1991) \( 5k+10 \) (59,456>>1035). Caution is advised on the result of this test. Even if we would have failed it, this would not have meant that a meta-analysis cannot be performed: drawing effect sizes from a random sample of economic papers might have produced a similar effect size distribution. In our case, we suspect ‘real’ moderating effects due to difference in the effectiveness of individual SCM practices on different performance measures. Therefore, these cases might conceptually not be suited to be added up (although this is done by for instance Frohlich and Westbrook (2001)). The quantitative tests are not well suited to be performed for the individual dimensions which have fewer observations. For example, Nair (2006) and Gerwin and Barrowman (2002) do not address publication bias at all. In our case, therefore the best way to ensure the absence of publishing bias was by a careful review of all possible publications.

![Figure 3: Effect size distribution for SCM Flow Practices](image-url)
3.4 Meta-analytic procedures

Data Analysis is performed in two steps. First, all findings need to be translated to the same effect size measure. Equations to recalculate common effect sizes can be found in Rosenthal (1995), Cooper and Hedges (1994) and Fern and Monroe (1996). Alternatively the Z-score and effect sign can be computed from the p-value. The resulting set of effect sizes needs to be approximately normally distributed as this an assumption in MA. In Operations Management, usually the correlation coefficient (Pearson’s r) or beta coefficient (Multiple least squares) are used to measure the effect size. We are interested in the one-on-one relationship and will therefore use Pearson’s r. If the beta coefficients are provided, we attempted to recalculate the correlation coefficient as applied by Nair (2006):

“The correlation coefficients were obtained by using the basic equation $R_{YX}R_{XX}^{-1}=\beta$ where $R_{YX} = [R_{YX1} \ R_{YX2} \ ...]$ is the correlation matrix of the independent and dependent variable; $R_{XX}^{-1} = [\begin{array}{cc} R_{x1,x1} & R_{x1,x2} \\ R_{x2,x1} & R_{x2,x2} \end{array}]$ is the correlation matrix of the independent variables and $\beta$ is the regression coefficient.” Correction for bivariate versus multivariate measurements is very much desirable but not necessary in MA (Peterson and Brown, 2005). When the use of r-values is not possible, standardized (beta) coefficients could also be used. But if it is possible to correct for multivariate effects this should definitely be implemented.

The observed effect size can differ from the true effect due to 11 study artifacts (e.g. measurement error, dichotomization, range variation) which can be corrected for
(Hunter and Schmidt, 2004). If this should be done is doubtful because as Rosenthal (1991) states “It does not strike me as the proper goal of a meta-analysis. The goal is to teach us better what is, not what might someday be in the best of all possible worlds when all our independent and dependent variables are perfectly measured, perfectly valid, perfectly continuous, and perfectly unrestricted in range”. Many authors thus choose not to include any corrections in MAs. On the other hand, there is a significant stream influenced by Hunter and Schmidt (1990) who believe that they can approximate the ‘general truth’ better by correcting. Papers related to Operations Management (Gerwin and Barrowman, 2002; Nair, 2006) try at least to correct for reliability (measurement error) of SCM and performance. Yet, as information is lacking in their source papers, the authors mentioned apply averaging of correlations between independent and dependent variables and calculate Cronbach’s alpha from the correlations between the independent variables. Averaging leads to worse results than keeping all values (Bijmolt and Pieters, 2001) which is a valid argument not to attempt to produce reliability measures yourself by this technique. From a practical point of view, next to the validity argument that can be made against of corrections, gathering the necessary information from SCM papers would be rather impossible due to very restricted reporting in most studies. There is not even enough information to perform the averaging method of Nair (2006) and Gerwin and Barrowman (2002). Therefore, we opted to use the Bare Bones approach (Hunter and Schmidt, 1990a) where corrections are only performed for differences in sample size.

Correlations between -1 and +1 are non-normal and therefore the Fisher transformation was used to make the data closer to a normal distribution:

$$z_r = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right) \quad \text{with} \quad s.e.(z_r) = \frac{1}{\sqrt{n-3}}$$

After the analysis, the findings were then back-transformed to obtain the final results.

To calculate the 95% confidence interval for the mean, we used the procedure by Lipsey and Wilson (2001): CI=average correlation +/- 1.96*SE. The standard error is calculated as the inverse of the sum of all study weights. The study weights equal the number of papers minus 3 for the Fisher transformation. Even if the average effect size calculated with the standard error is positive and lies in a rather narrow range, this does not mean that we can conclude that SCM is beneficial in all cases. Whether negative effects can occur should also be evaluated with the credibility interval which refers to the distribution of population means found. If the credibility interval includes zero, then the generalizability of the positive effect found through the confidence interval is not warranted and moderator effects should be investigated.

Calculation of the credibility interval is as follows. From the corrected estimate of sampling error variability $S_e^2$ and the variance of the corrected sample correlation $S^2_r$, the estimate for the population variance $S_p^2$ is obtained: $S_p^2 = S_r^2 - S_e^2$. The credibility interval is calculated similarly to the confidence interval but with the $S_p$ instead of the standard error. We use 1.96 $S_p$ which makes results significant at precisely 5% while others like Nair (2006) and Gerwin and Barrowman (2002) opt to use heuristics (RATIO1) which use a width of two $S_p$ as an approximation.

The randomness of variation can give clues whether the effect sizes come from a single population or if moderators influence the results. It can be assessed with the 75%
rule, the Q-statistic, and the I\(^2\) statistic. The 75% rule (Hunter and Schmidt, 1990a),
equals the heuristic called RATIO2 which is used by Nair (2006) and Gerwin and
Barrowman (2002): \( \text{RATIO2} = \frac{S^2_e}{S^2_{	ext{total}}} \). It assesses whether 75% or more of the variance of
the corrected sample correlation are caused by structural error. If this is the case, then no
moderators exist.

Another way to assess homogeneity of effect sizes is the Q-statistic. It can be used
to assess the spread of the real values around their mean:
\( Q = \sum_{i=1}^{k} w_i (T_i - \hat{T})^2 \). The degree
of homogeneity can then be calculated with the \( I^2 \) index which reflects the percentage of
total variability in a set of effect sizes due to true heterogeneity:
\( I^2 = \frac{(Q - (k - 1)) * 100%}{Q} \)

for \( Q > k - 1 \) and otherwise \( I^2 = 0 \). In the last case, all variability can be ascribed to sampling
errors. An \( I^2 \) value of 0.25, 0.5 and 0.75 can be interpreted as low, medium and high
heterogeneity (Huedo-Medina et al., 2006).

Testing for homogeneity has low power when the number of observed effects is
low and when the decision made is binary. We have included all three measures of
heterogeneity to ensure stability of our findings.

4. Results
For all 247 effect sizes based on all SCM types and performance measures combined, the
overall effect is positive with an average of 0.158 and 95% confidence interval between
0.149 and 0.167. The credibility interval indicates that not all effect sizes are expected to
be positive (lower bound = -0.107). High levels of moderation are perceived. Q equals
1003.23 at 246 degrees of freedom which means that with high significance (p=0.0000)
the null hypothesis that the population is homogeneous can be rejected. The \( I^2 \) statistic of
Huedo-Medina et al (2006) displays the value of 0.7547 which can be interpreted as high
levels of heterogeneity. The 75% condition for homogeneity is not met as well.

This heterogeneity is a valid reason to split our data according to performance
measure and SCM type.

Individual measures
As Table 2 shows, sufficient observations exist to perform a meta-analysis for many
combinations (we have taken \( k > 1 \) and \( n > 2 \) as a minimum requirement which is
comparable to Nair (2006)). All effect sizes have a significantly positive mean. Although
all effect size averages are positive, indications for moderators exist for over 50% of the
cases (11 out of 21 cases where sufficient observations existed) and sometimes even
negative effects can be expected due to moderation. We will discuss these homogeneity
issues per SCM type.
<table>
<thead>
<tr>
<th></th>
<th>K</th>
<th>N</th>
<th>participants</th>
<th>Min ES</th>
<th>Max ES</th>
<th>proportion</th>
<th>Ave ES</th>
<th>SD ES</th>
<th>SE</th>
<th>Confidence interval (95%)</th>
<th>Confi dence interval (95%)</th>
<th>G</th>
<th>p-value</th>
<th>Q</th>
<th>df</th>
<th>RAIT02</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td>26</td>
<td>24</td>
<td>49327</td>
<td>0.21</td>
<td>0.24</td>
<td>0.27</td>
<td>0.28</td>
<td>0.19</td>
<td>0.17</td>
<td>0.42</td>
<td></td>
<td></td>
<td>0.157412</td>
</tr>
<tr>
<td><strong>Financial Performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.17</td>
<td>0.42</td>
<td></td>
<td></td>
<td>0.157412</td>
</tr>
<tr>
<td>Supplier Selection</td>
<td>4</td>
<td>2</td>
<td>833</td>
<td>0.28</td>
<td>0.45</td>
<td>1.00</td>
<td>0.35</td>
<td>0.06</td>
<td>0.03</td>
<td>0.28</td>
<td>0.40</td>
<td>0.23</td>
<td>0.44</td>
<td>0.876757</td>
<td>0.0508</td>
<td>0.614578</td>
</tr>
<tr>
<td>Development</td>
<td>3</td>
<td>4</td>
<td>850</td>
<td>0.21</td>
<td>0.36</td>
<td>1.00</td>
<td>0.30</td>
<td>0.02</td>
<td>0.03</td>
<td>0.24</td>
<td>0.36</td>
<td>0.26</td>
<td>0.43</td>
<td>0.318318</td>
<td>0.2294</td>
<td>0.306504</td>
</tr>
<tr>
<td>Information Exchange</td>
<td>4</td>
<td>1</td>
<td>615</td>
<td>0.28</td>
<td>0.43</td>
<td>1.00</td>
<td>0.35</td>
<td>0.00</td>
<td>0.04</td>
<td>0.28</td>
<td>0.42</td>
<td>0.35</td>
<td>0.35</td>
<td>0.379538</td>
<td>0.2652</td>
<td>0.209318</td>
</tr>
<tr>
<td>Organizational Integration</td>
<td>2</td>
<td>2</td>
<td>391</td>
<td>0.26</td>
<td>0.28</td>
<td>1.00</td>
<td>0.27</td>
<td>0.00</td>
<td>0.05</td>
<td>0.27</td>
<td>0.36</td>
<td>0.27</td>
<td>0.333504</td>
<td>0.6550</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall Competitive Capabilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.17</td>
<td>0.42</td>
<td></td>
<td></td>
<td>0.157412</td>
</tr>
<tr>
<td>Supplier Selection</td>
<td>4</td>
<td>2</td>
<td>833</td>
<td>0.28</td>
<td>0.45</td>
<td>1.00</td>
<td>0.35</td>
<td>0.06</td>
<td>0.03</td>
<td>0.28</td>
<td>0.40</td>
<td>0.23</td>
<td>0.44</td>
<td>0.876757</td>
<td>0.0508</td>
<td>0.614578</td>
</tr>
<tr>
<td>Development</td>
<td>3</td>
<td>4</td>
<td>850</td>
<td>0.21</td>
<td>0.36</td>
<td>1.00</td>
<td>0.30</td>
<td>0.02</td>
<td>0.03</td>
<td>0.24</td>
<td>0.36</td>
<td>0.26</td>
<td>0.43</td>
<td>0.318318</td>
<td>0.2294</td>
<td>0.306504</td>
</tr>
<tr>
<td>Information Exchange</td>
<td>4</td>
<td>1</td>
<td>615</td>
<td>0.28</td>
<td>0.43</td>
<td>1.00</td>
<td>0.35</td>
<td>0.00</td>
<td>0.04</td>
<td>0.28</td>
<td>0.42</td>
<td>0.35</td>
<td>0.35</td>
<td>0.379538</td>
<td>0.2652</td>
<td>0.209318</td>
</tr>
<tr>
<td>Organizational Integration</td>
<td>2</td>
<td>2</td>
<td>391</td>
<td>0.26</td>
<td>0.28</td>
<td>1.00</td>
<td>0.27</td>
<td>0.00</td>
<td>0.05</td>
<td>0.27</td>
<td>0.36</td>
<td>0.27</td>
<td>0.333504</td>
<td>0.6550</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Meta-analytic results (combinations with insufficient observations are marked yellow)
Supplier Selection
The effect of Supplier Selection could only be assessed for Financial Performance, Overall Competitive Capabilities, and Quality. Especially for Quality, Supplier Selection has relatively low levels of heterogeneity (I2 statistic = 0.37) or its effect might perhaps be homogeneous (75% rule=0.77 and Q not significant at 5%).

The credibility interval for Supplier Selection does not include zero and therefore it is very likely (95%) that Supplier Selection positively affects Quality under all circumstances. Financial Performance and Overall Competitive Capabilities could also be deemed as homogeneous according to the Q statistic (p=0.0611 and p=0.0506) but the 75% rule does not exceed the critical value and therefore more research on the presence of moderators would be valuable. The confidence intervals (both not including zero) suggest that Supplier Selection positively affects performance under all circumstances.

Development
Input in the development of each other’s products and processes does not show a uniform result for moderation. Financial Performance has a medium level of moderation and the credibility interval includes zero. For Overall Competitive Capabilities, Development enhances performance equally in all cases. If the individual competitive capabilities are investigated, then the statement of the positive benefit of Development holds for Speed/Flexibility and Dependability in all cases. Customer Service displays an insignificant Q value but the effect does not pass the 75% rule test to state conclusively that the effect is homogeneous. The lower credibility bound exceeds zero and therefore although some moderation might exist, the net effect will always be positive.

Quality and Cost are strongly moderated (Q is highly significant, I2=0.75 respectively 0.84, 75% rule value <0.75) and might even assume negative values (lower credibility bounds <0). Moderation effects of these dimensions might be opposite for varying competitive environments. Firms who pursue different competitive capabilities, should therefore be aware of the consequences of Development of/with Suppliers on their most important performance dimensions.

Information Sharing
Information Sharing displays the same effects on an aggregate performance level as Development. While the effect on Financial Performance is moderated and can become negative (Q highly significant, I2=0.90, 75% rule value <0.75, lower credibility bound <0), Overall Competitive Capabilities show positive effects with adequate homogeneity (Q p-value=0.28, I2=0.20, 75% rule value =1.06, lower credibility bound >0). Quality could not be tested as insufficient observations exist. We found no evidence of moderation on Cost, in fact all indicators showed that the relationship was highly homogeneous and positive (Q p-value=0.41, I2=0, 75% rule value =40, lower credibility bound >0). As only two papers and four observations were used for this assessment, we would suggest that more work is performed before we can conclude that the effect on Cost is not moderated. Speed/Flexibility, Dependability, and Customer Service are all moderated (medium to high levels) and could assume negative values.
**Organizational Integration**

The effect of Organizational Integration could not be assessed on Overall Competitive Capabilities and Customer Service. For Financial Performance, there is again evidence of moderation (Q p-value=0.0014, I²=0.62, 75% rule value =0.45, lower credibility bound<0). Similar to Information Sharing, Speed/Flexibility and Dependability are moderated for Organizational Integration. The results for Quality and Cost are ambiguous regarding homogeneity (Q not significant, I²=0.03 respectively 0.32 but 75% rule value<0.75) but as the effect means might become negative (lower credibility bound<0), more research should be conducted in the mechanisms driving successful Organizational Integration.

5. Conclusion, limitations and further research directions

5.1 Conclusion

No single statement can be made about hypotheses H1 and H2. Whether the hypotheses should be accepted or rejected depends highly on the SCM-performance pair under investigation. It seems that for all heterogeneous cases, the level of moderation is so severe that negative effect sizes can be expected under certain circumstances. Therefore, if there is moderation and Hypothesis 2 is thus accepted, then Hypothesis H1 will be rejected.

Three important conclusions can be formulated from the results of this meta-analysis.

First, while in many studies different Supply Chain Management types are added to achieve one mega-construct, we found heterogeneity in the total of 247 SCM correlations. If we split the entire sample up into specific pairs of SCM type and performance measure, then about half of the combinations are still moderated. The fact that so much moderation can be explained simply by the SCM-performance combination implies that large issues exist in adding up the different SCM types and generalizing across performance measures which has been a common practice in many SCM papers thus far.

Furthermore, the specific combinations display different mean effects (e.g. for Dependability the mean effect of Information Sharing is 0.12 and Development it is 0.26). Depending on the goal of a company, the implementation of the some SCM practices should be prioritized over others. There has been little reflection on the practical ability of companies to implement various SCM types at once. By conducting empirical research at a lower level of aggregation in the future and providing guidelines on the best order, we might meet practitioner demand for a more practical understanding of SCM.

Third, although the general impression is that more partnering is better, SCM is definitely not always beneficial and might even contribute negatively to performance. An overview of the individual results is provided in Table 3. Supplier Selection is the only SCM practice that seems beneficial for all performance measures. For the remaining SCM types, the picture is not quite clear for both the sign of the effect sizes found and if they are moderated or not. The effects of SCM on the competitive capabilities are ambiguous: negative effects through moderation can again occur. But as firms strive for a mix of competitive capabilities, the effects of different performance dimensions are combined. In that case, the comforting conclusion can be drawn that taken as a whole the
The effect of the individual SCM practices on overall competitive capabilities seems to be positive. Although Supply Chains are managed in a way that is far from optimal (Fisher, 1997), on the whole it is a good idea for firms to put effort in SCM. Although firms might be a bit confused about how to manage their supply chain, overall they seem to engage in partnering that is beneficial for their overall competitive position.

When the overall competitive capabilities need to be leveraged to financial performance, moderating factors play a role yet again. One possible explanation for this is that the ability to build strong relationships, to create goodwill, etc., might be a skill that partners value just as much as or more than the realization of actual competitive capabilities (elaborated on in section 5.3).

<table>
<thead>
<tr>
<th>Supplier Selection</th>
<th>Mean</th>
<th>Distribution</th>
<th>Moderation</th>
<th>75% rule</th>
<th>Q</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Performance</td>
<td>+</td>
<td>+</td>
<td>yes</td>
<td>no*</td>
<td>low-medium</td>
<td></td>
</tr>
<tr>
<td>Overall Competitive Capabilities</td>
<td>+</td>
<td>+</td>
<td>yes</td>
<td>no*</td>
<td>medium-high</td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td>+</td>
<td>+</td>
<td>no</td>
<td>no</td>
<td>low-medium</td>
<td></td>
</tr>
</tbody>
</table>

* = significant at alpha = 10%

<table>
<thead>
<tr>
<th>Development</th>
<th>Mean</th>
<th>Distribution</th>
<th>Moderation</th>
<th>75% rule</th>
<th>Q</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Performance</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>medium</td>
<td></td>
</tr>
<tr>
<td>Overall Competitive Capabilities</td>
<td>+</td>
<td>+</td>
<td>no</td>
<td>no</td>
<td>low</td>
<td></td>
</tr>
<tr>
<td>Speed/Flexibility</td>
<td>+</td>
<td>+</td>
<td>no</td>
<td>no</td>
<td>very low</td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>high</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>very high</td>
<td></td>
</tr>
<tr>
<td>Dependability</td>
<td>+</td>
<td>+</td>
<td>yes**</td>
<td>no</td>
<td>medium</td>
<td></td>
</tr>
<tr>
<td>Customer Service</td>
<td>+</td>
<td>+</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
</tbody>
</table>

** = close to cut-off value of 0.75

<table>
<thead>
<tr>
<th>Information</th>
<th>Mean</th>
<th>Distribution</th>
<th>Moderation</th>
<th>75% rule</th>
<th>Q</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Performance</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>very high</td>
<td></td>
</tr>
<tr>
<td>Overall Competitive Capabilities</td>
<td>+</td>
<td>+</td>
<td>no</td>
<td>no</td>
<td>low</td>
<td></td>
</tr>
<tr>
<td>Speed/Flexibility</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>high</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>+</td>
<td>+</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Dependability</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>medium</td>
<td></td>
</tr>
<tr>
<td>Customer Service</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>very high</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organizational Integration</th>
<th>Mean</th>
<th>Distribution</th>
<th>Moderation</th>
<th>75% rule</th>
<th>Q</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Performance</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>medium-high</td>
<td></td>
</tr>
<tr>
<td>Speed/Flexibility</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>yes</td>
<td>very high</td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td>+</td>
<td>+</td>
<td>yes</td>
<td>no</td>
<td>very low</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>+</td>
<td>-</td>
<td>yes</td>
<td>no</td>
<td>low-medium</td>
<td></td>
</tr>
<tr>
<td>Dependability</td>
<td>+</td>
<td>+</td>
<td>yes</td>
<td>yes</td>
<td>medium-high</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Overview of the effect of SCM practices

5.2 Future research directions
Several future research directions can be considered.

As a starter, we need much more empirical research on Supply Chain Management. There was insufficient literature to test all combinations of SCM practices.
and performance measures. For some pairs that could be tested, very little data exists and more empirical papers would be in order.

Even if we sort out the effect sizes into SCM-performance pairs, for halve of the combinations strong evidence of the existence of moderators is present. Many theoretical arguments have been made for the influence of environmental dynamics on SCM and performance. In the next version of this meta-analysis, we will research the influence of the three substantive moderators addressed in Research propositions R1-R3. Sufficient data and range therein exist in our survey paper selection to test for these moderators.

**R1:** The involvement of Suppliers-only (without Buyers, distribution partners, etc.) moderates the effect of SCM on performance.

**R2:** The cultural region in which businesses operate moderates the effect of SCM on performance.

**R3:** The industry in businesses operate moderates the effect of SCM on performance.

In future empiric work on SCM and performance, specific moderators like culture, industry, etc., should be studied explicitly. When more moderators are included, it will be possible to identify combinations of moderator variables where the largest improvements in theory can be made (Farley *et al.*, 1995).

We advise future survey research papers to include the relevant Pearson’s r correlation matrices in order to provide better information for future meta-analyses. Besides uniformity in reporting, we should strive to balance the level of detail in Supply Chain Management research. The applicability of findings might benefit from a more narrow definition of concepts. On the other hand it is disorderly to measure all possible constructs of SCM types and performance measures individually. Moreover, SCM types respectively performance measures correlate within their concept and their combined effect should be investigated on the level of detail that is best suited for the research question at hand. The six types of SCM which we proposed are well founded in literature. Other combinations are also possible and the achievement of an isomorphic conceptualization of SCM types might be a collective priority so as to enable the future accumulation of knowledge in SCM.

### 5.3 Limitations

Our research has two limitations. First, none of the artifacts mentioned by Hunter and Schmidt (1990) have been accounted for. When the purpose of a MA is to summarize work done in the field, corrections should not be applied. Contrary, if the goal is to find a generalizable truth, then it could be advisable to correct for study artifacts. It depends on the field which corrections are applied. In marketing e.g. the original studies are used uncorrected. In the MAs in OM (Gerwin and Barrowman, 2002; Nair, 2006), the measurement error correction $\rho_{xy}^{corrected} = \frac{\rho_{xy}^{observed}}{\sqrt{r_x} \sqrt{r_y}}$ is applied. Besides the fact that theoretical arguments against the application of corrections can be made, in our study correcting for the Hunter and Schmidt artifacts was simply impossible due to a lack of information in the original papers. When a more uniform view on the use of single or multiple constructs is developed in Supply Chain Management reliability corrections as indicated by Nair (2006) and Gerwin and Barrowman (2002) will be possible.

Second, the interaction effect between types of SCM cannot be accounted for in this meta-analysis but should be included in future empirical work. The SCM practices in
our framework seem to interrelate in a way that has not been fully explored yet. Researchers suggest an incremental approach to SCM with increasing commitment of parties in the later stages (Sánchez-Rodríguez et al., 2005; van der Vaart and van Donk, 2004). Effective Information Sharing is generally viewed as a necessity for Joint Decision Making. By enhancing the accuracy and speed of information IT solutions for example might increase the quality of collaboration (Frohlich and Westbrook, 2002; Sanders and Premus, 2005). Other indicators that affect the quality of joint decision making are data consistency and cross-functional application integration (Rai et al., 2006) and the level of collaborative planning (Jhakharia and Shankar, 2007). Further work on the disposition of SCM practices compared to each other is necessary. Moreover, in this paper we have focused on SCM practices which concern the management of the production flow only. In Marketing and increasingly in Operations Management literature, the value of Relationship Management and underlying concepts such as Trust, Communication, Goodwill, etc., has been explored in a business-to-business context. A correlation between relational and production flow management factors has been widely found (Bagchi et al., 2005; Benton and Maloni, 2005; Chwen et al., 2006; Cousins and Menguc, 2006; Jhakharia and Shankar, 2007; Kulp et al., 2004; Prahinski and Benton, 2004; Sheu et al., 2006). More conceptual and empirical research, perhaps building on findings from Marketing literature (Geyskens et al., 1999; Geyskens et al., 2006), is necessary to construct an overall SCM-framework which comprises both relational and production flow factors.
### Table 1: Overview of papers used in MA

<table>
<thead>
<tr>
<th>Paper</th>
<th>Source</th>
<th>Sample (number of firms), size and unit of analysis</th>
<th>Region/ Country</th>
<th>Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagchi e.a. (2005)</td>
<td>IJLM</td>
<td>149 firms with at least 100 people and 10 M$ revenue; respondents were mixed: there was no specific person targeted within the companies</td>
<td>Europe</td>
<td>Mixed: 9 types where authors expected to find most advanced collaboration</td>
</tr>
<tr>
<td>Carr and Pearson (1999)</td>
<td>JOM</td>
<td>739 firms, members of the National Association of Purchasing Management, respondents were executives at Director and Vice President level</td>
<td>USA</td>
<td>Mixed</td>
</tr>
<tr>
<td>Chen and Paulraj (2004)</td>
<td>JOM</td>
<td>221 firms; all medium to large (36% had more than 100 employees and 60% had a gross income of over 100M$); respondents were mixed</td>
<td>USA</td>
<td>Mixed (SIC 34-39)</td>
</tr>
<tr>
<td>Chen et. al. (2004)</td>
<td>JOM</td>
<td>221 firms; all medium to large (36% had more than 100 employees and 60% had a gross income of over 100M$); respondents were mixed</td>
<td>USA</td>
<td>Mixed (SIC 34-39)</td>
</tr>
<tr>
<td>Chow et. al. (2006)</td>
<td>Omega</td>
<td>101 US firms with middle line managers enrolled in MBA; 109 firms of Taiwan’s top manufacturing firms; respondents and size not reported</td>
<td>US and Taiwan separately</td>
<td>Mixed</td>
</tr>
<tr>
<td>Cousins and Menguc (2006)</td>
<td>JOM</td>
<td>142 firms; size not reported; respondents are the highest ranking supply persons</td>
<td>UK</td>
<td>Mixed: 11 types</td>
</tr>
<tr>
<td>Devaraj et. al. (2007)</td>
<td>JOM</td>
<td>120 firms, size varies but over half had fewer than 250 employees; Respondents are senior managers of functional areas</td>
<td>?</td>
<td>Mixed (SIC 357, 367, 369, 371): four high tech industries but half is from automotive industry</td>
</tr>
<tr>
<td>Flynn and Flynn (2005)</td>
<td>IJPR</td>
<td>164 firms, size not reported, respondents mixed from laborers to managers</td>
<td>Globally mixed (USA, Germany, Italy, Japan, and UK)</td>
<td>Mixed Electronics, transportation components, machinery industries</td>
</tr>
<tr>
<td>Frohlich and Westbrook (2002)</td>
<td>JOM</td>
<td>187 manufacturers and 298 service firms, size not reported; respondents were senior managers (usually VPs of Operations or General Managers)</td>
<td>UK</td>
<td>Mixed: &gt;10 industries</td>
</tr>
<tr>
<td>Humphreys et. al. (2004)</td>
<td>Omega</td>
<td>142 companies randomly selected from the Directory of Electronics Industry; respondents were purchasing directors, size ranged from 50 to over 3000 employees and nearly half had gross sales of 100M$ or more.</td>
<td>China</td>
<td>Electronic manufacturing companies</td>
</tr>
<tr>
<td>Johnson et. al. (2007)</td>
<td>JOM</td>
<td>284 firms of which 92.2% had annual sales &gt; 500 M$; respondents were primarily CP or director of purchasing or supply chain management</td>
<td>US and Canada</td>
<td>Mixed: manufacturing and service firms</td>
</tr>
<tr>
<td>Authors</td>
<td>Journal/Source</td>
<td>Sample Description</td>
<td>Country</td>
<td>Industry or Sector</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>---------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Kulp et al. (2004)</td>
<td>MS</td>
<td>54 divisions of manufacturers; size not reported; respondents were Senior</td>
<td>USA?</td>
<td>Food and consumer packaged goods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Information/ Collaboration executives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li et al. (2005)</td>
<td>JOM</td>
<td>196 firms with &gt;100 employees, respondents were top level executives or managers of</td>
<td>Globally mixed?</td>
<td>Mixed: SIC 25, 30, 34, 35, 36, 37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the functional areas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lin et al. (2005)</td>
<td>IJPE</td>
<td>109 firms in Hong Kong, 103 in Taiwan; size not reported; most respondents were</td>
<td>Taiwan and Hong Kong separately</td>
<td>Mixed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>managers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moberg et al. (2004)</td>
<td>IJLM</td>
<td>46 firms who were members of the Council of Logistics Management; size not</td>
<td>USA?</td>
<td>Food and consumer packaged goods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>reported; respondents were logistics executives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narinsahim and Kim</td>
<td>JOM</td>
<td>244 firms in Korea and 379 in Japan; mean sales were 428 M$ for Korean and</td>
<td>Korea and Japan</td>
<td>Mixed: large manufacturing firms</td>
</tr>
<tr>
<td></td>
<td></td>
<td>958M$ for Japan; respondents were supply chain managers or top-level executives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peterson et al. (2005)</td>
<td>JOM</td>
<td>134 firms; average annual sales ranged from 3M$ to 160M$; respondents not provided.</td>
<td>Globally mixed</td>
<td>Mixed: both manufacturing and non-manufacturing</td>
</tr>
<tr>
<td>Petkova et al. (2008)</td>
<td>Euroma</td>
<td>457 firms; size ranges from &lt;50 employees to over 2000; respondents are Director of</td>
<td>Europe</td>
<td>Mixed: SIC codes 28-35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Operations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rai et al. (2006)</td>
<td>MIS Quarterly</td>
<td>110 members of Council of Logistics Management; median firm size was 4000</td>
<td>USA?</td>
<td>mixed: manufacturing (SIC 20-39) or retail (SIC 52-59)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>employees and median firm revenue was 1.5B$; respondents were usually managers of</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>logistics, distribution, SC or IT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosenzweig et al.</td>
<td>JOM</td>
<td>238 companies of the VIM database, Companies differ in size from &lt;100 and &lt;50M$ to</td>
<td>Globally mixed (35 countries)</td>
<td>Consumer products</td>
</tr>
<tr>
<td>(2003)</td>
<td></td>
<td>&gt;10000 employees and &gt;5B$ revenue; respondents were senior manufacturing executives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salvador et al. (2001)</td>
<td>IJOPM</td>
<td>164 plants with at least 100 employees; respondents were very diverse</td>
<td>Germany, Japan, UK, USA</td>
<td>Mixed: Electronic, machinery, transportation equipment</td>
</tr>
<tr>
<td>Sanchez-Rodriguez et</td>
<td>SCM</td>
<td>306 firms from Dun and Bradstreet database, average had 141M euro sales and 568</td>
<td>Spain</td>
<td>Mixed: 13 industries</td>
</tr>
<tr>
<td>al. (2005)</td>
<td></td>
<td>employees; respondents were mixed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sanders (2007)</td>
<td>JOM</td>
<td>245 firms; mean sales range from 4M$ to 59M$ with 3,810-121,000 employees;</td>
<td>USA</td>
<td>Mixed: &gt;15 industries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>respondents were mixed but usually higher executives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scannell et al. (2000)</td>
<td>JBL</td>
<td>57 suppliers of Big Three firms; average annual sales 501M$ with average sale to</td>
<td>USA</td>
<td>Automobile</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Big Three 83.67%; respondents were CEOs of companies or SBUs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stank et al. (2001)</td>
<td>JBL</td>
<td>306 members of the CLM; firm size is not reported; respondents were mixed</td>
<td>Canada, Mexico, USA</td>
<td>Mixed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>executives and managers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanley and Wisner</td>
<td>JOM</td>
<td>118 firms who were members of the National Association of Purchasing</td>
<td>USA</td>
<td>Mixed: &gt;16 industries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Management (NAPM); annual revenue varied from &lt; 25M$ to &gt;500M$; respondents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Journal</td>
<td>Study Design</td>
<td>Sample Details</td>
<td>Location</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>--------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>2001</td>
<td>IJOPM</td>
<td>Tan e.a.</td>
<td>313 firms from the American Society of Quality Control; the number of employees varied from 12 to 256,000; annual sales ranged from 1M$ to 65B$; respondents were quality directors and vice-presidents</td>
<td>USA</td>
</tr>
<tr>
<td>1999</td>
<td>IJOPM</td>
<td>Van Donk e.a.</td>
<td>144 firms; size not reported; respondents were general manager, production manager, or logistics manager</td>
<td>Spain and The Netherlands</td>
</tr>
<tr>
<td>2003</td>
<td>JOM</td>
<td>Vickery e.a.</td>
<td>57 firms which were top 150 (in terms of annual sales) independently owned first tier suppliers to General Motors, Ford and Chrysler; respondents were CEOs of SBU</td>
<td>North America</td>
</tr>
</tbody>
</table>
References


