In search of supplier flexibility performance measurement
Helena Forslund
School of Business and Economics, Linnaeus University, Växjö, Sweden and
School of Engineering, Jönköping University, Jönköping, Sweden, and
Stig-Arne Mattsson
School of Engineering, Jönköping University, Jönköping, Sweden

Abstract
Purpose – The purpose of this study is to identify, characterize and assess supplier flexibility measurement practices in the order-to-delivery process.
Design/methodology/approach – The study involved a survey; participants were 224 purchasing managers at Swedish manufacturing companies that had more than 20 employees.
Findings – Scrutiny of the details of measurement practices revealed that most respondents actually do not specifically measure supplier flexibility. Instead they measure other measures like delivery reliability, conduct qualitative follow-ups, or cannot specify how supplier flexibility is measured. It was acknowledged that they measure different supplier flexibility aspects, and the applied measures were characterized, e.g. in terms of which flexibility dimension they represent.
Research limitations/implications – Conceptual clarifications and adaptations to measuring supplier flexibility in the order-to-delivery process are provided. The identified measures can be a contribution in further developing literature on flexibility performance measurement.
Practical implications – Purchasing, logistics and supply chain managers in search of supplier flexibility performance measurement can find ways to measure and an extended flexibility vocabulary. This has the potential to improve flexibility in the supply chain.
Originality/value – Even though flexibility is claimed as being an important competitive advantage, few empirical studies and operationalized measures exist, particularly in the order-to-delivery process.
Keywords Supplier flexibility, Performance measurement, Order-to-delivery process, Survey

1. Introduction
Due to increased volatility in customer demands, the ability to flexibly adapt has become a top priority for many companies (Üstündag and Ungan, 2020; Kumar and Singh, 2020; Kuo et al., 2016; Jafari, 2015). Many types of flexibility exist, with different types – sets of situations for which flexibility is required (Abdelilah et al., 2018) – and on different levels. Engelhardt-Nowitzki (2012) therefore encourages that flexibility be studied in a well specified way. Type-wise, to handle customers’ flexibility demands, flexible production may not be enough (Üstündag and Ungan, 2020). The possibilities to achieve flexibility are also highly affected by suppliers (Kumar and Singh, 2020; Bag et al., 2018), as manufacturers typically spend 60–80% of their total cost of purchasing (Üstündag and Ungan, 2020). Manders et al. (2017) call for insights from studying flexibility from the customer’s perspective, implying supplier flexibility. Supplier flexibility enables companies to adapt efficiently to changes, supported by suppliers’ capabilities (Bag et al., 2018; Lao et al., 2010). Therefore, the flexibility

© Helena Forslund and Stig-Arne Mattsson. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licenses/by/4.0/legalcode
type of the current study is supplier flexibility. Level-wise, flexibility can be understood on different management levels; on a strategic level as handling long-term, systematic changes in customer demand, or on an operational level as handling stochastic, unpredictable, short-term changes (Kumar and Singh, 2020). On an operational level, it is possible to study how supplier flexibility directly affects customers in the order-to-delivery process (Forslund et al., 2021). Liao (2020) claims that supply chain flexibility often is understood as linked to processes on an operational level. Goyal et al. (2018) similarly linked upstream supply chain flexibility to certain processes. Supplier flexibility in the operational order-to-delivery process is, in line with that, focused in the current study.

Performance measurement plays a critical role in any company and appropriate measures are needed for, e.g. communication purposes (Elgazzar et al., 2019; Bourne et al., 2018). Koste and Malhotra (1999) state that to be able to understand and develop any type of flexibility, companies must be able to capture it in a reasonably accurate and credible way. Elgazzar et al. (2019) found few flexibility measures in their recent overview of supply chain measures. Maestrini et al. (2018) highlight a research gap in how companies apply supplier performance measurement in a broad sense. Despite its significance for most companies and the extensive research interest, measurement of flexibility is a difficult task (Kumar and Singh, 2020), contains shortcomings and is underdeveloped (Engelhardt-Nowitzki, 2012). Knowledge on how to measure, present, compare and evaluate supplier flexibility is needed and has been repeatedly highlighted as a research gap (Forslund et al., 2021; Afy-Shararah and Rich, 2018; Manders et al., 2017), such as for supplier evaluation purposes (e.g. Jääskeläinen, 2018; Lao et al., 2010).

Studies on flexibility performance measurement focus on developing and analyzing measurement scales for research purposes, rather than empirically studying companies’ flexibility measurement practices (Kumar and Singh, 2020; Elgazzar et al., 2019). One empirical study investigated the use of performance measures in seven large Swedish manufacturing companies (Landström et al., 2016) and identified 3151 measures, with none of them concerning flexibility. According to Bourne et al. (2018), performance measurement often develops from practice. Consequently, the current study examines companies’ descriptions of their supplier flexibility measurement practices.

Knowledge in identifying the extent to which supplier flexibility is measured, and the ways in which it is measured, fills a gap in the literature about supplier flexibility performance measurement. The lack of existing measurement discovered in earlier research, indicates that potentially applied measures deserve further description and characterization; this would be a valuable contribution to and expansion of the flexibility performance measurement literature. Toward that background, it is also relevant to assess to what extent the applied measures are valid measures of supplier flexibility in the order-to-delivery process. The results can be useful for practitioners, who wish to improve their practices and align measurement with the prioritized ability of being flexible. Thus, this study addresses gaps in both the literature and in practice.

In search of supplier flexibility performance measurement, the purpose of this study is to identify, characterize and assess supplier flexibility measurement practices in the order-to-delivery process. The purpose is fulfilled by responding to two research questions;

**RQ1.** To what extent and in which ways are supplier flexibility measured in the order-to-delivery process?

**RQ2.** How can the applied measures be characterized, and to what extent can they be applied as measures of supplier flexibility in the order-to-delivery process?

The remaining article is organized as follows; a literature review (section 2) is carried out in order to frame the empirical data. The methodology (section 3) describes the empirical study, and its findings are presented in (section 4), responding to RQ1. Section 5 responds to RQ2, and
the purpose is fulfilled in section 6 by discussing the results and contributions of the study. Conclusions are presented in section 7.

2. Literature review of supplier flexibility performance measurement
The literature review primarily covers aspects on supplier flexibility performance measurement in the order-to-delivery process. The section is initiated with definitions of supplier flexibility and similar flexibility types. Dimensions of supplier flexibility and ways to measure and grade flexibility are then presented. Throughout the section, specifications and characterizations are marked with italic text.

2.1 Definitions of supplier flexibility and similar flexibility types
Earlier research has identified different flexibility types, also within supplier flexibility and its similar types. Conceptual unclarities are discovered. Jafari (2015) conducted a systematic literature review and defined inbound logistics flexibility as the ability to quickly and efficiently respond to customer needs in inbound delivery. A number of other close types, like purchasing and supply flexibility, were also found. Purchasing flexibility was defined by Zhang et al. (2002, p. 571) as “the ability of the organization to provide the variety of materials and supplies needed by manufacturing quickly and performance-effectively through corporative relationships with suppliers”. This definition rather points at the intra-organizational relation between manufacturing and the purchasing department, than at the intended inter-organizational flexibility from suppliers. The same view was defined for supply flexibility as “the ability of the purchasing function to respond in a timely and cost-effective manner to changing requirements of purchasing components in terms of volume, mix and delivery date” (Tachizawa and Thomsen, 2007, p. 1116). Even the term supplier flexibility is defined by Lao et al. (2010, p. 8) as “the extent of responsive abilities through the use of supplier-specific capabilities”, which also is a definition of the customer’s internal ability to use supplier flexibility.

Upstream supply chain flexibility was defined by Goyal et al. (2018, p. 830) as “the responsiveness of the upstream supply chain system to deal with variations in demand”. Such a definition is close to our view on supplier flexibility, even if it does not stress the stochastic, order-related changes. Also the definition of Kumar (2020); supplier flexibility is the assessment of suppliers’ ability to accommodate respondents’ requests and changes efficiently and meet emergency orders, is close to our view. If respondents are replaced with customers, it has the targeted order-related view. Supplier flexibility in the order-to-delivery process was defined by Forslund et al. (2021) as suppliers’ ability to fulfill short-term changes in customer demand. This definition is used in the current study.

2.2 Dimensions of flexibility in general and of supplier flexibility
One way of characterizing flexibility measures is by the dimension they address. Several authors have developed dimensions – the axes according to which flexibility types can evolve (Abdelilah et al., 2018) – for the flexibility concept. There is a large variety in taxonomies and conceptual models, specifying flexibility in different dimensions (Kumar, 2020; Jafari, 2015); furthermore, flexibility dimensions are related to each other (Manders et al., 2017). According to Stevenson and Spring (2009), the main difficulties in measuring flexibility are this multidimensional phenomenon, and that a system can be flexible in one dimension, while simultaneously being inflexible in another dimension.

A fundamental study was conducted by Slack (1987), who suggested five dimensions of flexibility in general: product flexibility, mix flexibility, volume flexibility, delivery flexibility and quality flexibility. Product flexibility was defined as “the ability to introduce novel
products, or to modify existing ones”; mix flexibility was defined as “the ability to change the range of products made within a given time period”; volume flexibility was defined as “the ability to change the level of aggregated output”; delivery flexibility was defined as “the ability to change planned or assumed delivery dates”; and quality flexibility was defined as “the ability to change planned product quality levels” (Slack, 1987, p. 38).

Forslund et al. (2021) related flexibility dimensions to the order-to-delivery process. As it basically concerns how much to order and when to deliver, the most obvious dimensions to include in the framework are volume flexibility and delivery flexibility. Volume flexibility was found by Liao (2020) and Abdelilah et al. (2018) to be the ability of a company to operate at various production output levels, a dimension similar to Slack’s (1987), in that volume is seen on an aggregated, long-term level. According to Forslund et al. (2021), the volume flexibility dimension in the order-to-delivery process is defined as the ability to change volumes or quantities on individual orders. Delivery flexibility was by Abdelilah et al. (2018) defined as the capability of the company to adapt lead-times to customers’ requirements, while Liao (2020) defined delivery flexibility as the ability of a company to deliver products to customers in response to uncertainties in, e.g. delivery dates. Both these views on delivery flexibility fit well into an operational level. Forslund et al. (2021) defined the delivery flexibility dimension in the order-to-delivery process as the ability to change delivery dates. With the clarification of these dimensions, the current study’s definition of supplier flexibility is further specified.

2.3 Ways to measure and grade flexibility
Different ways of characterizing flexibility measures are described in this sub-section. Stevenson and Spring (2009) pointed out that no simple flexibility measure can be established for a system. According to Slack (1983), the difficulty of achieving a single flexibility measure stems from three different qualities: flexibility reflects potential rather than performance; flexibility is not an isolated achievement but must be related to other performances, such as quality, volume and delivery; and flexibility must be graded in terms of scope, cost and time to accomplish. For these reasons, Slack (1983) argued that it may be better to develop a systematic qualitative evaluation method than to try to determine quantitative measures. Later, Gerwin (1993) explained that it is common for companies to use one-dimensional, qualitative, value-based flexibility statements that are easy to understand. No identified study measured flexibility in the qualitative way suggested by, e.g. Gerwin (1993).

Slack (1987, p. 39) used two grading variables for each flexibility dimension: he defined range flexibility as the "range of states which the production system or resource is capable of achieving" and response flexibility as “the ease (in terms of cost, time or both) with which changes can be made”. Golden and Powell (2000, p. 376) applied four grading variables to the capacity to adapt: temporal, defined as “how long it takes an organization to adapt”; range, defined as “the number of options that an organization has open to it for change that was foreseen and the number of options it has available to react to unforeseen change”; intention, defined as “whether the organization is being proactive or reactive to changing conditions”; and focus, defined as “whether the flexibility is gained internally to the organization or by managing external relationships with trading partners”. These ways of grading are not related to the order-to-delivery process; Slack’s view on flexibility is long-term and Golden and Powell’s is highly general.

Maestrini et al. (2018) surveyed Italian customer companies and asked to what extent they measured supplier flexibility performance, rated on a Likert scale. The study did not reveal how supplier flexibility performance was defined or operationalized, which is why it was not possible to further use. Liao (2020) and Mishra et al. (2018) measured flexibility items related to, e.g. volume and delivery dimensions on multi-item Likert scales. Forslund et al. (2021) also used multi-item Likert scales to assess operationalized items related to the volume and delivery flexibility dimensions.
According to Stevenson and Spring (2009), the literature on flexibility measurement suffers from the fact that flexibility measures are inadequately defined. The case study by Jääskeläinen (2018) on supplier performance measurement did not identify any flexibility measures. Few studies have defined flexibility measures, although one exception is the supply chain operations reference (SCOR) model. The SCOR measure of supply chain flexibility measures the average time it takes to respond to an unplanned demand increase (APICS, 2018); it is a way of capturing the dimension of volume flexibility, but it does not explicitly address the supplier. The similar SCOR measure of supply chain adaptability measures the supplier’s ability to manage changes in total delivery volumes within a given timeframe (APICS, 2018)—seen as another way to capture the dimension of volume flexibility, but it does not fit the order-related changes in the current study.

The practice to measure flexibility by using other existing measures has been revealed in earlier studies. Baird and Su (2018) suggested to measure flexibility performance in manufacturing companies as a broad and complex index based upon other measures; asset turnover, inventory days of supply, cash-to-cash cycle time and production flexibility. However, these measures were not clearly defined and not explicitly related to supplier flexibility. Mahadevan (2017) suggested to measure supply chain flexibility as the average time to respond to unplanned orders, another measure that does not explicitly address the supplier as it does not mention whose average time. Similarly, Forslund and Jonsson (2010) studied performance measurement practices between customers and suppliers and found flexibility aspects in delivery reliability. In terms of delivery reliability—the share of orders that are delivered on confirmed or wished delivery date (Forslund and Jonsson, 2010)—the majority of surveyed suppliers (72%) based their measures definition on the confirmed delivery date, which is agreed upon by both the customer and the supplier. However, basing the definition on the wished delivery date—i.e. the delivery date that the customer initially demanded before negotiating and agreeing with the supplier—would better reflect the suppliers’ ability to flexibly deliver (Forslund and Jonsson, 2010). Delivery reliability was defined by APICS (2018) as measuring how consistently goods are delivered at or before the promised time. Elgazzar et al. (2019) found many supply chain performance measures to be theoretical and to have limited practical implementation. Jääskeläinen (2018) further concluded that performance measurement research seldom is explicit with the related practices. The same is valid for the suggested flexibility measures; they are more conceptual than operational, which was also concluded in the recent systematic literature review by Kumar and Singh (2020), as well as indicated as a research gap by Liao (2020). Hence no operational or operationalized supplier flexibility measures for practical use in a company are identified in earlier research.

3. Methodology
3.1 Research design
The lack of operationalized supplier flexibility measures (Kumar and Singh, 2020; Liao, 2020; Jääskeläinen, 2018) and the recommendations for qualitative approaches (Gerwin, 1993; Slack, 1983) direct the research design toward a qualitative research approach (Bryman and Bell, 2007). At the same time, the case studies of Jääskeläinen (2018) and Landström et al. (2016) did not identify any flexibility measure, which directs the research design toward a broader quantitative survey study (Hair et al., 2019; Bryman and Bell, 2007). It was decided to combine the approaches in research design and to add a qualitative part to a quantitative survey study. The first research question of the study was hence addressed through an added, mainly qualitative part of a survey, where the quantitative part was presented in a previous study on supplier flexibility in the order-to-delivery process (Forslund et al., 2021).
3.2 Survey instrument
In the cover letter, supplier flexibility was defined as the supplier’s ability to handle short-term changes in demand. The survey instrument, procedures and results from the quantitative part of the survey are described in Forslund et al. (2021). The respondents were thereafter asked with a qualitative, nominal yes/no scale (Hair et al., 2019) “Do you measure, or in a systematic way evaluate, your suppliers’ flexibility?” Since the literature on flexibility performance measures provides unclear operationalizations, it was not possible to give pre-defined response alternatives. If the answer was yes, the respondents were further asked “Please describe how these practices look like” in open text. The respondents had no limits on the possible amount of text to write.

3.3 Sampling method and data collection
The study was carried out in Sweden, a geographic context in which few empirical studies on supplier flexibility have been identified (the exception is Landström et al., 2016). A critical issue was finding respondents with the right prerequisites to respond. Respondents with the position of “purchasing manager” were addressed, as they were perceived as having the best knowledge about supplier flexibility. Contact information in the shape of e-mail addresses was found on public Internet webpages and in company databases, which constitute the sampling frame. The sampling criteria sought only purchasing managers who purchase material. As it was difficult to only include these, the researchers decided to reach out to all identified purchasing managers displaying an e-mail address and afterward exclude responses of managers who purchase services.

A cover letter was e-mailed to, in total, 24,293 respondents (the gross population, which included “all” industries and company sizes). The cover letter included a link to an on-line questionnaire. One wave of e-mailing resulted in 555 complete responses. Due to the sampling method, it was not possible to conduct an analysis of non-response bias. The response rate was here 2.3%. A large number of responses (158) were received from companies with less than 20 employees. It can be assumed that small companies have different organizations and operations from larger companies, in terms of resources, process formalization, knowledge level and management. In order to establish a more homogenous sample, responses from the smaller companies were excluded. Also responses from wholesalers/retailers, service and government-controlled companies, were excluded for higher homogeneity. Respondents who purchased services were also excluded. This left the focus on 224 respondents that represented material purchasing in larger manufacturing companies in Sweden, on which the following analysis is based. The sample in described by company size and supplier flexibility measurement practices in Table 1.

3.4 Analysis
The open-ended text answers were coded into response classes, a process further described in section 4. Using such an inductive approach implies that classes are derived from the data under study and that an iterative process of class building, testing and revising is conducted,

<table>
<thead>
<tr>
<th>Company size (employees)</th>
<th>Respondents</th>
<th>Measure supplier flexibility: No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–49</td>
<td>75</td>
<td>26 (35)</td>
</tr>
<tr>
<td>50–249</td>
<td>103</td>
<td>31 (30)</td>
</tr>
<tr>
<td>250+</td>
<td>46</td>
<td>20 (43)</td>
</tr>
<tr>
<td>Total</td>
<td>224</td>
<td>77 (34)</td>
</tr>
</tbody>
</table>
in line with recommendations for systematic and transparent content analysis by Seuring and Gold (2012). This was carried out by the two researchers individually, after which the individual codings were compared and the few discrepancies were addressed; the inter-coder reliability (Seuring and Gold, 2012) was high. Responses were translated from Swedish into English. Differences between company sizes were analyzed. The analysis of the second research question of the study built upon the statement from Bourne et al. (2018) that performance measurement often develops from practice. RQ2 was fulfilled by characterizing (flexibility dimensions as well as ways of measuring and grading flexibility) and assessing the extent to which the applied practices are valid supplier flexibility measures.

4. RQ1. To what extent and in which ways are supplier flexibility measured in the order-to-delivery process?

This section is initialized by describing the sample in Table 1.

The sample was found to be a good mix of different company sizes. The largest companies are more likely to measure supplier flexibility. This result is in line with a general belief that larger companies have more resources, more formalized processes, higher levels of knowledge and more professional management than smaller companies.

In total, 77 respondents responded yes on the question if they measured supplier flexibility. However, nine respondents did not provide any example of how they did so in the open-ended text, so they were excluded from further analysis. Of the remaining 68 respondents, seven provided an unspecific description (e.g. “we measure in the ERP system”, “we use KPIs”, “I know we measure but I do not know how”, or “we have just started measuring”). These responses were also excluded from further analysis. As a result, the number of companies who are able to describe how they measure supplier flexibility is considered to be 61, or 27% of the respondents. It was also found that some companies measured supplier flexibility in two ways (e.g. using two quantitative measures as in three companies), or using one quantitative measure and one qualitative follow-up as in four companies). Therefore, the total number of supplier flexibility measures applied is 68. A first analysis of the open answers showed that the alternative ways of measuring supplier flexibility to a large extent built upon measuring other order-to-delivery process-related measures. This is similar to the approach of Baird and Su (2018), who based their way of measuring flexibility on capturing other existing measures. The following classes were empirically identified as shown in Table 2.

The largest class includes those respondents whose answers showed that they measure delivery reliability – the share of orders that are delivered on confirmed or wished delivery date (Forslund and Jonsson, 2010). They represent 47% of the supplier flexibility measurements applied. The delivery reliability measurers were found to have different practices. Some measure delivery reliability by comparing confirmed delivery date to actual delivery date and other measure delivery reliability by comparing wished delivery date to actual delivery date. The latter sub-class consists of eight responses or 25% of the delivery reliability measurers. This is a practice which better reflects the customers’ demands than

<table>
<thead>
<tr>
<th>In which ways is supplier flexibility measured</th>
<th>No. of measurements applied (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure delivery reliability</td>
<td>31 (47)</td>
</tr>
<tr>
<td>Follow up: audits/evaluations/assessments</td>
<td>28 (41)</td>
</tr>
<tr>
<td>Measure delivery lead-time</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Measure supplier flexibility in other ways</td>
<td>5 (7)</td>
</tr>
<tr>
<td>Total</td>
<td>68 (100)</td>
</tr>
</tbody>
</table>

Table 2. Classification of alternative ways of measuring supplier flexibility
using confirmed date as the comparison date (Forslund and Jonsson, 2010). In the study of Forslund and Jonsson (2010), a similar share of 28% of delivery reliability measurers based their delivery reliability measurement on wished delivery date.

The second-most common practice, with 41% of the supplier flexibility measurements, includes respondents who conduct supplier follow-ups, audits, evaluations, or assessments. This is typically described as being carried out qualitatively and infrequently. Slack (1983) and Gerwin (1993) argued that qualitative ways of measuring flexibility have the advantage of being easy to understand.

The smallest class includes four responses (5%) whose measures are related to delivery lead-time – defined by APICS (2018) as the time from the receipt of a customer order to the delivery of the product.

Next class includes five exemplified (7%) supplier flexibility measurements applied (SF 1–5) that do not fall into any of the other classes.

(1) Analyze administrative costs and freight costs related to changed order quantities
(2) Report how many deliveries are on time when we have wished shorter lead-time, changed order quantity, or changed means of transportation
(3) Measure “change order request” in time for production
(4) Measure the number of changes within the lead-time and compare to the wished delivery date
(5) Conduct capacity/flexibility analysis using our own tools (demanding a volume flexibility of +20% within three months). As the timeframe is set to three months, this measure reflects long-term volume adaptations rather than the intended order-to-delivery process changes. It is therefore not analyzed further.

5. RQ2. How can the applied measures be characterized, and to what extent can they be applied as measures of supplier flexibility in the order-to-delivery process?

One way of characterizing the applied measures is into qualitative and quantitative ways. Measures of follow-up: audits/evaluations/assessments are shown in Table 2 as 41% of the ways of measuring. They are qualitative ways of measuring (e.g. Gerwin, 1993). Some examples of responses are “we continuously evaluate our suppliers’ flexibility”, “all types of deviations are captured in audits” and “annual subjective evaluation of suppliers’ flexibility”. The qualitative ways of measuring were difficult to analyze due to lack of elaboration in answers. Accordingly, these are not analyzed further. Even if all applied quantitative measures, 59% of the ways of measuring, are not supplier flexibility measures per se, the analysis assesses the possibilities for applying them as supplier flexibility performance measures. In 5.1, delivery reliability as a measure for supplier flexibility is characterized and assessed. In 5.2, one delivery lead-time measure is characterized and assessed. In 5.3, the empirically identified supplier flexibility measures are characterized and assessed.

5.1 Delivery reliability as a measure for supplier flexibility – characterization and assessment

The largest share of companies was found to apply delivery reliability as a measure for supplier flexibility. The practice can now be characterized. Delivery reliability can be interpreted as a supplier’s ability to flexibly adapt lead-times in the order-to-delivery process; therefore, it can be considered a measure of supplier flexibility in the delivery dimension (Forslund et al., 2021; Abdelilah et al., 2018). It was obvious that those respondents who
provided some more details about their measurement practices, focused on delivery dates ("we compare wished and confirmed delivery date", "we measure towards our wished date", "we measure delivery reliability against confirmed delivery date") and hence the delivery flexibility dimension. To refer to the dimension as “delivery flexibility” is misleading, as a delivery includes aspects of both volume and lead-time, in line with the definition of Liao (2020). For better clarity, we refer to this dimension as “delivery lead-time flexibility” hence-forth.

The volume dimension was only addressed in two responses, where both at the same time addressed the delivery lead-time dimension ("we measure delivery reliability in terms of both time and volume" and "we measure suppliers delivering too early, too late and with deviations in quantity"). The established term volume dimension includes the ability to adapt volumes long-term (Slack, 1987) or short-term (Forslund et al., 2021). In order to create a more context-specific definition, the term “order quantity flexibility” instead implies an adaptation of the volume dimension to an order-to-delivery process context.

No explicit mentioning of ways to grade flexibility (Slack, 1987; Golden and Powell, 2000) were found among the responses. Here it must be acknowledged that none of these ways to grade are explicitly related to the order-to-delivery process. When using delivery reliability as a way to measure supplier flexibility, a percentage is the result. This is one potential way to grade and compare different suppliers’ flexibility, with other suppliers or with themselves over time.

Neither did any respondent explicitly refer to the defined SCOR measures, or to Likert scale measures. Interestingly, the two operationalized flexibility measures identified in the literature review, the SCOR measures of supply chain flexibility and supply chain adaptability (APICS, 2018), both relate to the volume flexibility dimension. Operationalized flexibility measures covering the delivery lead-time dimension are consequently lacking, particularly toward the frequent use of time-related measures.

The practice to measure flexibility by other measures was revealed in earlier studies (e.g. Baird and Su, 2018; Mahadevan, 2017). Measuring delivery reliability is a mature and well-developed measurement practice (Forslund and Jonsson, 2010), that has support in most ERP systems (Forslund, 2010), which may be the reasons for this widespread use. This however seems to indicate that companies think that flexibility is the same thing as delivery reliability, or that they think that when a supplier is flexible, it means that the supplier is delivering reliably. At first glance, this can be seen as a faulty way of thinking. Delivery reliability reflects the ability to deliver according to a confirmed (APICS, 2018) or wished delivery date (Forslund and Jonsson, 2010). Definitions of supplier flexibility include the ability to adapt to variations in demands (Goyal et al., 2018), requests and changes (Kumar, 2020) or short-term changes in customer demand (Forslund et al., 2021). So, what happens when the customer wants to change an already-confirmed delivery date or wish a new delivery date – what type of supplier performance is then measured? This creates a gray zone between flexibility and delivery reliability.

Four different delivery reliability measures (DR1-4) can be distinguished, depending on the point in the order-to-delivery process (at order, after order confirmation, at delivery) that the measurement takes place. They all relate to the delivery lead-time dimension. The fact that many measures are provided confirms Stevenson and Spring (2009), who claim that flexibility cannot be presented as one single measure.

DR1. Supplier delivery reliability at order is defined as the number of orders for which the wished delivery date could be confirmed, despite the fact that the delivery lead-time is shorter than normally, divided by the total number of orders with the wished delivery lead-time shorter than normally during a period (a day, a week or whatever period that measurement takes place). This measure represents delivery reliability in the form of the difference between the wished and the confirmed delivery dates. It can be seen as having a reactive intention and
an external focus (Golden and Powell, 2000). It is assessed as expressing the supplier’s flexibility regarding the extent to which it can meet customers’ requests at order of a certain delivery lead-time. Since this flexibility measure only reflects what the supplier aims to provide, it must be supplemented by a meaningful measure of delivery reliability at delivery.

**DR2.** Supplier delivery reliability after order confirmation but before delivery is defined as the number of orders for which the confirmed delivery date has been accepted to be brought forward or backward according to customer requests, divided by the total number of orders with requests to bring forward confirmed delivery dates during a period. The measure represents delivery reliability in the form of the difference between a new wished delivery date (earlier or later) and a previously confirmed delivery date. It has a reactive intention and an external focus (Golden and Powell, 2000). It is assessed to express the extent to which the supplier is able to respond to customers’ wishes after order confirmation. Since also this measure only reflects what the supplier aims to provide, it must be complemented with a meaningful measure of delivery reliability at delivery.

**DR3.** Supplier delivery reliability at delivery versus confirmed delivery date is defined as the number of orders delivered at confirmed delivery date, divided by the total number of delivered orders during a period. Since this measure involves a comparison between confirmed and actual delivery date, it is a pure delivery reliability measure and can in that sense not be seen as a flexibility measure. It has also a reactive intention, but an internal focus (Golden and Powell, 2000). However, it is assessed to in one way capture supplier flexibility. Since all production consumes time between order and delivery, the supplier’s manufacturing conditions can be changed and affect its ability to deliver at confirmed delivery date. This can occur due to disruptions in the availability of capacity, disruptions in the supply of materials, discrepancies between planned and actual production times and variations in the amount of incoming orders. Therefore, this measure can also be regarded as an expression of the extent to which the supplier is flexible and able to handle changed operating conditions, consistent with Slack’s (1983) observation that flexibility expresses a potential rather than a performance.

**DR4.** Supplier delivery reliability at delivery versus wished delivery date is defined as the number of orders delivered at the wished delivery date, divided by the total number of delivered orders during a period. This measure represents delivery reliability in the form of the difference between the wished and the actual delivery dates. This measure can be seen as having a reactive intention and an external focus (Golden and Powell, 2000). It is assessed to address the extent of flexibility in being able to meet customers’ wishes to deliver at a certain delivery date and flexibility in handling changed operating conditions in the same way as DR3.

One could distinguish between customer-related delivery lead-time flexibility (DR1 and 2) and operating condition-related lead-time flexibility (DR3 and 4).

5.2 Delivery lead-time as a measure for supplier flexibility – characterization and assessment

Delivery lead-time (DL) is used to some extent as a supplier flexibility measurement practice. It is defined by APICS (2018) as the time from the receipt of a customer order to the delivery of the product; it is primarily a measure of a supplier’s delivery performance in terms of the time it takes to respond to customers’ demands. Some examples of responses from the study are “we focus on short lead-times” and “we measure our lead-times”. It is obviously characterized as dealing with the delivery lead-time dimension. Accordingly, delivery lead-times can be considered as a measure of supplier flexibility. Also, Baird and Su (2018) included cycle time as one part of measuring flexibility, and Mahadevan (2017) based their definition of supply chain flexibility on lead-time. Therefore, the literature supports using lead-time as a way to measure supplier flexibility.
DL1. Short delivery lead-time can hence be an enabler for being a flexible supplier – e.g. having short lead-times means greater potential to adjust to changing customer demands. But in itself, short delivery lead-time does not reflect the extent to which the supplier is flexible. First, the length of the delivery lead-time does not concern the volume flexibility dimension. Second, the only way for a supplier to truly accomplish lead-time flexibility is to add extra lead-time to the “normal” lead-time when establishing and communicating current delivery lead-times to customers. When customers then occasionally demand shorter delivery lead-times, flexibility can be provided by reactively reducing the added extra lead-time. Then, delivery lead-time can represent a delivery lead-time flexibility measure, and the size of the extra lead-time can correspond to Slack’s (1987) range flexibility. Used like that, it has the only found proactive intention and an external focus (Golden and Powell, 2000). However, a prerequisite is that the “normal” lead-time is short enough to allow adding extra lead-time while remaining competitive.

5.3 The empirically identified supplier flexibility measures – characterization and assessment

Four measures (SF1-4) are here identified.

SF1. “We analyze administrative costs and freight costs related to changed order quantities.” It was found in a company with 20–49 employees. This measure addresses the volume dimension (Forslund et al., 2021) and relates it to the company’s own costs. This is in line with the definition of supply flexibility by Tachizawa and Thomesen (2007) that includes a cost-effective response to changing requirements of purchased components in terms of volume, mix and delivery date. It can also be linked to response flexibility – e.g., the ease in terms of cost with which changes can be made – suggested as a way to grade flexibility by Slack (1987) and to an external focus (Golden and Powell, 2000). This measure captures one aspect of supplier flexibility: the cost of achieving it. When these costs are low, the supplier is flexible and able to meet the customer’s demands. The measure thus indicates the extent to which the supplier is flexible.

SF2. “We report how many deliveries are on time when we have wished shorter lead-time, changed order quantity, or changed means of transportation.”, was found in a company with 50–249 employees. This company keeps track of the number of changes (the demand for supplier flexibility) that result in fulfilled orders after changes in the delivery lead-time and order quantity dimensions (Forslund et al., 2021; Liao, 2020) together with a third type of change: means of transportation. This is a type of delivery reliability measure, but it seems as changes are treated individually and separated from the delivery reliability measure. This approach to measuring can possibly be linked to the grading variables range and to an external focus (Golden and Powell, 2000). This measure also indicates the extent to which the supplier is flexible.

SF3. “We measure “change order request” in time for production” was found in a company with 50–249 employees. According to Jääskeläinen (2018) and Stevenson and Spring (2009), the literature on flexibility measurement lacks adequate definitions of flexibility measures. The same is true for this response. The measure seems to be valid for the delivery lead-time dimension. “Change order request” can be a company-internal practice reflecting the number of demanded order-related changes, thus capturing the demand for supplier flexibility. The “change order requests” are then compared to whether a supplier is able to still deliver on time for production, which creates a quota or a percentage indicating the extent to which the supplier has been able to fulfill the changed need.

SF4. “We measure the number of changes within the lead-time and compare to the wished delivery date” was found in a company with more than 250 employees. This approach is similar to SF2, in that it keeps track of the demand for supplier flexibility. It is unclear how to compare a
number with a delivery date to create a quota. This measure focuses on the delivery lead-time flexibility dimension and can be linked to the grading variables range and to an external focus (Golden and Powell, 2000). It reflects the extent to which the supplier is flexible.

6. Discussion and contributions
This section discusses the contributions to literature and the practical implications.

6.1 Contributions to literature
Supplier flexibility can directly and frequently be perceived in the order-to-delivery process, as suppliers’ response to customers’ short-term order-related changes. A starting point was that literature seldom focuses on supplier flexibility in the order-to-delivery process. Consequently, also how to measure supplier flexibility in the order-to-delivery process was sparsely treated in extant research. This study gives some contributions to these areas. Measures are particularly critical for communication reasons (Bourne et al., 2018), such as communicating with existing or potential suppliers. Therefore, a specific and adapted vocabulary is required.

One specific contribution to the flexibility vocabulary is to develop or adapt the established but misleading dimension “delivery flexibility” to the order-to-delivery process. As a delivery includes aspects of both volume and lead-time “delivery lead-time flexibility” could be used, for better clarity. In the same vein, the other established dimension “volume flexibility” often represents the ability to change the level of aggregated volumes over time. The term “order quantity flexibility” instead implies an adaptation of the volume dimension to an order-to-delivery process context. This contribution could be valid for supplier flexibility, but it could also be applied downstream for other types and scopes of flexibility, such as the company’s flexibility toward customer.

Many respondents focused on the delivery lead-time flexibility dimension when measuring supplier flexibility. At the same time, it was seen that the few operationalized flexibility measures from literature only covered the order quantity flexibility dimension. Measures addressing the delivery lead-time dimension are consequently lacking. The study identified and operationalized a number of measures, responding to gaps pointed out by Kumar and Singh (2020) and Elgazzar et al. (2019). Four different delivery reliability measures, one delivery lead-time measure and four practically applied supplier flexibility measures are identified, giving starting points to contributions to literature. All supplier flexibility measures have an external focus and are reactive (Golden and Powell, 2000). They all measure the extent to which the supplier is flexible, but by focusing on different flexibility aspects. An overview of the measures is shown in Table 3.

The study bridges a gap between theory and practice, in the sense that Bourne et al. (2018) suggested that performance measurement often develops from practice. The empirically identified and characterized supplier flexibility measures can contribute to the body of knowledge on flexibility performance measurement and lead to theory development.

Finally, it was found that many similar definitions exist for supplier flexibility, with different meanings. To clarify that supplier flexibility is studied from a customers’ inter-organizational perspective and not how the intra-organizational purchasing department handles flexibility, a possibility could be to refer to still another synonym—suppliers’ flexibility. Such a concept could contribute to the flexibility vocabulary and literature with a lower risk for misunderstanding.

6.2 Practical implications
Landström et al. (2016) stated that no studied companies measure flexibility, which could seem true at first sight. The findings of the current study shed a new light over that study.
The findings that practitioners to a large extent measure supplier flexibility by applying other measures, can be an explanation to the Landström et al. (2016) findings and can give practitioners’ measurement practices a restoration.

The high dependence on suppliers imply their important role in achieving flexibility (Kumar and Singh, 2020; Üstündag and Ungan, 2020; Bag et al., 2018). This study therefore has practical implications. Other purchasing, logistics, or supply chain managers can gain ideas and inspiration from the supplier flexibility measures identified. They can select from the proposed measures and find measures covering the needed dimension(s). Stevenson and Spring (2009) pointed out that, e.g. a supplier can be flexible in one dimension while being inflexible in another, therefore the measurement system need to capture this. Practitioners can also select measures that included different aspects of measuring supplier flexibility. This can develop and improve supplier measurement and evaluation practices, in line with the research gap outlined by Jääskelainen (2018). It can also lay the foundation for supplier development.

The suggested clarifications of flexibility-related concepts and vocabulary have not the least practical implications and are suggested to be used also in the practical situation. Language and communication problems related to supply chain performance measurement was noticed by, e.g. Bourne et al. (2018) and Forslund and Jonsson (2010). Language problems can direct the focus on discussing unclarities in measurement results, instead of analysis and improvements between customers and suppliers. With an agreed and clear vocabulary, all efforts can be directed to improvement work.

The ability to flexibly adapt has become a top priority for many companies (e.g. Üstündag and Ungan, 2020; Kumar and Singh, 2020). At the same time, the possibilities to achieve

<table>
<thead>
<tr>
<th>Measure</th>
<th>Characteristics</th>
<th>Extent to which it measures supplier flexibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR1. Supplier delivery reliability at order</td>
<td>Delivery lead-time flexibility dimension</td>
<td>Reflects the aim for supplier flexibility measures customer-related flexibility</td>
</tr>
<tr>
<td>DR2. Supplier delivery reliability after order confirmation but before delivery</td>
<td>Delivery lead-time flexibility dimension</td>
<td>Reflects the aim for supplier flexibility measures customer-related flexibility</td>
</tr>
<tr>
<td>DR3. Supplier delivery reliability at delivery versus confirmed delivery date</td>
<td>Delivery lead-time flexibility dimension</td>
<td>Measures operating condition-related supplier flexibility</td>
</tr>
<tr>
<td>DR4. Supplier delivery reliability at delivery versus wished delivery date</td>
<td>Delivery lead-time flexibility dimension</td>
<td>Measures operating condition-related supplier flexibility</td>
</tr>
<tr>
<td>DL1. Short delivery lead-time</td>
<td>Delivery lead-time flexibility dimension</td>
<td>Enabling supplier flexibility</td>
</tr>
<tr>
<td>SF1. Analyze administrative costs and freight costs related to changed order quantities</td>
<td>Order quantity flexibility dimension response flexibility</td>
<td>Measures costs for achieving flexibility, and thereby the extent to which the supplier is flexible</td>
</tr>
<tr>
<td>SF2. Report how many deliveries are on time when we have wished shorter lead-time, changed order quantity, or changed means of transportation</td>
<td>Delivery lead-time and order quantity flexibility dimension range flexibility</td>
<td>Measures need for and the extent to which the supplier is flexible</td>
</tr>
<tr>
<td>SF3. Measure “change order request” in time for production</td>
<td>Delivery lead-time flexibility dimension</td>
<td>Measures need for and the extent to which the supplier is flexible</td>
</tr>
<tr>
<td>SF4. Measure the number of changes within the lead-time and compare to the wished delivery date</td>
<td>Delivery lead-time flexibility dimension range flexibility</td>
<td>Measures need for and the extent to which the supplier is flexible</td>
</tr>
</tbody>
</table>

Table 3. Identified measures for supplier flexibility in the order-to-delivery process
flexibility are to a large extent affected by suppliers (Kumar and Singh, 2020; Üstündag and Ungan, 2020; Bag et al., 2018). Without valid performance measurement practices, for suppliers’ and potentially also for customer flexibility, this priority is difficult to ensure and realize. The recent COVID-19 pandemic has painfully highlighted the need for flexibility in the supply chain, and many manufacturers have experienced uncertainties in the interface with suppliers. In line with the well-known expression “what gets measured gets done”, supplier flexibility performance measurement should be aligned with competitive priorities and strategies. A well-developed supplier performance measurement system can be associated with higher performance levels (Forslund and Jonsson, 2010). Therefore, higher levels of supplier flexibility can be expected, which in turn has the potential to improve flexibility in the whole supply chain.

7. Conclusion, limitations and further research
In order to target the research gaps related to measuring flexibility laid out by Forslund et al. (2021), Kumar and Singh (2020), Elgazzar et al. (2019), Afy-Shararah and Rich (2018), Manders et al., 2017, the purpose of this study was to identify, characterize and assess supplier flexibility measurement practices in the order-to-delivery process. The details of the supplier measurement practices revealed that most purchasing managers actually measure delivery reliability or delivery lead-time, conduct qualitative follow-ups, or cannot specify how flexibility is measured. The researchers characterized and assessed to what extent the applied practices actually represent measures for supplier flexibility. A number of measures for supplier flexibility in the order-to-delivery process were identified. It was acknowledged that they represent opportunities to measure different aspects of flexibility, in line with the suggestions of Baird and Su (2018).

As for all research, this study has its limitations. Due to the applied sampling method, it was not possible to describe which manufacturing sub-industry the respondents belonged to. The applied method implied that non-response bias could not be assessed. We do therefore not set out to provide a state-of-the-art description of practices. Instead, we identify, characterize, exemplify and assess supplier flexibility measurement practices. The study reflects the practices in Sweden. The companies are all manufacturing companies with over 20 employees purchasing materials. Small manufacturers, wholesaling/retailing, service and public organizations are not included in the article, in order to present more homogenous findings. Neither are those purchasing managers who purchase services included. Purchasing managers were addressed as they were assumed to have the best prerequisites to assess supplier flexibility. It is unknown if they also have the best prerequisites to assess the need for supplier flexibility. The study focused supplier flexibility. The order-to-delivery process takes place both upstream the supply chain toward suppliers and downstream the supply chain toward customers. Therefore, there are good possibilities to apply the findings of the study; vocabulary, definitions and measures, also for customer flexibility, or downstream supply chain flexibility in line with Goyal et al. (2018).

As for further research, it would be of interest to further develop, assess and test the usefulness of the above-mentioned practices in measuring supplier flexibility in some companies – not least because the literature lacks operationalized measures and because flexibility is of great interest and importance to companies. The fact that the surveyed companies largely equate delivery reliability and flexibility also suggests that this could be a topic of interest and it would be interesting to know more about the reasons why. Another path for future research would be to contact the respondents who described supplier flexibility measures and to conduct deeper case studies with them. This would provide more information about the details of flexibility measurement practices. It was seen that grading supplier flexibility in the order-to-delivery process, using grading variables related to
long-term flexibility definitions (Slack, 1987), had some applicability. Grading variables related to general flexibility definitions (Golden and Powell, 2000), like intention and focus, were applicable. More adapted grading variables for the order-to-delivery process could be developed by doing case studies in further research. A final suggestion for future research would be to specifically focus on the flexibility performance measurement capabilities of common ERP systems, in a similar way as in Forslund (2010), in order to support companies in search of flexibility performance management.

References
APICS (2018), Module 3 Supply Chain Improvement and Best Practice, Certified Supply Chain Professional - Learning System 2018, APICS, Chicago.


**About the authors**

Helena Forslund is Professor in Logistics and Supply Chain Management at School of Business, Linnaeus University, and Guest Professor at School of Engineering, Jönköping University, Sweden. She received her PhD from Linköping University, Sweden. She has published over 30 journal articles in e.g. International Journal of Physical Distribution and Logistics Management, International Journal of Operations and Production Management and International Journal of Productivity and Performance Management. Her research interests are within Supply Chain Management, Performance Management and Process Management. Helena Forslund is the corresponding author and can be contacted at: helena.forslund@lnu.se

Stig-Arne Mattsson has some 30 years of company experience in Production and Inventory Management, Supply Chain Management and Information Systems. Parallel to his industrial career he was Adjunct Professor in Supply Chain Management at Linnaeus University, Lund University and Chalmers University of Technology and is currently Senior Professor at Jönköping University in Sweden. He has written a number of textbooks and papers in internationally acknowledged journals within the area of Production and Inventory Management. He is certified in Production and Inventory Management (CFPIM) by The Educational Society for Resource Management, APICS and in Logistics (ESLog) by European Logistics Association, ELA.

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)