Assessing Outcomes from Business-to-Business Selling

Per Ivar Seljeseth

NORD UNIVERSITY BUSINESS SCHOOL

NORD University

www.nord.no

Assessing Outcomes from Business-to-Business Selling

Per Ivar Seljeseth

PhD in Business
Nord University Business School

PhD in Business no. 84 (2021)
Assessing Outcomes from Business-to-Business Selling
Per Ivar Seljeseth

To my grandmother and grandfather

ACKNOWLEDGMENTS

ABSTRACT

The importance of personal selling as a marketing tool has increased in recent decades, along with the maturing sales research discipline. However, researchers have noted the need for sales research to improve some of its research design practices regarding the conceptualizations and operationalizations of constructs and the use of data sources and respondents. This thesis responds to these issues by investigating the conceptualization, operationalization, data sources, and respondents used to assess outcomes from business-to-business (B2B) selling—also termed B2B sales performance outcomes.

Sales performance outcomes represent the outcomes that salespeople produce and range from economic outcomes, such as sales revenue, to outcomes associated with salespeople's customer interactions, such as customer satisfaction. Sales researchers frequently use these outcomes as dependent variables to estimate antecedents' effects on the outcomes and thereby identify antecedents to improved selling. Consequently, developing dependable knowledge of successful selling relies on reliable and validly assessed outcomes. Also, sales managers depend on reliable and validly assessed outcomes because of the intense managerial focus on optimizing outcomes from selling. For example, precise assessments of sales performance outcomes enable managers to detect low performance on critical outcomes and to take actions for improvement.

Despite these outcomes' importance, previous research provides little guidance or consensus on how they should be assessed. Further, although the antecedents of improved selling have been investigated extensively and are the subject of reviews and meta-analyses, sales performance outcomes have not

been reviewed. Therefore, this thesis provides the first investigation and review of this topic by addressing the following overarching research question: how do researchers assess outcomes from B2B selling?

A systematic literature review was conducted to answer this research question. To be included in the review, studies need to assess the outcomes from B2B selling, be empirical, be quantitative, and be published relatively recently (2001–2015). The search resulted in 139 studies. Data were extracted from these studies, and a unique dataset was created describing how researchers assess the outcomes, including the studies' measures, use of objective and/or subjective measures, number of measures, respondents, and data sources. Each of these methodological issues required specific data analysis, examination, and evaluation in relation to particular previous research and were thus handled in four research papers.

The first paper investigates the measures used to assess the outcomes. The reviewed studies use a large variety of measures, and a large portion of the studies use a few measures of sales revenue to assess the outcomes. Using such few measures disregards the multiple types of outcomes desired from B2B selling. Further, many studies fail to measure outcomes beneficial to customers, such as offer value and customer satisfaction. This paper contributes with recommendations for improving these measures and reveals the need to develop theory explaining which outcomes are desired from B2B selling.

The second paper suggests such a theory by developing the B2B Sales Performance Outcomes Chain. This chain contributes as the first complete theoretical framework conceptualizing desired outcomes from B2B selling. The framework identifies seven main types and 21 subtypes of outcomes and can be used to select measures with stronger construct validity.

The third paper investigates key methodological issues related to assessing the outcomes—namely, the number of measures, objective versus subjective measures, and respondent types. Further, this paper examines differences in methods published across journals. The reviewed studies use methods ranging from best-practice methods published in the highest-ranked journals to those associated with biased assessments. The review reveals an inconsistency in sales research as many reviewed studies use methods that previous research has associated with biases, for example, the use of few measures, subjective measures, salespeople's self-ratings, and single-source ratings. This paper contributes to future sales research by proposing guidelines for improved methods to assess the outcomes.

The fourth paper investigates the data sources used to assess the various outcomes from B2B selling. The evaluation reveals the widespread use of salespeople and sales managers to rate economic outcomes and outcomes related to salespeople's customer interactions. These are among the most critical outcomes from B2B selling, but company records and customers, respectively, can provide considerably more reliable and valid assessments of these outcomes than salespeople and sales managers. This paper contributes by suggesting the most reliable and valid data sources to assess specific types of outcomes from B2B selling.

In summary, this thesis shows the large variety of quality and sophistication in the methods to assess outcomes from B2B selling. Moreover, this thesis reveals the widespread use of methods that, according to previous research, do not provide the most reliable and valid assessments—for example, the use of few revenue-focused measures, subjective measures, self-ratings, and single-source measures as well as a mismatch between data sources and collected measures. This finding indicates the need for many researchers to

reevaluate their methods. Further, this finding appeals to a future debate and research on the methodological warnings and recommendations relevant to sales research. This thesis contributes to such future debate and research by suggesting theoretical frameworks, guidelines, and future research directions to improve the assessed outcomes from B2B selling.

TABLE OF CONTENTS

1. INTRODUCTION

This first chapter presents the background of this thesis' overarching research question and the research papers comprising this thesis. First, this chapter introduces sales research's need to improve its research design practices and how this thesis aims to contribute to such progress. Next, this chapter defines the central construct in this thesis—B2B sales performance outcomes—and explain the construct's importance for research and management. Further, this chapter outlines possible problems for researchers and managers due to the little guidance on methods to assess these outcomes. Then, this chapter explains the complexity of assessing outcomes from B2B selling.

Next, the overarching research question is presented as well as how this thesis aims to answer this question and which aspects of the respective methods are investigated. Moreover, this chapter introduces two vital concepts for research and measurement quality—validity and reliability—used in this thesis to evaluate researchers' methods to assess these outcomes. Last, this chapter introduces the four research papers and outlines the remaining chapters in this thesis.

1.1. Background of the thesis' research question

1.1.1. How this thesis responds to previous research

As a subdiscipline of marketing, personal selling has increased in importance considerably over the last few decades (Moncrief, Marshall, and Watkins 2000) along with the maturing sales research discipline (Asare, Yang, and Alejandro 2012). As any research discipline matures and expands, a critical

examination of its research methods becomes necessary due to more complex research questions and the need for more sophisticated methods to answer such questions (Crook et al. 2010).

A review of research methods in sales research (Asare et al. 2012) between 1980 and 2008 reveals that sales research needs to improve some of its research design practices. Specifically, the authors encourage future sales research to conceptualize and operationalize constructs. Further, the authors express concerns regarding bias stemming from single-source surveys and urge future sales research to investigate the use of respondent types. The importance of construct operationalizations and concerns regarding single-source bias are supported by Rapp, Gabler, and Ogilvie (2020).

This thesis addresses these issues by investigating the conceptualization, operationalization, data sources, and respondents used to assess an important construct in sales research—B2B sales performance outcomes, also termed outcomes from B2B selling. Outcomes from selling are broadly defined as the outcomes that salespeople produce (Anderson and Oliver 1987) and range from economic outcomes, such as sales revenue (Zallocco, Pullins, and Mallin 2009), to outcomes related to customer interactions, such as customer satisfaction (Wang, Hoegg, and Dahl 2018).

1.1.2. The importance of B2B sales performance outcomes

Organizational performance outcomes are the ultimate dependent variables in just about every management research area (Richard et al. 2009). Likewise, sales performance outcomes are important and frequently used dependent variables (Asare et al. 2012) in the extensive research field investigating antecedents that can influence and improve selling (Limbu et al. 2016; Verbeke, Dietz, and Verwaal 2011).

The frequent use of sales performance outcomes in sales research demonstrates the importance of these outcomes. Asare et al. (2012) review 1,346 empirical sales research studies published between 1980 and 2008 and find that 18% of these studies use sales performance outcomes as dependent variables. Further, this review reveals the increasing use of these outcomes as dependent variables: Among the studies published by the end of the review period, 26% use sales performance outcomes as dependent variables.

The importance of sales performance outcomes is also related to the outcomes' function in sales research. Researching the antecedents to improved selling is a widespread goal among sales researchers (Limbu et al. 2016). When investigating such antecedents, sales performance outcomes are used as dependent variables to identify the effect of or response to a change in antecedents (Robson and McCartan 2016). By detecting such relationships (covariance), researchers can identify antecedents that can improve selling (Ohiomah, Benyoucef, and Andreev 2020). Consequently, the reliability and validity of the assessed antecedents and outcomes influence the reliability of the estimated covariances (Hair et al. 2010). Reliably and validly assessed outcomes are therefore fundamental for identifying dependable antecedents to improved selling (Rapp et al. 2020; Ghauri and Grønhaug 2002). A recent report reveals that fewer than 20% of firms investing in sales enablement were able to effectively determine their return on investment (Miller Heiman Group 2018), thus demonstrating the importance of research on these antecedents and the measurements conducted in such research (Rangarajan et al. 2020).

For sales managers, assessing outcomes from selling is essential (Zallocco et al. 2009) because of the strong managerial focus on optimizing sales outcomes (Zoltners, Sinha, and Lorimer 2008). In many firms, personal selling is an essential part of marketing, ultimately judged by its contributions to firms'

3

overall organizational performance. Reliable and valid measures of performance outcomes are essential for evaluating firms' and managers' specific actions (Richard et al. 2009) as well as for detecting low performance on essential outcomes and determining necessary managerial actions to improve such performance (MacInnis 2011).

1.1.3. The lack of guidance on methods and probable consequences

Even though the methods used to assess variables are important for research quality (Ghauri and Grønhaug 2002), previous research offers little guidance on methods to assess outcomes from selling—for example, which outcomes to measure (Siguaw, Kimes, and Gassenheimer 2003). Further, to the best of my knowledge, a review of the methods researchers use to assess sales performance outcomes has not been conducted. This lack of such a review may have led to this little guidance on methods and stands in contrast to the reviews (e.g., Herjanto and Franklin 2019) and meta-analyses (e.g., Ohiomah et al. 2020; Verbeke et al. 2011; Albers, Mantrala, and Sridhar 2010; Churchill et al. 1985) on the antecedents of sales performance.

The meta-analyses on the antecedents of sales performance use different outcomes in their analyses. While Ohiomah et al. (2020) and Albers et al. (2010) only use economic outcomes, such as sales revenue and profits, Herjanto and Franklin (2019) also use outcomes related to buyer-seller relationships. Verbeke et al. (2011) take one step further and call upon future researchers to address the fundamental question of what constitutes sales performance outcomes in today's economy. The present thesis responds to this question by suggesting a conceptualization of the outcomes from B2B selling.

The little guidance on methods to assess these outcomes may create serious problems for sales researchers. Regarding measures, it is advisable to

use multiple measures to capture the different types of outcomes desired from selling (Henard and Szymanski 2001; Churchill et al. 1985). Thus, without a proper conceptualization of the outcomes, researchers may use measures that cannot capture the essential outcomes from the specific type of selling investigated (Richard et al. 2009). If essential outcomes remain unobserved, subsequential model testing may be inappropriate, and the results can lead to incorrect conclusions (Fornell and Larcker 1981). For example, salesforce incentives may positively influence short-term sales revenue but may negatively influence long-term customer relationships (Zoltners, Sinha, and Lorimer 2012). Thus, if incentive research assesses the outcomes by solely measuring sales revenue, the incentives' possible adverse effects on other critical outcomes will remain unobserved, and the conclusions may be incorrect (Fornell and Larcker 1981). Further, including insufficient measures in research models may lead to research models that are too simple for our complex reality (MacInnis 2011). Such simplified models may only provide a partial understanding of the research problem being studied and likely generate deficient conclusions (Hult et al. 2008; Richard et al. 2009).

Further, the little guidance may cause researchers to fail to assess outcomes that managers deem essential, which may in turn threaten sales research's managerial relevance and applicability (Zallocco et al. 2009; Richard et al. 2009). Indeed, previous research outlines an apparent gap between how sales researchers and practitioners view sales performance outcomes (Zallocco et al. 2009). Furthermore, researchers may assess different outcomes across studies, thereby making it difficult, if not impossible, to synthesize findings across studies and achieve cumulative knowledge building (Katsikeas et al. 2016). Also, the use of different outcomes across studies limits researchers' ability to classify outcomes in meta-analyses and investigate how antecedents

may influence specific types of sales performance outcomes (Verbeke et al. 2011).

Finally, regarding methods to assess the outcomes, researchers may, for example, use less reliable and valid data sources and respondents, which could bias the assessed outcomes. Such biased assessments of the outcomes represent a serious threat to the reliability of research findings and can accentuate inaccurate or less important antecedents of sales performance. In summary, predictions and models are only as strong as the data collected to test them (Rapp et al. 2020). Thus, weak measures and methods to assess the outcomes represent a severe threat to theory testing (Katsikeas et al. 2016) and knowledge building in sales research (Hult et al. 2008).

Sales managers may also suffer from the little guidance on measures and methods to assess outcomes from B2B selling. For example, research reveals that managers lag behind research on sales performance (Zallocco 2009) and may have problems selecting measures to assess sales success (Haines 2004; Ingram et al. 2005). Invalid and biased measures may cause managers to overlook low performance on critical outcomes, which can in turn hinder managerial decisions and actions for improving such performance (MacInnis 2011).

1.1.4. The complexity of assessing B2B sales performance outcomes

Several factors make it complex to assess outcomes from B2B selling. First, these assessments are complicated because B2B selling's strategic role requires B2B salespeople to participate in numerous activities and produce multiple types of outcomes (Cron, Baldauf, and Leigh 2014). This multiplicity of outcomes is confirmed by sales managers and salespeople surveyed in two studies suggesting 19 (Zallocco et al. 2009) and 31 (Behrman and Perreault 1982) relevant outcomes to assess. Consequently, assessing multiple types of outcomes requires a set of measures reflecting these outcomes and data sources or respondents to provide reliable and valid measures (Groves et al. 2009; Ghauri and Grønhaug 2002).

Second, these assessments may be complicated by the dynamics and fundamental changes over the last few decades (Cuevas 2018) in external and internal organizational environments, setting new and rising standards for the sales profession (Jones et al. 2005). Further, because of the growing recognition of the importance of customer satisfaction, customer loyalty, and long-term customer relationship management, today's salespeople are asked to do more, and the job has become more complex. Thus, firms look beyond the transaction-based concept of immediate sales revenue when measuring and evaluating sales performance outcomes (Zallocco et al. 2009).

Third, these assessments are complex because B2B selling can take various forms across different sales contexts, such as different industries, products, and/or organizational philosophies (Singh and Abraham 2010). Different sales contexts may require different outcomes to be produced. Thus, the specific context should influence which outcomes should be assessed (Richard et al. 2009). B2B selling often takes two primary forms: transactional and consultative B2B selling (Davie, Stephenson, and Valdivieso De Uster 2010).

While transactional B2B selling typically involves selling off-the-shelf products (Parvinen et al. 2013), consultative B2B selling typically involves customizing solutions. Such customizing requires, for example, co-creation (Töytäri and Rajala 2015) and customer relationships (Storbacka et al. 2009), which may lead to a more diverse set of outcomes relevant for assessment compared to transactional B2B selling.

Fourth, these assessments are complex because of the little guidance from previous research on appropriate methods to assess outcomes from selling. Further, the methods relevant to assess the outcomes are treated inconsistently in sales research as the methods frequently used in published studies are simultaneously criticized in the literature for often causing biased assessments. Previous research suggests that such bias is associated with, for example, the use of too few measures (Hult et al. 2008; Richard et al. 2009), subjective measures (Rich et al. 1999; Jaramillo, Carrillat, and Locander 2005), single types of respondents (Jap and Anderson 2004; Hulland, Baumgartner, and Smith 2018), and salespeople's self-ratings (Rich et al. 1999; Paulhus 2002; Jaramillo et al. 2005; Tourangeau and Yan 2007; Steenkamp, De Jong, and Baumgartner 2010). In summary, the complexity of assessing outcomes from B2B selling enhances the importance of the present thesis' investigation of how these outcomes are assessed.

1.1.5. The overarching research question and how this thesis answers it

The importance of assessing outcomes from B2B selling and the potential problems from assessments with weak reliability and validity lead to the overarching research question of this thesis: how do researchers assess outcomes from B2B selling?

To answer this research question, this thesis investigates how researchers attend to the following methodological issues related to assessing outcomes from B2B selling: the conceptualization of the outcomes and operationalization of measures, the number of measures and types of measures (objective versus subjective measures), and the types of data sources and respondents.

These methodological issues are examined using a literature review, more precisely termed a methodological literature review, as this is the most effective way to become familiar with research methods (Onwuegbuzie and Frels 2016). Further, a literature review is an effective tool for identifying conflicts and gaps in research (Boot, Sutton, and Papaioannou 2016) as well as issues that can improve research (Onwuegbuzie and Frels 2016). Furthermore, a literature review can be used to develop theoretical frameworks and guidelines to improve future research (Snyder 2019).

The present thesis is based on a systematic literature review in contrast to a traditional (scoping and narrative) or integrative literature review (Onwuegbuzie and Frels 2016). There are multiple reasons for this. First, a systematic literature review aims to identify all relevant studies (Jesson, Matheson, and Lacey 2012) and may therefore have stronger internal validity by avoiding bias from subjectively selecting studies (Boot et al. 2016) or only reviewing single studies (Jesson et al. 2012). Further, a systematic review enables tabular features, making it easier to interpret large amounts of data (Boot et al. 2016). Last, a systematic review includes transparent methods for collecting, including, and evaluating studies (Jesson et al. 2012), ensuring that the conclusions are grounded in the gathered data and not fabricated (Boot et al. 2016).

The studies included in the present review assess sales performance outcomes as dependent variables and only investigate B2B selling because of the differences between B2B and business-to-consumer (B2C) selling (Lilien 2016). Further, the reviewed studies are solely quantitative because of the dominance of quantitative studies in sales research (Asare et al. 2012) and the differences between quantitative and qualitative research methods (Onwuegbuzie and Frels 2016; Ghauri and Grønhaug 2002). Although some qualitative studies include quantified data and that qualitative data can be coded and quantified to allow statistical analysis, quantitative and qualitative research differ regarding their perspectives on knowledge, research objectives, information of interest, measures, and data collection (i.e., how and where to collect data) (Ghauri and Grønhaug 2002). As such, adding qualitative studies into this review may have made a manageable review in terms of time and resources unfeasible, created distractions from the main focus of the review, and threatened the accuracy of the data collection and data analysis (Boot et al. 2016). Last, as the overarching research question asks how researchers assess the outcomes, the included studies were published relatively recently (2001–2015) in contrast to a historical examination far back in time. The review includes 139 studies that fulfill these inclusion criteria.

The measures and methods used to assess the outcomes in the reviewed studies are evaluated in relation to two primary issues associated with research quality—the validity and reliability of assessments (e.g., Seale 2009; McGivern 2013). Validity refers to the degree to which research designs, measures, and methods deliver accurate and unambiguous evidence (McGivern 2013) and reflects whether the reported results are true (Seale 2009). Further, validity refers to the degree to which a study measures what it intends to measure (e.g., McGivern 2013).

There are multiple types of validity evaluations (e.g., Voorhees et al. 2016; Ghauri and Grønhaug 2002). However, researchers often evaluate research quality in terms of internal and external validity. Internal validity refers to the extent to which causal relationships between variables can be inferred, while external validity refers to the extent to which findings can be generalized to populations and other settings (Seale 2009; McGivern 2013; Ghauri and Grønhaug 2002). Researchers can use a third type of validity test—measurement validity (Seale 2009)—which refers to the degree to which measures successfully measures concepts (Seale 2009; Ghauri and Grønhaug 2002).

Evaluating measurement validity is important when assessing the construct of outcomes from B2B selling because it is an abstract construct that cannot be directly observed because of its multiple components (revenue, profit, customer satisfaction, etc.) (Groves et al. 2009). The most crucial form of validity for such an abstract construct is construct validity (Ghauri and Grønhaug 2002), which refers to the extent to which measures reflect or represent the components constituting the construct (Groves et al. 2009; Ghauri and Grønhaug 2002). Thus, if a study lacks construct validity, the findings are worthless, and the internal and external validity of the research findings are also destroyed (Ghauri and Grønhaug 2002).

Among, the several types of construct validity evaluations (Seale 2009; Ghauri and Grønhaug 2002), this thesis evaluates construct validity by determining how well the measures conform to expectations from previous research/theory (Seale 2009). This evaluation is applied when examining researchers' use of measures to assess the outcomes.

The validity-related objective of reporting accurate research results (Seale 2009) also relies on measures' reliability, defined as measures' stability across repetitive assessments (Groves et al. 2009; Ghauri and Grønhaug 2002). The differences between an observed score in a survey and the "true" score is systematic bias and random error (Ghauri and Grønhaug 2002). Systematic bias can occur from respondents' stabile underreporting and overreporting (Groves et al. 2009), for example, self-ratings that tend to overreport personal achievements. Random error can occur from personal and situational factors (Groves et al. 2009), for example, ratings influenced by a positive or negative incident close to the survey. Systematic bias and random error represent the criteria used to evaluate reliability associated with objective and subjective measures and the various data sources and respondents used to assess the outcomes.

1.2. Overview of the four research papers

1.2.1. How the research papers answer the thesis' research question

Each of the four research papers answers the overarching research question, with separate research questions investigating particular methodological issues related to assessing these outcomes, as shown in Figure 1 and explained in later paragraphs.

Figure 1. The four research papers
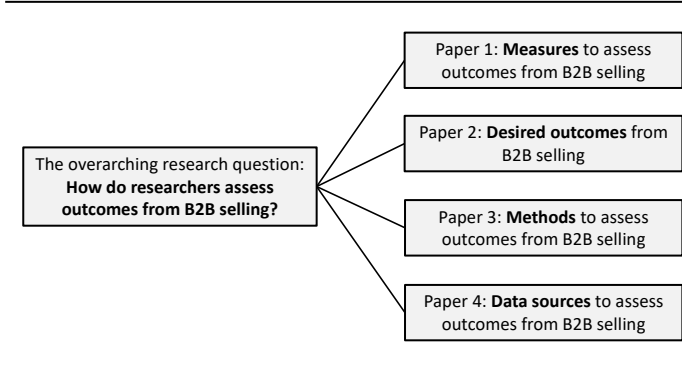
These methodological investigations required collecting various data categories from the reviewed studies to explore how sales researchers assess the outcomes. The data was organized in accordance with the investigated methodological issues, which resulted in a unique and large dataset. The analysis of the data representing each methodological issue required specific data analysis. Further, the evaluation of the results had to be conducted in relation to previous research particularly relevant to each methodological issue investigated. Thus, a thorough investigation meant that each methodological issue needed to be investigated in a separate research paper, leading to four research papers. The following paragraphs introduce the research questions and contributions of each research paper.

The first paper investigates the measures used to assess the outcomes from B2B selling and therefore explores an essential aspect of how the outcomes are assessed. The paper aims to answer the following research questions: *which measured do researchers use to assess outcomes from B2B selling?* The research question was answered through a systematic review of the measures used to assess outcomes from B2B selling in 139 studies published in 17 journals. The paper shows the large variety of measures used to

assess the outcomes and the most frequent sets of measures used in the reviewed studies. The paper contributes by evaluating the construct validity provided by these measures by examining how the measures correspond to previous research on outcomes from B2B selling. Further, the paper contributes recommendations on how researchers can improve these measures.

The second paper responds to the finding in the first paper, which reveals that there are currently no complete theoretical frameworks suggesting which outcomes are desired from B2B selling. The second paper suggests such a theoretical framework by answering the following research question: *which outcomes are desired from B2B selling?* The paper answers this question by organizing the outcomes measured in the reviewed studies, thereby creating and contributing a complete theoretical framework conceptualizing the desired outcomes from B2B selling. This framework can be used for multiple purposes and is a proper tool for operationalizing measures to assess outcomes from B2B selling.

The third paper investigates three key methodological issues vital for assessing outcomes from B2B selling and addresses two research questions: *how appropriate are the methods researchers use to assess B2B sales performance outcomes, and are there differences in methods published in different journals?* The paper examines the following three key methodological issues: the number of measures, the type(s) of measures (objective and subjective, and the type(s) of respondents. Further, the paper examines differences in methods published across the 17 journals that contributed studies to the review. The examination reveals substantial variation in the quality and sophistication of methods—from those that may provide biased assessments to best-practice methods published in the highest-ranked

journals. The paper contributes by evaluating the methods and providing guidelines on improved methods to assess the outcomes.

The fourth paper investigates the data sources used to assess various outcomes from B2B selling by addressing the following research question: *which data sources do researchers use to assess the various types of outcomes from B2B selling, and which data sources are appropriate to assess the various types of outcomes?* The paper answers these research questions by examining the data sources (e.g., company records, sales managers, salespeople, and customers) used to assess various types of outcomes in the reviewed studies. The examination reveals the widespread use of data sources that do not provide the most reliable and valid assessments of the outcomes they assess. The paper contributes by suggesting the most appropriate data sources to assess various types of outcomes.

Table 1. Overview of the research papers in the thesis

	Paper 1	Paper 2	Paper 3	Paper 4
Title	Measures to assess B2B sales performance outcomes: A systematic review and future directions	Desired outcomes from B2B selling: A systematic review and conceptualization	Methods to assess outcomes from B2B selling: A systematic review, cross-journal examination, and guidelines	Data sources to assess sales performance outcomes
Author(s)	Seljeseth, Korneliussen, Greenacre	Seljeseth	Seljeseth, Korneliussen, Greenacre	Seljeseth
Research question(s)	Which measures do researchers use to assess outcomes from B2B selling?	Which outcomes are desired from B2B selling?	How appropriate are the methods researchers use to assess B2B sales performance outcomes, and are there differences in methods published in different journals?	Which data sources do researchers use to assess the various outcomes from B2B selling, and do researchers use the most valid data sources to assess the various outcomes?
Method(s)	Cluster analysis	Quantitative and conceptual	Cluster analysis and correspondence analysis	Cross-tabulations and ranking
Key findings/ contributions	Researchers use 151 different measures to assess the outcomes, and seven sets of measures are frequently used. A large portion of the studies measure only sales revenue and thus disregard the multiple types of outcomes from B2B selling. The paper suggest how researchers can improve the measures used to assess the outcomes.	The study suggests the first complete theoretical framework conceptualizing the outcomes desired from B2B selling with the B2B Sales Performance Outcomes Chain. The chain suggests seven main types/categories and 21 subtypes/subcategories of outcomes desired from B2B selling.	Researchers use methods with substantial variations in quality and sophistication—from methods that may provide biased assessments to best-practice methods published in the highest-ranked journals. This study suggest guidelines on methods to assess the outcomes.	There is a widespread use of data sources that do not provide the most reliable and valid assessments of the outcomes. This study suggests guidelines on the most appropriate data sources to assess the various outcomes from selling.
Publication status	Previous versions of the paper are presented at the 48th EMAC Annual Conference 2019 and the 16th Conference of the International Federation of Classification Societies 2019. Previous versions submitted to Industrial Marketing Management (ABS level 3) and Journal of the Academy of Marketing Science (ABS level 4). After passing the review processes it was not accepted for publication. Also, previous version submitted to Journal of Personal Selling and Sales Management (ABS level 2). The editor invited us to resubmit a new version of this paper, which will be done in 2021.	Preveous version presented at the 20th Conference of the European Association for Education and Research in Commercial Distribution 2019. Previous version submitted to Industrial Marketing Management (ABS level 3). After passing the review process it was not accepted for publication. The paper will be submitted to Journal of Business and Industrial Marketing (ABS level 3) in 2021.	Previous version submitted to Journal of the Academy of Marketing Science (ABS level 4). After passing the review process it was not accepted for publication. The paper will be submitted to Industrial Marketing Management (ABS level 3) in 2021.	The paper is submitted to Journal of Personal Selling and Sales Management (ABS level 2) in December 2020.

1.2.2. The interrelatedness of the research papers

The four research papers comprising this thesis are interrelated in the following ways: The first research paper investigates the measures used to assess outcomes from B2B selling. The search for appropriate theoretical frameworks to evaluate the measures used by the reviewed studies reveals a lack of frameworks conceptualizing outcomes desired from B2B selling.

The second research paper contributes to overcoming this lack of frameworks by suggesting a theoretical framework that conceptualizes desired outcomes from B2B selling and the construct of B2B sales performance outcomes. Consequently, the first and second research papers investigate two interrelated theoretical and methodological issues. The first paper investigates how researchers operationalize measures of outcomes from B2B selling, while the second paper develops a conceptualization of these outcomes. This conceptualization can be used to operationalize measures to assess the outcomes, which is the subject of the first research paper.

Once the investigation of the measures and conceptualizing the outcomes were completed, the third research paper takes a relevant next step by investigating three key methodological issues vital for assessing the outcomes: the number of measures, the types of measures (objective and subjective measures), and the types of respondents. Further, the third research paper examines how journals attend to these three key methodological issues by examining differences in methods used in studies published in different journals.

Regarding the number of measures used to assess the outcomes, the first and third research papers are interrelated. The first paper examines the number of measures used to assess the outcomes and reveals that researchers

use from one to 30 measures. This large variety of measures and the importance of the number of measures for assessing the outcomes supported the inclusion of this issue in the third research paper's examination of methods published across journals. Thus, the third paper examines differences in the number of measures across journals.

The third research paper examines the use of objective measures from company records and subjective measures from multiple types of respondents. The result from this examination is evaluated in relation to previous research, which show considerable differences in these data sources' ability to provide reliable and valid assessments of outcomes from selling. Each data source may provide reliable and valid assessments of certain outcomes while likely providing less reliable and valid assessments of other outcomes. These differences reveal the need to investigate which data sources are used to assess the various types of outcomes.

This investigation is conducted in the fourth research paper. Thus, the third and fourth research papers are interrelated as they investigate objective and subjective measures and various data sources and respondents used to assess the outcomes. The fourth research paper goes one step further and "connects" these data sources with the measured outcomes. More concretely, the fourth research paper examines which data sources (e.g., company records, sales managers, salespeople, and customers) are used to assess various outcomes in the reviewed studies. Previous research on various data sources' ability to assess different types of outcomes reliably and validly are used to evaluate and suggest the most appropriate data sources to assess various types of outcomes.

1.3. Outline of the thesis

This thesis is structured into six chapters. Chapter 2 defines the key term B2B sales performance outcomes and provides the theoretical background for assessing these outcomes. Chapter 3 presents the methods used in the thesis and research papers, including philosophical approaches, as well as how the systematic review was conducted and how the data was analyzed. Chapter 3 also evaluates the validity, reliability, and ethics associated with the research conducted in the thesis. Chapter 4 summarizes the findings and contributions of the four research papers. Chapter 5 discusses conclusions, implications for researchers and managers, limitations, and suggestions for future research. Then, the references are outlined and Chapter 6 presents the four research papers composing this thesis.

2. THEORETICAL BACKGROUND

This chapter introduces relevant research related to the assessing of B2B sales performance outcomes. First, this chapter introduces key definitions and the nature of B2B selling, outcomes from B2B selling, and measures to assess these outcomes, which are applied in the first and second research papers. The chapter then introduces relevant research on the number of measures, objective and subjective measures, and data sources and respondents used to assess outcomes from B2B selling, which are applied in the second and third research papers.

2.1. Definitions

Sales performance outcomes can be broadly defined as the outcomes that salespeople produce (Anderson and Oliver 1987) and range from economic outcomes, such as sales revenue, to outcomes associated with salespeople's customer interactions, such as customer satisfaction. Despite the outcomes' frequent and increasing use as dependent variables in sales research (Asare et al. 2012) and even though researchers have discussed numerous measures to assess the outcomes, no theoretical solution has yet been suggested to measure the outcomes (Siguaw et al. 2003).

The conceptualization of the outcomes-from-B2B-selling construct has a widespread impact on how it should be assessed. First, a conceptualization outlines the components of a construct (Groves et al. 2009; Ghauri and Grønhaug 2002), which, in this case, means outlining the various types of outcomes. This outlining provides guidance for operationalizing measures of the outcomes (Ghauri and Grønhaug 2002) and shows how the

conceptualization and operationalization of this construct are interconnected (Groves et al. 2009; Ghauri and Grønhaug 2002). Thus, to select measures to assess outcomes from B2B selling, one needs to understand the nature of B2B selling and subsequently conceptualize the desired outcomes from this type of selling. In turn, these outcomes indicate which data sources or respondents are relevant or most appropriate to assess the themselves (Groves et al. 2009). For example, using company records to assess sales revenue, and customers to assess customer satisfaction.

2.2. The nature of B2B selling

B2B and B2C interactions are both parts of complex marketing contexts (Gummesson and Polese 2009). Like B2B customers, consumers/B2C customers can buy complex and customized products and services and operate in complex networks and relationships with their families, friends, and numerous suppliers. Further, similar to participants in B2B customers' buying centers, consumers often interact with household "buying centers" comprising family members who act as buyers, payers, users, and shareholders (Gummesson and Polese 2009).

However, B2B selling differs from B2C selling in several ways. First, B2B salespeople often work with value chain intermediaries' networks, while B2C salespeople work with end consumers. Thus, B2B marketers face fewer customers and engage in far larger transactions in terms of economic value compared to B2C marketers. To a more considerable degree, these larger transactions are technical and economic value propositions rather than perceptual brand value propositions (Lilien 2016). Thus, B2B selling, as opposed to B2C selling, is likely to involve more rational buying criteria, more complex

and lengthy decision processes (Dawes, Lee, and Dowling 1998; Manning, Reece, and Ahearne 2010), more people (Gartner 2019), and better trained buying-decision participants (Dawes et al. 1998; Manning et al. 2010).

Whereas B2C selling often takes place within organizational boundaries (e.g., retail stores), B2B salespeople often operate as "boundary spanners" inside and outside their selling companies (Nygaard and Dahlstrom 2002). Thus, compared to B2C selling, B2B selling involves a far more extensive range of stakeholders, such as financial analysts, purchasing agents, engineers, manufacturing managers, and lawyers (Lilien 2016). In summary, these differences between B2B and B2C selling suggest that B2B selling is required to produce more complex and numerous sets of outcomes compared to B2C selling.

2.3. Outcomes from B2B selling

As personal selling is a marketing function, theory on marketing performance outcomes can help conceptualize outcomes from B2B selling and operationalize valid measures to assess such outcomes (Groves et al. 2009; Seale 2009). Katsikeas et al. (2016) provide a theoretical framework on marketing performance outcomes that largely correspond with the outcomes desired from B2B selling, as suggested in prior research (e.g., Cuevas 2018). Figure 2 exhibits this theoretical framework.

Figure 2. The Marketing Performance Outcome Chain
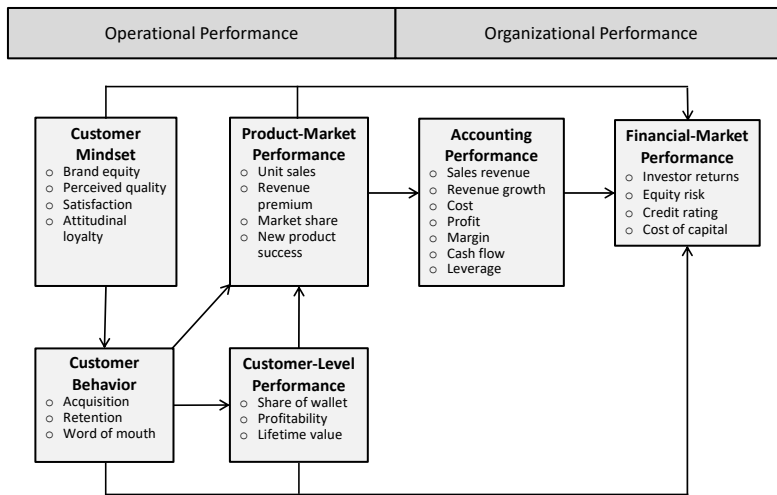
Adapted from Katsikeas et al. (2016)

This framework suggests six main categories of outcomes divided into operational and organizational performance. The outcomes can also be divided into outcomes directly beneficial for customers and the selling company (Zoltners et al. 2008). The customer outcomes are shown in the customer mindset category, and the company outcomes are shown in the five remaining main categories.

The framework organizes the outcomes in a value chain structure that shows how various outcomes from marketing relate to and influence each other. The framework illustrates well how B2B marketing and selling outcomes influence each other, for example, how customer mindset, such as customer satisfaction, influences customer behavior, such as customer retention (e.g., Rauyruen and Miller 2007; Blocker et al. 2011). Further, the framework shows how customer behavior, such as customer retention, influences customer-level performance, such as customers' lifetime value; product-market performance,

such as unit sales (Rauyruen and Miller 2007); and accounting performance, such as sales revenue, cost, and profit (Lam et al. 2004; Rauyruen and Miller 2007). Last, the framework shows how accounting-performance, such as sales revenue, cost, and profits, are antecedents of financial-market performance, such as investor returns.

2.4. Measures to assess outcomes from B2B selling

Because B2B selling is expected to produce multiple types of outcomes, the outcomes-from-B2B-selling construct is an abstract construct that cannot be directly observed (Groves et al. 2009). To assess such an abstract construct, one must use measures that reflect the multiple components (i.e., types of outcomes) that constitute the construct, also termed construct validity (Groves et al. 2009; Ghauri and Grønhaug 2002). Thus, research that conceptualizes the various types of outcomes desired from B2B selling could provide a basis for the operationalization of measures. However, to the best of my knowledge, no previous research provides such conceptualization.

Therefore, to operationalize measures of these outcomes, one has to rely on several research studies. Two previous research studies attempt to operationalize these measures by interviewing practitioners in B2B selling—namely, sales managers and salespeople (Zallocco et al. 2009; Behrman and Perreault 1982). These two studies confirm the need for multiple measures. However, the studies suggest different measures and thus do not contribute to a consensus on valid measures nor on which outcomes are desired from B2B selling.

The primary measure of sales performance outcomes should be sales revenue, which is the most important outcome from selling (Zallocco et al.

2009). However, the salesperson's expanded role over the last few decades (Cuevas 2018) requires measures with a more long-term focus, such as relationship expansion and customer satisfaction (Hughes and Ogilvie 2020).

Further, the measures must be adapted to the outcomes desired from the particular type of selling investigated and the growing diversity of customer expectations in B2B markets. This diversity suggests that B2B salespeople should conduct two main types of selling: transactional and consultative B2B selling (Davie et al. 2010). Transactional B2B selling typically involves selling off-the-shelf products (Parvinen et al. 2013), while consultative B2B selling typically involves selling customized solutions (Cuevas 2018). Any selling needs to create basic outcomes, such as service quality (Töytäri and Rajala 2015), offer value (Blocker et al. 2012), customer satisfaction (Wang et al. 2018), and customer loyalty (Lam et al. 2004). However, with consultative B2B selling, the products and services are customized, which requires salespeople to create additional outcomes. Such additional outcomes could include salespeople's and customers' ability to co-create products and services (Töytäri and Rajala 2015) and salespeople's ability to cooperate with other departments within the selling firm to create customized solutions (Steward et al. 2010; Guenzi and Panzeri 2015). Further, customization requires strong buyer-seller relationships (Storbacka et al. 2009; Mullins et al. 2014) and market intelligence regarding customer needs (Flint, Woodruff, and Gardial 2002). Consequently, to assess outcomes from consultative B2B selling, one needs a larger variety and number of measures compared to transactional B2B selling.

2.5. Number of measures to assess outcomes from B2B selling

Previous paragraphs have introduced research on outcomes from B2B selling, showing that B2B salespeople are required to produce multiple types of outcomes (Cuevas 2018). These multiple types of outcomes indicate that the outcomes-from-B2B-selling construct is a multi-dimensional abstract construct that needs multiple measures to be assessed (Martinez-Martin 2010; Groves et al. 2009; Ghauri and Grønhaug 2002). Practitioners confirm this multiplicity by suggesting 19 (Zallocco et al. 2009) and 31 (Behrman and Perreault 1982) measures to assess outcomes from B2B selling. Consequently, using only one or a few measures to assess B2B selling outcomes typically provides only a partial assessment of the outcomes from B2B selling.

2.6. Types of measures to assess outcomes from B2B selling

Sales performance outcomes can be assessed using both objective and subjective measures. Objective measures usually consist of numbers extracted from company records of "hard" economic outcomes, such as sales revenue, sales quota compliance, and profits (Churchill et al. 1985). Subjective measures are ratings from sales managers, salespeople, and customers and are valid to assess "soft" outcomes associated with salespeople's customer interactions, such as offer value, customer satisfaction, and customer relationships.

Two meta-analyses show that subjective measures assess sales performance outcomes quite differently than objective measures, revealing a shared variance of only 20% (Rich et al. 1999) and 11.6% (Jaramillo et al. 2005) between subjective and objective measures of the outcomes. As such, subjective measures may be poor indicators of economic outcomes, so it is

preferable to assess such outcomes using objective measures from company records when available (Dess and Robinson 1984).

However, subjective measures are valuable and relevant, but they should be combined with an additional source of measures (Rapp et al. 2020). The complementary benefits of subjective and objective measures make it valuable to combine both types of measures when assessing sales performance outcomes (Bagozzi, Verbeke, and Gavino 2003) as subjective measures are vital to assess soft outcomes while objective measures provide reliable assessments of hard economic outcomes.

2.7. Data sources and respondent to assess outcomes from B2B selling

Four data sources are used to assess sales performance outcomes: company records, sales managers, salespeople, and customers. Thus, when using subjective measures, the choice of data source(s) involves selecting which types of respondents to use.

Single or multiple data sources. The choice of data source involves choosing whether to use single or multiple data sources in the same study. Multiple data sources can be combinations of objective and subjective measures or combinations of several types of respondents. The combining of ratings from multiple types of respondents is recommended to attenuate respondent bias (Hulland et al. 2018) because ratings from a single type of respondent at a specific point in time may entail such bias (Jap and Anderson 2004). Schmitz, Lee, and Lilien (2014) recommend using three types of respondents—salespeople, sales managers, and customers—to overcome respondent bias. Further, when assessing buyer-seller relationships, using

multiple types of respondents, such as salespeople and customers, is recommended to provide richer assessments (Hughes, Le Bon, and Rapp 2013; Hulland et al. 2018).

Company records. Company records are preferred over respondent ratings to assess economic outcomes due to their more robust reliability (Dess and Robinson 1984). Company records are typically subject to detailed government regulations regarding accounting, auditing, and reporting, which may be one cause for this robust reliability. The main limitation of objective measures is their inability to assess soft outcomes from selling, such as customer satisfaction and customer relationships. However, company records can provide reliable assessments of actual customer loyalty from customer repurchase data over time.

Company records may not always provide 100% accurate assessments as they can be manipulated on purpose or by accident, at least in the short term. For example, salespeople addicted to bonuses may manipulate sales reports on purpose (Zoltners et al. 2012), while incorrect accruals of sales revenue could accidentally bias company records.

Salespeople. Salespeople have the best insights into their tactics, efforts, and interactions with customers, as well as their interactions with departments, managers, and colleagues within their own company. However, salespeople's self-ratings are associated with respondent bias from various causes. First, the meta-analysis by Jaramillo et al. (2005) reveals that salespeople rate sales performance outcomes quite differently from objective measures. Thus, it is preferable to avoid using salespeople's self-ratings to assess economic outcomes when such measures are available from company records (Dess and Robinson 1984). Second, this meta-analysis reveals that salespeople rate sales performance outcomes differently from sales managers.

This difference may be caused by salespeople's narrower definition of the outcomes compared to sales managers (Rich et al. 1999). Third, self-ratings from salespeople are associated with respondent bias from socially desirable responding (Steenkamp et al. 2010), which may cause salespeople to rate sales performance outcomes better than they are in reality (Paulhus 2002; Tourangeau and Yan 2007).

Sales managers. Sales managers are likely to be knowledgeable about sales performance outcomes as they have easy access to relevant information sources, such as company records, customers, and salespeople. Further, sales managers expect a wider variety of outcomes than salespeople (Rich et al. 1999), which may strengthen the validity of sales managers' ratings. Furthermore, sales managers typically have a bird's-eye view to compare their subordinates' outcomes (Jaramillo et al. 2005).

Similar to salespeople, sales managers rate sales performance outcomes differently than objective measures, as shown in the meta-analysis by Jaramillo et al. (2005). Thus, it is also preferable to avoid using sales managers' ratings to assess economic outcomes when such measures are available from company records. However, sales managers' ratings are shown to be twice as reliable as salespeople's ratings to assess economic outcomes (Jaramillo et al. 2005). Further, sales managers rate salespeople's customer interactions, such as trustworthiness, technical knowledge, product knowledge, and availability, quite different from how customers rate these outcomes. Thus, it is preferable to refrain from using sales managers' ratings to assess such outcomes when customer ratings are available (Cannon and Spiro 1991).

Customers. Customers are naturally in the best position to provide reliable evaluations of salespeople's customer interactions, such as gaining trust, providing advice, and providing high-quality customer service (Cannon

and Spiro 1991). Further, customers may be the best data source to evaluate salespeople as salespeople are required to satisfy customer needs (Lambert, Sharma, and Levy 1997). The high importance of salespeople's customer interactions (e.g., Williams and Attaway 1996; Wang, Dou, and Zhou 2012; Wang et al. 2018) suggests using the most reliable and valid data source to assess such outcomes, which is customers.

3. METHODS

This chapter presents the methods used in this thesis. First, this chapter explains why a systematic literature review is appropriate to answer the research questions. Next, the chapter discusses the philosophy of science associated with a systematic literature review and this thesis before presenting the research steps involved in the review. Then, the chapter describes the data analyses in each research paper and evaluates the validity and reliability of the thesis before finally addressing ethical considerations regarding the thesis.

3.1. Using a systematic literature review to answer the research questions

The overarching research question and the research questions in the four research papers require an examination of how researchers assess outcomes from B2B selling. This examination of researchers' methods can be conducted with a literature review (Cooper 2010; Boot et al. 2016) for the following reasons. First, a literature review is the most effective way to become familiar with the research methods used in previous research (Onwuegbuzie and Frels 2016). Second, a literature review is an effective tool for identifying conflicts and gaps in previous research (Boot et al. 2016) as well as issues that can improve future research (Onwuegbuzie and Frels 2016). Third, a literature review is suitable for developing theoretical frameworks and guidelines to improve future research (Snyder 2019).

A literature review can be conducted as a traditional (scoping) review or as a systematic review (Boot et al. 2016). While traditional literature reviews entail a purposive selection of studies by the reviewer and a discursive

examination, systematic literature reviews aim to identify all relevant studies and examine each study according to predefined criteria (Jesson et al. 2012). A systematic literature review was chosen for the present thesis due to the following strengths. First, a systematic literature review includes organized and transparent methods for collecting, including, synthesizing, and evaluating studies (Jesson et al. 2012). Such transparency makes it easier to judge the methods and findings in this thesis (Boot et al. 2016). Second, such transparency strengthens the conclusions' auditability, ensuring the conclusions are grounded in the gathered data rather than fabricated to support a prior assumption (Boot et al. 2016). Third, a systematic literature review may have stronger internal validity as this approach avoids bias from a subjective selection of studies (Boot et al. 2016) or from only reviewing single studies, which could be done in a traditional review (Jesson et al. 2012). Finally, a systematic review enables graphical and tabular features, making it easier to interpret large amounts of data and findings (Boot et al. 2016), which is the case in the present review and thesis.

A systematic and quantitative literature review applied to answer clearly defined research questions includes a descriptive research design (McGivern 2013; Onwuegbuzie and Frels 2016). The research questions in this thesis' research papers require a literature review that describes the following methodological issues related to assessing the outcomes: the measures, the number of measures, the types of measures (objective and subjective), and the types of respondents and data sources. Consequently, the data collected from all the reviewed studies to describe all these methodological issues resulted in a relatively large dataset. Further, each methodological issue requires specific data analysis and examination, resulting in multiple analyses and evaluations. Furthermore, the findings regarding each methodological issue must be

evaluated in relation to previous research with specific relevance for each issue. Thus, a thorough investigation of how researchers assess outcomes from selling requires investigating these methodological issues over four research papers.

Systematic literature reviews share similarities with meta-analyses regarding their quantitative procedures and statistical analyses. However, meta-analyses emphasize synthesizing the results and findings in studies (e.g., Cooper and Hedges 2009; Cooper 2010), while this thesis requires synthesizing the research methods used in studies, termed a methodological literature review (Onwuegbuzie and Frels 2016). Thus, a systematic review is appropriate over a meta-analysis to answer the research questions of this thesis (Cooper and Hedges 2009; Cooper 2010).

3.2. Philosophy of science

Philosophy of science is the systematic study of scientific activity and knowledge, which includes different scientific paradigms (Gilje and Grimen 1993). These paradigms represent various scientific perspectives on what can be counted as facts (Kuhn 2012) and how to study and understand the world (Patton 2015). Thus, scientific paradigms determine the frameworks and principles for research methods (Guba and Lincoln 1994).

Each scientific paradigm covers three scientific levels. First, the ontology level includes assumptions about how reality actually is and what can be known about it. Second, the epistemology level expresses how we can acquire knowledge about reality. Third, the methodology level includes techniques for generating information about reality (Easterby-Smith, Thorpe, and Jackson 2012).

The present thesis and systematic literature review are rooted in the post-positivistic paradigm. The post-positivistic ontology argues that social science research should be objective despite acknowledging human limitations in providing objective and irrefutable knowledge (Lincoln and Guba 2000). According to Onwuegbuzie and Frels (2016), literature reviews align with the post-positivistic view of objective but imperfect knowledge. According to post-positivists, "the absolute truth" that positivists claim exists is nowhere to be found (Wildemuth 1993).

The post-positivistic drive toward objectivity (Lincoln and Guba 2000) aligns with systematic literature reviews' ability to give the most trustworthy answers to specific review questions (Boot et al. 2016). This trustworthiness is strengthened by the ambition of reviewing all relevant studies in systematic literature reviews (Jesson et al. 2012). This ambition strengthens such reviews' internal validity by reducing bias from the subjective selection of studies (Boot et al. 2016).

The post-positivistic acknowledging of human limitations in providing objective knowledge (Lincoln and Guba 2000) aligns with systematic literature reviews' limitations to providing objective and value-neutral knowledge. These limitations are related to the series of decisions that must be made when conducting such reviews. These decisions are influenced by researchers' perspectives, such as what studies to include and what fragments of the research to emphasize or criticize (Onwuegbuzie and Frels 2016). The present literature review is based on several subjective decisions regarding, for example, study inclusion criteria and the methodological issues to focus on. Thus, rather than attempting to be objective and value-neutral, literature reviewers should strive to be systematic to minimize biases (Onwuegbuzie and Frels 2016).

The post-positivistic epistemology proposes that knowledge is built by adding building blocks, such as generalizations, to the existing edifice of knowledge (Lincoln and Guba 2000). The present systematic review on how the reviewed studies assess outcomes from B2B selling aims to be a generalization of how *all* B2B sales researchers assess these outcomes.

Furthermore, the post-positivistic epistemology establishes that researchers should try to eliminate biases in empirical data (Onwuegbuzie and Frels 2016) to generate findings that are "probably true" (Lincoln and Guba 2000). Such efforts to eliminate biases aligns strongly with the present thesis' primary goal—to contribute to reducing biases and improving the reliability and validity of quantitative assessments in empirical research.

The review reveals the widespread use of subjective ratings from respondents, which may be associated with respondent bias. The present evaluation acknowledges that subjective ratings are based on linguistic interpretations of survey questions (Solberg 2001) and that answers to such questions are often socially constructed (Gilje and Grimen 1993) and can never be absolutely certain (Gilje and Grimen 1993; Slagstad 1995). Further, the present evaluation acknowledges that salespeople's self-ratings of sales performance outcomes may suffer from respondent bias (Podsakoff et al. 2003). Such respondent bias may come from overreporting of socially desirable behaviors and underreporting of behaviors that are socially undesirable (Fishbein and Ajzen 2010).

On the methodological level, post-positivists use both quantitative and qualitative methods (Lincoln and Guba 2000). However, a post-positivistic approach to literature reviews is likely to emphasize quantitative studies (Onwuegbuzie and Frels 2016), which is the scope for the present review. Further, to analyze, handle, and synthesize quantitative data, the obvious

choice it to use quantitative approaches (Boot et al. 2016), which are conducted in the present thesis.

To analyze the quality of research, post-positivists use conventional benchmarks, such as validity, reliability, and objectivity (Lincoln and Guba 2000). Similarly, in this thesis, reliability and validity are used to evaluate the methods used to assess the outcomes from B2B selling. Further, this thesis' push for objectivity is enhanced by following the strict standards of conducting systematic literature reviews (Boot et al. 2016; Littell, Corcoran, and Pillai 2008).

Post-positivism accepts both inductive and deductive research approaches, and this thesis' systematic literature review encompasses a research paradigm of open-mindedness in observations and inductive bottom-up generalizations from a large number of collected data (Ladyman 2002).

3.3. The research steps in the systematic literature review

The present systematic literature review provides data to answer the overarching research question and the research questions in the four research papers. Thus, the review's research steps in the research papers are similar: for example, the collection of studies, inclusion criteria, and data recording. Therefore, the method sections in the four research papers share large similarities. The systematic review is conducted in accordance with the guidelines suggested by Palmatier, Houston, and Hulland (2018) and Littell et al. (2008) and consists of the following research steps.

The first step was to gather empirical studies measuring B2B sales performance outcomes. The included studies had to be published in scientific journals as such studies represent the highest level of research (Nord and Nord

1995; Ngai 2005), and systematic literature reviews tend to focus on the highest-quality research available (Boot et al. 2016). Furthermore, researchers generally use such journals to disseminate studies and acquire knowledge (Nord and Nord 1995; Ngai 2005). Thus, the present review excluded master theses, doctoral dissertations, conference papers, unpublished papers, and textbooks (Ngai 2005).

The primary sources were scientific journals that, according to the review by Asare et al. (2012), publish research on personal selling and sales management. The six journals that Asare et al. (2012) claim publish the highest number of sales research studies were examined issue by issue (i.e., *Journal of Personal Selling & Sales Management, Industrial Marketing Management, Journal of Marketing, Journal of the Academy of Marketing Science, Journal of Business & Industrial Marketing, and Journal of Business Research*). The other eleven journals that publish sales research according to Asare et al. (2012) were examined through an online keyword search on each journal's website using separate keywords like "sales performance," "sales," "selling," and "sales effectiveness" (i.e., *Journal of Marketing Theory and Practice, International Journal of Research in Marketing, Journal of Marketing Research, European Journal of Marketing, Journal of Applied Psychology, Marketing Intelligence & Planning, Journal of Business-to-Business Marketing, Journal of International Marketing, Marketing Science, Psychology & Marketing, Women in Management Review*).

In the next step, the collected studies were examined in relation to the following study inclusion criteria. First and foremost, the studies must assess sales performance outcomes as dependent variables associated with at least one sales-related independent variable (Katsikeas et al. 2016). Researchers use various labels for the sales-performance-outcomes construct; thus, these

various labels were accepted for inclusion, including substitutes for "sales" (e.g., "salesperson," "sales team/force," "key account manager," "sales trainee," and "sales territory") and substitutes for "performance outcomes" (e.g., "outcomes," "productivity," "success," "effectiveness," and "excellence").

Further, the included studies only investigate B2B selling because of its differences from B2C selling (Lilien 2016; Dawes et al. 1998; Manning et al. 2010). Furthermore, the included studies are quantitative (as opposed to qualitative) for the following reasons: Foremost, the scope of the review needs to reflect the review audience (Boot et al. 2016), which is the sales research community. The studies published by this community are 68% purely quantitative and only 6% purely qualitative (Asare et al. 2012), indicating a major focus on quantitative research methods among sales researchers. Additionally, the review excludes qualitative studies to avoid being too wide and unmanageable. Conducting a single, completely comprehensive review is generally not feasible given time and resource constraints. The differences between qualitative and quantitative studies in such a comprehensive review could also create distractions from the review's main focus (Boot et al. 2016). Indeed, a clear distinction between quantitative and qualitative research is their different research methods (Onwuegbuzie and Frels 2016), which is the main focus of this thesis. These differences would have increased the workload at the expense of the selectivity and accuracy of the data analysis and evaluations (Boot et al. 2016). Thus, such differences often require splitting the pool of relevant studies (Boot et al. 2016), which is accomplished by including only quantitative studies in the present review. Lastly, as systematic literature reviews are most closely aligned with the quantitative research tradition, such reviews typically emphasize quantitative research rather than qualitative research (Onwuegbuzie and Frels 2016).

Finally, the overarching research question and the goal of investigating how researchers assess these outcomes required a review of relatively recent research instead of a historical review far back in time. Thus, the included studies were published in the last 15 years. Since the studies were collected in 2016, the review includes studies published between 2001 and 2015. Meta-analyses and literature reviews were excluded, along with studies that did not provide information regarding this review's inclusion criteria.

Screening the text of the collected studies to ensure they met all the inclusion criteria generated 139 studies published in 17 journals, as shown in Table 2.

Table 2. The journals and the number of studies contributing to the review

Journals	Total n = 139	%
Journal of Personal Selling & Sales Management	30	21.6
Industrial Marketing Management	29	20.9
Journal of Marketing	16	11.5
Journal of the Academy of Marketing Science	15	10.8
Journal of Business & Industrial Marketing	14	10.1
Journal of Business Research	9	6.5
Journal of Marketing Theory and Practice	5	3.6
International Journal of Research in Marketing	4	2.9
Journal of Marketing Research	4	2.9
European Journal of Marketing	3	2.2
Journal of Applied Psychology	3	2.2
Marketing Intelligence & Planning	2	1.4
Journal of Business-to-Business Marketing	1	0.7
Journal of International Marketing	1	0.7
Marketing Science	1	0.7
Psychology & Marketing	1	0.7
Women in Management Review	1	0.7

The table can be read like this: *Journal of Personal Selling and Sales Management* contributes 30 studies to this review, which constitutes 21.6% of the reviewed studies.

In the final step, the data of interest was extracted and organized into an evaluative framework. Such a framework is key in systematic reviews for data extraction and data categorization and for ensuring that the reviewed studies are handled in a consistent manner (Boot et al. 2016). In accordance with the research questions in each paper, the evaluative framework included the following sections to record the extracted data: measures, types of measures (objective and subjective measures), data sources, and types of respondents.

During the review process, the coding protocol was advanced and refined with new coding classes to record new categories of data as they arose. To ensure accuracy and transparency during the review process, a coding protocol was developed in Excel following the procedure recommended by Lipsey and Wilson (2001) regarding how to code extracted data. The protocol organized the studies by year of publication and by author(s). To ensure quality and reliability, the coding process was repeated a second time.

3.4. Data analyses

3.4.1. Data analysis in Paper 1

Paper 1 examines the measures used in the reviewed studies to assess B2B sales performance outcomes. The final dataset from the review included 139 studies using 151 different measures to assess the outcomes. A data matrix was used to record whether each measure was present or absent in each study, coded as 1 or 0, respectively. The final data matrix comprised 139 rows

(the studies) and 151 columns (the measures), showing which measures were used in each study.

For interpretation purposes, a cluster analysis was conducted on the measures to break down the large dataset of studies and measures into groups of studies using similar measures. This categorization was conducted as groups of similar studies are easier to examine than each individual study and measure alone (Hair et al. 2010). The cluster analysis was conducted on the presence/absence matrix of measures using a hierarchical clustering algorithm, which required two decisions: 1) how to quantify the differences between the studies in terms of the dichotomous observations (presence/absence of measures) and 2) which clustering criteria to use to join clusters of studies in the hierarchical process.

Concerning the first decision, the Jaccard Index of Similarity (Greenacre and Primicerio 2013) is suitable for quantifying the similarities between two studies as the dataset consists of mostly absent measures and only a few measures are present in each study. The problem with conducting a cluster analysis on such a dataset is that the most common similarity among the studies is a large number of absent measures, which need to be ignored in the present cluster analysis.

Concerning the second decision, the Ward (1963) criterion was used to combine clusters, which displayed the hierarchical clustering of the data structure in the form of a dendrogram (e.g., Tan, Steinbach, and Kumar 2006; Greenacre and Primicerio 2013). The dendrogram shows clusters of studies with similar (homogeneous) measures of B2B sales performance outcomes. Then, the level of homogeneity within each of the clusters and between the clusters was measured. The clusters of studies using homogeneous measures

provided a typology of the sets of measures researchers use to assess the outcomes (Hair et al. 2010).

3.4.2. Data analysis in Paper 2

Paper 2 aims to develop a theoretical framework of the desired outcomes from B2B selling. This theoretical framework was developed by categorizing the measured outcomes in the reviewed studies into an initial framework for marketing performance outcomes (Katsikeas et al. 2016). This initial framework was adjusted and adapted to B2B selling during the review process of recording and categorizing the measured outcomes. The adaptation was conducted by removing categories of outcomes that were not measured by the reviewed studies and adding new categories for measured outcomes that did not fit into the initial framework.

3.4.3. Data analysis in Paper 3

Paper 3 examines three key methodological issues in the reviewed studies as well as differences in the methods published in different journals. The first step of data analysis for this paper involved creating three cross-tabulations for the three key methodological issues and for the 17 journals. In the next step, three cluster analyses were performed to group journals that publish similar methods. The clustering algorithm uses the same distance measure between the row profiles—the Chi-square distance—as the distance measure used in the correspondence analyses (see below), which we performed to visualize the tables (Greenacre 2016). The specific algorithm used was Ward clustering (Ward 1963), which optimizes between-cluster distance variance in each step of the clustering process. Finally, three correspondence analyses were conducted to visualize the row profiles in correspondence

analysis maps (CA-maps) (Greenacre 2016). CA-maps aid in data interpretation by visualizing essential patterns (Hair et al. 2010). The present analysis used the version of correspondence analysis known as the contribution biplot (Greenacre 2013). This analysis generates arrows in CA-maps representing the various types of methods investigated. A journal's location in line with the arrow and location between the center cross and the arrowhead indicate the value of the journal profiles (Greenacre 2010) and, thus, the method used by the studies published in each journal. Journals located close to the center in the CA-maps tend to publish studies with methods that are close to the average for the reviewed studies. A location away from the center and toward a specific type of method shows that a journal publishes studies that use this particular type of method to a greater degree than average for the reviewed studies.

3.4.4. Data analysis in Paper 4

Paper 4 examines the data sources used to assess B2B sales performance outcomes in the reviewed studies. The measured outcomes were categorized within an initial framework with the four data sources (company records, sales managers, salespeople, and customers). The initial framework was expanded during the review process by adding two new categories of data sources to record measures assessed by combinations of respondents (sales managers and salespeople as well as sales managers, salespeople, and customers). Thus, adding these two categories of data sources, the final framework included six categories of data sources. Finally, the measured outcomes were reordered in accordance with the frequency that they were assessed in the reviewed studies.

3.5. Validity and reliability

Internal validity is one aspect of a quality assessment or a critical appraisal of a systematic review. Essentially, internal validity refers to the degree to which one can believe a study's results—in other words, how close the results are to the truth (Boot et al. 2016). Ensuring a high level of internal validity requires the researcher to avoid systematic bias and errors, for example, from reviewing only single studies (Jesson et al. 2012) or from selecting studies subjectively, which is often done in traditional literature reviews (Boot et al. 2016). To avoid such bias, a systematic review was conducted in this thesis.

In a systematic review, internal validity is facilitated by using a standard set of guidelines to execute the review (Boot et al. 2016). To ensure internal validity, the present review and thesis followed the guidelines from Palmatier et al. (2018) and Littell et al. (2008). These guidelines require the research to formulate clear research questions; provide clear definitions of the construct, population (studies), and inclusion criteria; identify all relevant studies; record the data into a predefined protocol; use data analyses to identify clusters, patterns, and relationships; present the data clearly and completely in tables and figures; interpret and discuss the core results to provide a deeper understanding; and discuss the implications for researchers, practitioners, and future research.

Early on in the systematic literature review's scientific development process, it was considered desirable to collect all relevant studies. However, more recently, there has been increasing recognition that even the most exhaustive search cannot collect the entire universe of studies (Boot et al. 2016). Thus, "fitness for purpose" is the appropriate aspiration for the collected

studies' coverage of the review scope, which is underpinned by the trade-off of rigour versus relevance (Bennet et al. 2005).

Through an extensive search for studies, including an issue-by-issue review of the journals and keyword searches on the journals' websites, the present review aimed to include all relevant studies. This search for studies resulted in 139 studies. This number of studies and the aim of collecting all relevant studies through an extensive search provide good coverage of the review scope and satisfactory internal validity (Boot et al. 2016).

However, two particular biases may have influenced this thesis' internal validity. First, by excluding unpublished studies, this review may suffer from the reasoning editors use when selecting which studies to publish, also called publication bias (Gilbody and Song 2000). Second, by excluding non-English studies, this review overlooks potentially valuable information from such studies (Song et al. 2010).

External validity (also termed generalizability or applicability) is another aspect essential to a quality assessment of studies conducting systematic reviews. External validity refers to the degree to which the results from a study can be applied to the population identified by the research question (Boot et al. 2016) or applied to another population (population validity), another setting (ecological validity), or over time (historical validity) (Dekkers et al. 2010).

The present review used a strict and detailed set of study inclusion criteria, such as the requirement that the studies research B2B selling and be empirical and quantitative. Using a strict set of inclusion criteria provided a clear definition of the population investigated (i.e., the particular type of sales research). This clear definition of the population and the large sample (of reviewed studies) provides satisfactory external validity (generalizability) to reviews with similar inclusion criteria (Boot et al. 2016).

The population validity and ecological validity of the investigated outcomes, measures, and number of measures have uncertain generalizability to, for example, reviews of B2C sales research. This uncertain generalizability is caused by the often wider variety of desired outcomes from B2B selling compared with B2C selling (e.g., Lilien 2016; Dawes et al. 1998; Manning et al. 2010). However, the present investigations of subjective and objective measures and the types of respondents and data sources may have satisfying population validity and ecological validity to reviews of B2C sales research because of similar methods to assess the outcomes. Regarding historical validity, changes in the sales profession and the advance of sales research have not occurred rapidly (Cuevas 2018; Asare et al. 2012). Thus, the present investigations have relatively robust historical validity over a short historical period, but this validity is more uncertain over a longer historical period.

Reliability is also a vital aspect of a quality assessment of a systematic literature review. The reliability of such review refers to the trustworthiness of the results and the degree to which both the review and the study results are reproducible (Boot et al. 2016). This study is highly reproducible because of the rigorous procedures and study inclusion criteria as well as the good coverage of the population with 139 reviewed studies (Boot et al. 2016; Cooper 2010).

Further, the data investigated in this review are categorical and such data require little evaluation from the researcher to be recorded correctly in the data protocol. Also, the coding process was repeated a second time to ensure the accuracy of the recorded data. Furthermore, the reliability of systematic reviews relates to whether the study findings are substantial enough to have a practical impact and thus to be meaningful (Boot et al. 2016). The present thesis' analyses reveal patterns of considerable size regarding how

researchers assess the outcomes and represent clear findings with reliable implications for sales researchers.

3.6. Ethical considerations

The present thesis uses secondary data sources and does not include information about firms or people from interviews, questionnaires, or observations. Thus, the present thesis has few ethical considerations regarding, for example, negative influences on respondents, the anonymity of respondents, or the confidentiality of firms (Bell and Bryman 2007).

However, one ethical consideration that needs to be addressed is the assurance that this research was conducted in accordance with the professional responsibilities of a researcher (Steneck 2006). These responsibilities include searching for the truth in an independent, honest, and forthright way (Bunge 1996). Independence and honesty are ensured by strictly following the principles of systematic reviews in this review. These principles require open-mindedness in observations and inductive bottom-up generalizations from an extensive collection of data (Ladyman 2002).

The Nord University Business School financed the work conducted for this thesis, so there are few potential conflicts of interest regarding the findings. By using public funding, this thesis is obligated to return the best possible results to both the public and the scientific community, which includes reporting findings correctly and honestly.

4. SUMMARY OF THE RESEARCH PAPERS

This chapter consists of four sections, with each section outlining each research paper's research question(s), methods, findings, and key contributions to the thesis. Each research paper is outlined in the context of the overarching research question, and all the papers are interconnected as they all investigate various methodological issues related to assessing B2B sales performance outcomes. The research papers are written to align with various scientific conferences' and journals' requirements and are therefore independent papers with different formats.

4.1. Paper 1. "Measures to assess B2B sales performance outcomes: A systematic review and future directions"

Paper 1 investigates an essential issue of assessing B2B sales performance outcomes: the measures used to assess such outcomes. Despite the importance of assessing outcomes from selling, the literature provides little consensus or guidance on which measures should be used to assess such outcomes (Siguaw et al. 2003; Ingram et al. 2005). Thus, Paper 1 addresses the following research question: *which measured do researchers use to assess outcomes from B2B selling?*

The study answers the research question through a systematic review of the measures used to assess outcomes from B2B selling in 139 published studies. The review shows that researchers assess these outcomes very differently by using 151 different measures. This large variety of measures

demonstrates the lack of consensus among researchers regarding measures to assess the outcomes.

A cluster analysis of the measures was conducted to simplify the large dataset of studies and measures, which disclosed seven sets of measures used to assess the outcomes. These sets of measures were evaluated in relation to previous research explaining outcomes from B2B selling. The evaluation shows that a large portion of the reviewed studies basically only measure sales revenue and overlook outcomes associated with salespeople's customer interactions. Consequently, these studies disregard the multiple types of outcomes desired from B2B selling and the critical importance of fruitful customer interactions in B2B selling addressed in research and by practitioners.

This paper provides three contributions. First, by demonstrating the lack of consensus on measures and the narrow assessment of outcomes, this paper addresses the need for many sales researchers to reevaluate their measures of these outcomes. Second, this study categorizes the measures into those that assess outcomes beneficial to customers and those that assess outcomes beneficial to selling companies. This categorization provides detailed insights into frequently assessed company outcomes and less frequently assessed customer outcomes. These insights offer help for researchers to conduct more balanced assessments of company and customer outcomes in future research. Third, this paper suggests recommendations for improved measures of the outcomes.

4.2. Paper 2. "Desired outcomes from B2B selling: A systematic review and conceptualization"

Paper 2 contributes to further narrowing the research gap regarding outcomes from B2B selling and valid measures to assess such outcomes by developing a theoretical framework conceptualizing the outcomes desired from B2B selling. The sales literature does not offer such a conceptualization despite the importance of these outcomes to research and management. This study addresses the following research question: *which outcomes are desired from B2B selling?*

The research question was answered by organizing the reviewed studies' measured outcomes into a value chain framework. This organizing led to the creation of the B2B Sales Performance Outcomes Chain. This chain is the first complete theoretical framework conceptualizing the outcomes desired from B2B selling and suggest seven main types and 21 subtypes of outcomes. Further, the chain shows how the various outcomes relate to and influence each other. Researchers can use the chain to select measures to assess the outcomes or investigate hypotheses and research models that include the outcomes. Managers can use the chain as an overarching executive tool for targeting and monitoring outcomes and for directing sufficient effort toward realizing the various desired outcomes.

4.3. Paper 3. "Methods to assess outcomes from B2B selling: A systematic review, cross-journal examination, and guidelines"

Paper 3 investigates three key methodological issues vital for assessing B2B sales performance outcomes: how many measures to use, what type(s) of

measures to use, and what type(s) of respondents to use. This study investigates these methodological issues because the sales literature offers little guidance and appears inconsistent on these issues. This inconsistency implies that sales researchers frequently use methods that previous research associates with biased assessments. This study addresses the following two research questions to resolve these gaps: *how appropriate are the methods researchers use to assess B2B sales performance outcomes, and are there differences in methods published in different journals?*

This study answers these research questions through a systematic review of the three key methodological issues in the reviewed studies. Further, this study examines differences in these methods across the 17 journals publishing the reviewed studies.

This study provides several contributions to improve researchers' methods for assessing the outcomes. First, it exhibits a large variety of possible research methods and pinpoints their various sophistication and quality. Second, this study reveals the widespread use of methods that may not provide the most reliable and valid assessments. This finding indicates the need for many sales researchers to scrutinize and reevaluate their methods. Third, the examination shows how studies in the highest-ranked journals provide more reliable and valid assessments by combining objective and subjective measures and multiple types of respondents to a larger degree than average for the reviewed studies. Last, this study contributes by suggesting guidelines for improved methods to assess the outcomes.

4.4. Paper 4. "Data sources to assess sales performance outcomes"

Paper 4 investigates the data sources used to assess the various types of outcomes from B2B selling. The outcomes can be assessed using four data sources: company records, sales managers, salespeople, and customers. This investigation is motivated by the data sources' various ability to provide reliable and valid assessments of different types of outcomes. Further, the investigation is motivated by the little guidance on which data sources should be used to assess various types of outcomes. The study addresses the following two research questions: *which data sources do researchers use to assess the various outcomes from B2B selling, and do researchers use the most valid data sources to assess the various outcomes?*

This study answers the research questions by examining and evaluating the data sources used to assess the outcomes in the reviewed studies. The study provides two main contributions. First, it reveals the widespread use of salespeople and sales managers to rate economic outcomes and outcomes associated with salespeople's customer interactions. However, these are outcomes that company records and customers, respectively, can provide considerably more reliable assessments of. By identifying such widespread mismatches between the data sources used in research and the outcomes they assess, this study should encourage sales researchers to reevaluate the data sources they use to assess various outcomes. Second, this study suggests the first concrete guidelines on the most valid data sources to assess various outcomes and outcomes that preferably could be assessed using each data source.

5. CONCLUSIONS AND IMPLICATIONS

5.1. Conclusions

This thesis is based on a systematic literature review of the methods used to assess outcomes from B2B selling in 139 published studies. The review reveals variation in the quality and sophistication of methods researchers use to assess these outcomes, demonstrating the lack of consensus regarding these methods. Further, this thesis evidences a sizeable inconsistency in sales research—namely, the widespread use of methods to assess the outcomes that have been criticized in research for being associated with biases. This large variety of methods and this inconsistency in sales research may stem from the little guidance on methods to assess these outcomes.

This thesis also reveals that studies published in the highest-ranked journals use best-practice methods to a larger degree than average among the reviewed studies. This finding may be expected, but it is valuable to identify which concrete methods these studies conduct better. Thus, these studies can serve as best-practice examples of research methods that ensure more robust reliability and validity of assessed outcomes.

The lack of consensus and the variation in the quality of the methods used to assess the outcomes from B2B selling call for many sales researchers to reevaluate and improve their research methods. This thesis contributes to such a reevaluation by suggesting methods that may need improvement, theoretical frameworks and guidelines on improved methods, and particular issues for further investigations. The following paragraphs complete the discussion on each of the investigated methodological issues in this thesis.

5.1.1. Conceptualization of the outcomes

Proper conceptualization of the outcomes from B2B selling are important because these outcomes constitute an abstract and complex construct containing multiple types of outcomes (Groves et al. 2009). Further, such conceptualization provides a foundation to enable an operationalization of measures with solid construct validity (Groves et al. 2009; Ghauri and Grønhaug 2002).

This thesis extends previous research conceptualizing outcomes from marketing (Katsikeas et al. 2016) by adapting this conceptualization to B2B selling. This adaption provides the first complete theoretical framework that conceptualizes desired outcomes from B2B selling, as represented by the B2B Sales Performance Outcomes Chain. This framework suggests seven main types and 21 subtypes of outcomes from B2B selling. This framework can help many researchers develop measures that reflect the multiple types of outcomes and improve the construct validity of future assessments.

This framework also extends and systemizes the multiple outcomes from B2B selling suggested in earlier research on B2B selling and marketing. The framework particularly systemizes outcomes directly beneficial to customers, such as customer satisfaction (Cravens 1995), offer value (Blocker et al. 2012), service quality (Töytäri and Rajala 2015), customer relationships (Storbacka et al. 2009), customer loyalty (Lam et al. 2004), and co-creation of products and services (Töytäri and Rajala 2015).

5.1.2. Operationalization of measures

The large variety of measures and the large variety of the number of measures used to assess the outcomes show that researchers have different understandings and conceptualizations of this same construct. These

differences raise concerns regarding how sales performance outcomes are labeled in research. For example, when measuring only sales revenue to assess the outcomes, it would be more literally and scientifically accurate to name the dependent variable "sales revenue" and not "sales performance outcomes." Further, this wide variety of measures and number of measures raise questions regarding whether the various measures perform equally well and which differences in construct validity they will eventually involve.

A large portion of the reviewed studies use relatively few sales revenue measures to assess the outcomes. The need for such measures is evident by the fact that sales revenue is the most important outcome from selling. Further, such measures are highly reliable when using sales revenue data from company records (Dess and Robinson 1984).

However, the reviewed studies' widespread use of relatively few sales revenue measures stands in contrast to the salesperson's expanded role (Cuevas 2018) and focus on customer relationships and customer satisfaction (Hughes and Ogilvie 2020). Further, using few sales revenue measures stands in contrast to the multiple measures suggested by practitioners in B2B selling (Zallocco et al. 2009; Behrman and Perreault 1982; Cron et al. 2014). Also, this contrast may indicate a valuable strength of qualitative research designs as two of these studies (Zallocco et al. 2009; Cron et al. 2014) are qualitative studies. These two studies used in-depth interviews via an open-ended interview approach that enabled respondents to freely articulate their experiences and knowledge (McGivern 2013). This interview approach may have stimulated the practitioners to express a broad range of desired outcomes relevant to assess. Behrman and Perreault's (1982) study is mainly quantitative, but they used qualitative research techniques to prepare their research instrument with a panel of judges comprising, among others, sales and marketing managers.

Also, solely assessing sales revenue contrasts the multiple types of outcomes suggested by the present conceptualizations in the B2B Sales Performance Outcomes Chain. Further, such measurement overlooks how critical outcomes influence each other (Katsikeas et al. 2016), as suggested in this chain. For example, creating customer satisfaction is important to achieve customer retention (e.g., Rauyruen and Miller 2007; Blocker et al. 2011), which is essential for creating unit sales (Rauyruen and Miller 2007) and sales revenue (Lam et al. 2004; Rauyruen and Miller 2007).

Another central tendency among the reviewed studies is the much more frequent measurement of company outcomes and the less frequent measurement of customer outcomes. This tendency reveals the need for sales researchers to increase their measurement of customer outcomes in future sales research.

Choosing which measures to assess such an abstract and multi-dimensional construct as the outcomes from B2B selling involves the fundamental conflict of reliability and validity in survey research (Franke, Rapp, and Andzulis 2013). In this context of assessing outcomes from selling, this conflict refers to balancing reliability and construct validity. For example, using multiple measures to assess one type of outcome may strengthen that specific outcome's reliability and reveal stronger relationships between antecedents and the outcome (e.g., Churchill 1979). However, using multiple measures to assess one type of outcome may come at the expense of using the measures to assess multiple types of outcomes (Richins 2004), which is crucial for construct validity (Groves et al. 2009; Ghauri and Grønhaug 2002). The balancing of reliability and validity is essential to avoid using too many measures. Overly long questionaries could result in respondents having to dedicate a significant

amount of effort and time, thus potentially lowering response rates and measurement reliability (e.g., Richins 2004; Drolet and Morrison 2001).

5.1.3. The number of measures

The review reveals substantial variability in the number of measures used to assess the outcomes from B2B selling and the widespread practice of using few measures. Using too few measures is inadequate to capture and reflect abstract constructs with multiple components (Groves et al. 2009: Ghauri and Grønhaug 2002), such as the multiple types of outcomes from B2B selling (e.g., Zallocco et al. 2009; Behrman and Perreault 1982; Cron et al. 2014). Thus, construct validity—namely, how well the measures conform to previous research—is not satisfying for those studies using few measures to assess such constructs that previous research has suggested consist of multiple components/types of outcomes (Seale 2009).

5.1.4. Objective and subjective measures

This thesis extends the research on subjective and objective measures in sales research. Subjective measures are the most commonly used measures in the reviewed studies even though economic outcomes are the most frequently measured outcomes. The use of subjective measures to assess economic outcomes stands in contrast to the fact that subjective measures can be poor indicators of objective sales performance outcomes (Jaramillo et al. 2005; Rich et al. 1999). Thus, researchers strongly recommend using objective measures to assess economic outcomes when such measures are available from company records (Dess and Robinson 1984).

To a large extent, the reviewed studies published in the highest-ranked journals follow the best practice of combining subjective and objective

measures, as recommended in research (Rapp et al. 2020; Bagozzi et al. 2003). These studies combine subjective and objective measures to a larger degree than average among the reviewed studies. Such combination enables researchers to harness the complementary benefits of subjective and objective measures (Bagozzi et al. 2003), for example, by using reliable, objective measures of economic outcomes and valid subjective measures of sales peoples' customer interactions.

One possible explanation of the dominant use of subjective measures in the reviewed studies may stem from researchers' limited access to objective measures from company records. Further, this frequent use of subjective measures rated by respondents may arise from the convenience of using one data source that can assess both the antecedents and the outcomes of B2B selling and thus provide all the data necessary for a research study.

5.1.5. Single or multiple respondent types

This thesis extends previous research regarding the use of single or multiple respondent types in research surveys. Previous research recommends using multiple respondent types to reduce respondent bias (Hulland et al. 2018; Jap and Anderson 2004) and provide richer assessments (Hughes et al. 2013; Hulland et al. 2018). In contrast, half of the reviewed studies use one sample of salespeople as respondents.

However, the present review shows that the highest-ranked journals tend to publish studies that follow the best practice of using multiple respondents, as suggested in previous research (Hulland et al. 2018; Schmitz et al. 2014; Hughes et al. 2013; Jap and Anderson 2004). The studies published in these highest-ranked journals use two (dyadic) respondent types to a much larger degree than average among the reviewed studies. Further, these

highest-ranked journals are the only journals that publish studies using three (triadic) respondent types.

One plausible explanation for the frequent use of one type of respondents in the reviewed studies may be the additional effort needed for data collection and analysis when using multiple types of respondents. However, the studies published in the highest-ranked journals appear to have overcome this convenience barrier, apparently recognizing the robust reliability and validity of using multiple respondent types.

5.1.6. Respondent types

The present thesis extends previous research on appropriate respondents to use in sales research surveys. Most of the current reviewed studies use salespeople and sales managers as respondents, either separately or in combination. However, previous research reveals several concerns regarding such respondents. First, salespeople and sales managers are not preferrable for assessments of economic outcomes (Jaramillo et al. 2005; Rich et al. 1999). Further, salespeople's self-ratings are generally associated with respondent bias from socially desirable responding (Steenkamp et al. 2010). Last, sales managers are not preferrable for rating several outcomes related to salespeople's customer interactions (Cannon and Spiro 1991). Thus, it is interesting that only a small portion of the reviewed studies use customers as respondents even though customers can provide the most reliable and valid assessments of outcomes related to salespeople's customer interactions (Cannon and Spiro 1991; Lambert et al. 1997), which are critical outcomes for successful B2B selling (Wang et al. 2012; Wang et al. 2018; Töytäri and Rajala 2015; Blocker et al. 2012; Lam et al. 2004; Storbacka et al. 2009; Mullins et al. 2014). The importance of these outcomes and the little use of customers as

respondents indicate the need to increase the use of customers as respondents in research surveys.

5.1.7. Data sources to assess various types of outcomes

This thesis extends previous research on various data sources' ability to assess different types of outcomes. The present analysis connected the data sources used to assess various outcomes in the reviewed studies. This analysis provided a categorization of the measured outcomes by each of the data sources.

This categorization reveals the widespread use of salespeople and sales managers to rate sales revenue even though using company records to assess such outcomes provides the most reliable and valid assessments (Jaramillo et al. 2005; Rich et al. 1999). Further, salespeople and sales managers are frequently used to rate outcomes related to salespeople's customer interactions despite customers' ability to provide the most reliable and valid assessments of such outcomes (Cannon and Spiro 1991; Lambert et al. 1997).

The following reasons may explain this widespread use of salespeople and sales managers to assess economic outcomes and outcomes related to salespeople's customer interactions. Most importantly, researchers usually have limited access to company records and customer files. Thus, salespeople and sales managers may be chosen due to convenience and tradition. Second, such outcomes are critical in B2B selling, so researchers may tend to measure them despite lacking access to the most appropriate data sources. Finally, this non-optimal use of data sources to measure outcomes may be influenced by the lack of research, debate, and frameworks on appropriate data sources to assess various types of outcomes from selling. This thesis contributes such a review and evaluation and suggests such a framework.

The examination also reveals that relatively few outcomes are measured using company records and customers. This finding indicates the need to increase the variety of outcomes measured using company records and customers to utilize these data sources' ability to provide reliable and valid assessments of essential outcomes from B2B selling.

5.2. Implications for researchers

Sales researchers should recognize that sales research and practitioners believe B2B selling should produce multiple types of outcomes and that a large portion of the present reviewed studies use measures that ignore this multiplicity. This difference indicates that many sales researchers may need to reevaluate their measures to avoid a theory gap related to such an essential construct in sales research. In such a reevaluation, the selected measures should be based on robust theory (Richard et al. 2009) and construct validity to ensure the measures reflect the multiple types of outcomes (Groves et al. 2009; Ghauri and Grønhaug 2002).

The B2B Sales Performance Outcomes Chain developed in this thesis is suitable to conceptualize and operationalize the outcomes and can contribute to this reevaluation and improved construct validity. Further, researchers can use this framework to develop research models and hypotheses as the framework suggests how the outcomes are related to and influence each other.

The present literature review provides researchers an extensive outline of the large variety of survey methods available to assess the outcomes from B2B selling. Researchers can use this outline to select methods to assess outcomes from selling and to develop sales and marketing survey designs in

general. Further, the present evaluation of methods should alert sales researchers to recognize the varying sophistication of sales research methods. In particular, researchers need to recognize methods associated with biased assessments and exercise caution when using such methods. Furthermore, the guidelines and frameworks presented in this thesis provide researchers valuable insights into methods to generate the most reliable and valid assessments of the outcomes from B2B selling.

Moreover, researchers can improve their assessments of these outcomes by adopting the method-related best practices used in the reviewed studies published in the highest-ranked journals. In particular, these studies attenuate respondent bias and strengthen the reliability and validity of the assessed outcomes by combining objective and subjective measures (Bagozzi et al. 2003) and using multiple types of respondents (Hulland et al. 2018; Jap and Anderson 2004; Schmitz et al. 2014; Hughes et al. 2013).

Researchers may need to more fully recognize the problems associated with the widespread use of salespeople and sales managers to rate economic outcomes and outcomes related to salespeople's customer interactions. Such outcomes are critical in B2B selling; thus, sales researchers should consider increasing their use of company records and customers to achieve more reliable and valid assessments of such outcomes. Also, researchers should preferably increase the number and variety of outcomes assessed using company records and customers. Finally, researchers can use the suggested guidelines on the most reliable and valid data sources to assess various outcomes.

5.3. Implications for managers

This thesis is based on an examination of how researchers assess outcomes from B2B selling but does not examine how managers assess these outcomes. However, effective sales management depends on reliable and valid assessments of these outcomes as B2B selling has become more challenging (Paesbrugghe et al. 2020) and fewer salespeople are failing to meet their sales goals (Hyken 2018). Reliable and valid assessments of the outcomes are vital for managers to detect failing outcomes and dysfunctional sales efforts and to take necessary action to improve selling (Zoltners et al. 2008).

When assessing outcomes from selling, managers collect the same measures from the same data sources as researchers using questionnaires, interviews, and company records. Thus, the evaluations and recommendations in this thesis provide valuable and actionable guidance to managers as well.

Managers should acknowledge firms' problems in selecting appropriate measures of sales success (Haines 2004) and avoid assessing solely economic outcomes due to the growing importance of, for example, customer relationships and customer satisfaction (Hughes and Ogilvie 2020; Zallocco et al. 2009). Managers should also recognize the measures suggested by practitioners in B2B sales research (Zallocco et al. 2009; Behrman and Perreault 1982; Cron et al. 2014).

To ensure to select measures with broad coverage of the outcomes, managers can use the B2B Sales Performance Outcomes Chain with its seven main types and 21 subtypes of outcomes. Further, this framework can be used as a management tool, outlining the multiple responsibilities for sales managers and salespeople. The framework can help firms develop strategies for their sales operations, select areas for improvement, and allocate sufficient

resources to realize targeted outcomes. Further, managers can use this framework to educate, train, monitor, and manage salespeople regarding the outcomes they are required to produce.

Managers have easy access to all data sources relevant to assess the outcomes: company records, sales managers, salespeople, and customers. This access is a valuable benefit as using the most appropriate data sources is vital for reliable and valid assessments of the outcomes. Managers are advised to combine objective measures from company records with subjective measures (Bagozzi et al. 2003) from several types of respondents (Hulland et al. 2018; Jap and Anderson 2004; Schmitz et al. 2014; Hughes et al. 2013).

Company records can be used to gain the most reliable and valid assessments of economic outcomes, and customers can be used to gain the most reliable and valid assessment of outcomes associated with salespeople's customer interactions. Further, sales managers and salespeople can be used to rate outcomes associated with sales operations. However, such ratings should be interpreted with caution because of the probable respondent bias, particularly when using salespeople's self-ratings. The present suggested data sources appropriate to assess the various types of outcomes provides managers with actionable guidelines.

5.4. Limitations and future research

This thesis investigates the methods used to assess outcomes from B2B selling in published quantitative research. Thus, this research frame does not include unpublished studies, qualitative studies, or studies on B2C selling, all of which may have contributed to validating the present findings (Onwuegbuzie and Frels 2016). Further, regardless of the attempt to collect all published

studies satisfying the present literature review's inclusion criteria, some eligible studies may not have been detected and included in the review.

Further, despite recognizing the relevance of the 21 types of outcomes suggested in the B2B Sales Performance Outcomes Chain, this thesis does not validate these outcomes. Such validation is an essential task for future research and could involve qualitative and quantitative research techniques to allow practitioners in B2B selling to rate the importance of the outcomes and/or the measures used to assess the outcomes. Such validation could aim to develop a standard set of measures adapted to, for example, statistical analyses frequently used in sales research, such as structural equation modeling (Asare et al. 2012). Such a validated standard set of measures would guide researchers on appropriate measures and improve future assessments and sales research. Further, such an investigation could help researchers create research models that include the outcomes managers deem important. Such adaptation could narrow the research-practice gap that sales researchers have pinpointed throughout history (Pullins et al. 2017). Finally, if researchers could agree on applying such a standard set of measures, this would help synthesize findings across studies and improve cumulative knowledge building in sales research (Katsikeas et al. 2016; Verbeke et al. 2011).

This thesis reveals the widespread use of salespeople's self-ratings despite the potential respondent bias associated with such ratings. Thus, future research could investigate the relationships between ratings from salespeople, sales managers, and customers, for example, ratings of outcomes associated with salespeople's customer interactions. Such an investigation could provide guidance on combining such respondent to reduce respondent bias.

REFERENCES

Anderson, E., & Oliver, R. L. (1987). Perspectives on Behavior-Based Versus Outcome-Based Salesforce Control Systems. *Journal of Marketing, 51*(4), 76-88.

Albers, S., Mantrala, M. K., & Sridhar, S. (2010). Personal Selling Elasticities: A Meta-Analysis. *Journal of Marketing Research, 47*(5), 840-853.

Asare, A. K., Yang, J., & Brashear Alejandro, T. G. (2012). The State of Research Methods in Personal Selling and Sales Management Literature. *Journal of Personal Selling & Sales Management,* 32 (4), 473-489.

Bagozzi, R. P., Verbeke, W., & Gavino J. C. Jr. (2003). Culture Moderates the Self-Regulation of Shame and its Effects on Performance: The Case of Salespersons in the Netherlands and the Philippines. *Journal of Applied Psychology, 88*(2), 219-322.

Behrman, D. N., & Perreault, W. D. Jr. (1982). Measuring the Performance of Industrial Salespeople. *Journal of Business Research, 10*(September), 355-369.

Bell, E., & Bryman, A. (2007). The Ethics of Management Research: An Exploratory Content Analysis. *British Journal of Management, 18*(1), 63-77.

Bennet, J., Lubben, F., Hogarth, S., & Campbell, B. (2005). Systematic Reviews of Research in Science Education: Rigour or Rigidity? *International Journal of Science Education, 27*(4), 387-406.

Blocker, C. P., Cannon, J. P., Panagopoulos, N. G., & Sager, J. K. (2012). The Role of the Sales Force in Value Creation and Appropriation: New Directions for Research. *Journal of Personal Selling & Sales Management, 32*(1), 15-27

Blocker, C. P., Flint, D. J., Myers, M. B., & Slater, S. F. (2011). Proactive Customer Orientation and its Role for Creating Customer Value in Global Markets. *Journal of The Academy of Marketing Science, 39*(2), 216-233.

Boot, A., Sutton, A., & Papaioannou, D. (2016). *Systematic Approaches to a Successful Literature Review*. London: SAGE Publications Ltd.

Bunge, M. A. (1996). *Finding Philosophy in Social Science*. New Heaven, Conn: Yale University Press.

Cannon, J. P., & Spiro, R. (1991). The Measurement of Salesperson Performance: Comparing Self-Evaluations with Customer Evaluations. In Gilly, M. et al. (Eds.). *Enhancing Knowledge Development in Marketing. Summer Educators' Proceedings* (pp 1-10). Chicago, IL: American Marketing Association.

Chan, A.-W., & Altman, D. G. (2005). Identifying Outcome Reporting Bias in Randomised Trials on PubMed: Review of Publications and Survey of Authors, *BMJ, 330*(7494, 753), 1-6.

Churchill, G. A., Jr. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research, 16*(1), 64-73.

Churchill, G. A., Ford, N. M., Hartley, S. W., & Walker, O. C. Jr. (1985). The Drivers of Salesperson Performance: A Meta-Analysis. *Journal of Marketing Research, 22*(2), 103-118.

Cooper, H. (2010). *Research Synthesis and Meta-Analysis: A Step-by-Step Approach.* Thousand Oaks, California: SAGE Publications Inc.

Cooper, H., & Hedges, L. V. (2009). Research Synthesis as a Scientific Process. In Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). *The Handbook of Research Synthesis and Meta-Analysis* (pp 3-16). New York: Russell Sage Foundation.

Cravens, D. W. (1995). The Changing Role of the Sales Force. *Marketing Management, 4*(2), 49-54.

Cron, W. L., Baldauf, A., & Leigh, T. W. (2014). The Strategic Role of the Sales Force: Perceptions of Senior Sales Executives. *Journal of the Academy of Marketing Science, 42*, 471-489.

Crook, R. T., Shook, C. L., Morris, M. L., & Madden, T. M. (2010). Are We There Yet? An Assessment of Research Design and Construct Measurement Practices in Entrepreneurship Research. *Organizational Research Methods, 13*(1), 1-15.

Cuevas, J. M. (2018). The Transformation of Professional Selling: Implications for Leading the Modern Sales Organization. *Industrial Marketing Management, 69*(February), 198-208.

Davie, C., Stephenson, T., & Valdivieso De Uster, M. (2010). Three Trends in Business-to-Business Sales. *McKinsey Quarterly 19*(May), 1-4.

Dawes, P. L., Lee, D. Y., & Dowling, G. R. (1998). Information Control and Influence in Emergent Buying Centers. *Journal of Marketing, 62*(3), 55–68.

Dekkers, O. M., von Elm, E., Algra, A., Romijn, J. A., & Vandenbroucke, J. P. (2009). How to Assess the External Validity of Therapeutic Trials: A Conceptual Approach. *International Journal of Epidemiology, 39*(1), 89-94.

Dess, G. G., & Robinson, R. B. Jr. (1984). Measuring Organizational Performance in the Absence of Objective Measures: The Case of the Privately-Held Firm and Conglomerate Business Unit. *Strategic Management Journal, 5*(3), 265-273.

Drolet, A. L., & Morrison, D. G. (2001). Do We Really Need Multiple-Item Measures in Service Research? *Journal of Service Research, 3*(3), 196-204.

Easterby-Smith, M., Thorpe, R., & Jackson, P. R. (2012). *Management Research*. London: SAGE Publications Ltd.

Fishbein, M., & Ajzen, I. (2010). *Predicting and Changing Behavior. The Reasoned Action Approach*. New York: Psychology Press.

Flint, D. J., Woodruff, R. B., & Gardial, S. F. (2002). Exploring the Phenomenon of Customers' Desired Value Change in a Business-to-Business Context. *Journal of Marketing, 66*(4), 102-117.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research, 18*(1), 39-50.

Franke, G. R., Rapp, A., & Andzulis, J. M. (2013). Using Shortened Scales in Sales Research: Risks, Benefits, and Strategies. *Journal of Personal Selling & Sales Management, 33*(3), 319-328.

Gartner. (2019). *CSO Update: The New B2B Buying Journey and Its Implication for Sales.* Gartner Report. https://emtemp.gcom.cloud/ngw/globalassets /en/sales-service/documents/trends/cso-update.pdf.

Ghauri, P., & Grønhaug, K. (2002). *Research Methods in Business Studies.* Harlow, UK: Pearson Education Limited.

Gilbody, S. M., & Song, F. (2000). Publication Bias and the Integrity of Psychiatry Research. *Psychological Medicine, 30*(2), 253-258.

Gilje, N., & Grimen, H. (1993). Samfunnsvitenskapens Forutsetninger. Innføring i Samfunnsvitenskapenes Vitenskapsfilosofi. Oslo: Universitetsforlaget.

Greenacre, M. (2010). *Biplots in Practice.* Bilbao: BBVA Foundation. Free download from www.multivariatestatistics.org.

Greenacre, M. (2013). Contribution Biplots. *Journal of Computational and Graphical Statistics, 22*(1), 107-122.

Greenacre, M. (2016). *Correspondence Analysis in Practice.* Florida: Chapman & Hall / CRC Press.

Greenacre, M., & Primicerio, R. (2013). *Multivariate Analysis of Ecological Data*. Bilbao: Fundación BBVA. Free official download from www.multivariatestatistics.org.

Groves, R. M., Fowler, F. J. Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Guba, E. G. & Lincoln, Y. S. (1994). Competing Paradigms in Qualitative Research. In Denzin, N. K., & Lincoln, Y. S. (Eds.). *Handbook of Qualitative Research* (pp 105-117). Thousand Oaks, CA: SAGE Publications Inc.

Guenzi, P., & Panzeri, F. (2015). How Salespeople see Organizational Citizenship Behaviors: An Exploratory Study Using the Laddering Technique. *Journal of Business and Industrial Marketing, 30*(2), 218-232.

Gummesson, E., & Polese, F. (2009). B2B is not an Island! *Journal of Business & Industrial Marketing, 24(*5/6), 337-350.

Haines, M. (2004). Sales Metrics and the Sales Dashboard: Gear up for Improved Accountability. *VARBusiness, 20*(12), pp. 60.

Hair, J. F. Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis*. Upper Saddle River, New Jersey: Prentice-Hall.

Henard, D. H., & Szymanski, D. M. (2001). Why Some New Products are More Successful Than Others. *Journal of Marketing Research, 38*(3), 362-375.

Herjanto, H., & Franklin, D. (2019). Investigating Salesperson Performance Factors: A Systematic Review of the Literature on the Characteristics of Effective Salespersons. *Australasian Marketing Journal, 27*(2), 104-112.

Hughes, D. E., Le Bon, J., & Rapp, A. (2013). Gaining and Leveraging Customer-Based Competitive Intelligence: The Pivotal Role of Social Capital and Salesperson Adaptive Selling Skills. *Journal of The Academy of Marketing Science, 41*(1), 91-110.

Hughes, D. E., & Ogilvie, J. L. (2020). When Sales Becomes Service: The Evolution of the Professional Selling Role and an Organic Model of Frontline Ambidexterity. *Journal of Service Research, 23*(1), 22-32.

Hulland, J., Baumgartner, H., & Smith, K. M. (2018). Marketing Survey Research Best Practices: Evidence and Recommendations from a Review of JAMS Articles. *Journal of the Academy of Marketing Science, 46*(1), 92-108.

Hult, G. T. M., Ketchen, D. J. Jr., Griffith, D. A., Chabowski, B. R., Hamman, M. K., Dykes, B. J., Pollitte, W. A., & Cavusgil, S.T. (2008). An Assessment of the Measurement of Performance in International Business Research. *Journal of International Business Studies, 39*(6), 1064-1080.

Hyken, S. (2018). *57% of Salespeople Missed Their Quotas Last Year*. Forbes. https://www.forbes.com/sites/shephyken/2018/09/02/77-of- sales-reps-missed-their-quotas-last-year/#6e7633a052e4.

Ingram, T. N., LaForge, R. W., Locander, W. B., MacKenzie, S. B., & Podsakoff, P. M. (2005). New Directions in Sales Leadership Research. *Journal of Personal Selling & Sales Management, 25*(2), 137-154.

Jap, S. D., & Anderson, E. (2004). Challenges and Advances in Marketing Strategy Field Research. In Moorman, C., & Lehmann, D. R. (Eds.). *Assessing Marketing Strategy Performance* (pp 269-292). Cambridge, MA: Marketing Science Institute.

Jaramillo, F., Carrillat, F. A., & Locander, W. B. (2005). A Meta-Analytic Comparison of Managerial Ratings and Self-Ratings. *Journal of Personal Selling & Sales Management, 25*(4), 315-328.

Jesson, J. K., Matheson, L., & Lacey, F. M. (2012). *Doing Your Literature Review: Traditional and Systematic Techniques.* London: SAGE Publications Ltd.

Jones, E., Brown, S. P., Zoltners, A. A., & Weitz, B. A. (2005). The Changing Environment of Selling and Sales Management. *Journal of Personal Selling & Sales Management, 25*(2), 105-111.

Katsikeas, C. S., Morgan, N. A., Leonidou, L. C., & Hult, G. T. M. (2016). Assessing Performance Outcomes in Marketing. *Journal of Marketing, 80*(2), 1-20.

Kuhn, T. S. (2012). *The Structure of Scientific Revolutions.* Chicago: The University of Chicago.

Ladyman, J. (2002). *Understanding Philosophy of Science.* Abingdon, Oxon: Routledge.

Lam, S. Y., Shankar, V., Erramilli, M. K., & Murthy, B. (2004). Customer Value, Satisfaction, Loyalty, and Switching Costs: An Illustration from a Business-to-Business Service Context. *Journal of the Academy of Marketing Science, 32*(3), 293-311.

Lambert, D. M., Sharma, A., & Levy, M. (1997). What Information can Relationship Marketers Obtain from Customer Evaluations of Salespeople? *Industrial Marketing Management, 26*(2), 177-187.

Lilien, G. (2016). The B2B Knowledge Gap. *International Journal of Research in Marketing*, *33*(3), 543-556.

Limbu, Y. B., Jayachandran, C., Babin, B. J., & Peterson, R. T. (2016). Empathy, Nonverbal Immediacy, and Salesperson Performance: The Mediating role of Adaptive Selling Behavior. *Journal of Business & Industrial Marketing, 31*(5), 654–667.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic Controversies, Contradictions, and Emerging Confluences. In Denzin, N. K., & Lincoln, Y. S. (Eds.). *The Handbook of Qualitative Research* (pp. 163-188). Thousand Oaks, CA: Sage Publications Inc.

Littell, J. H., Corcoran, J., & Pillai V. (2008). *Systematic Reviews and Meta-Analysis*. New York: Oxford University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis.* Thousand Oaks, CA: SAGE Publications Inc.

MacInnis, D. J. (2011). A Framework for Conceptual Contributions in Marketing. *Journal of Marketing, 75*(4), 136-154.

Manning, G., Reece, B., & Ahearne, M. (2010). *Selling Today: Creating Customer Value.* Upper Saddle River, New Jersey: Prentice Hall.

Martinez-Martin, P. (2010). Composite Rating Scales. *Journal of the Neurological Sciences, 289*(1-2), 7-11.

McGivern, Y. (2013). *The Practice of Market Research: An Introducti*on. Harlow, UK: Pearson Education Limited.

Miller Heiman Group (2018). *Sales Enablement Grows up: 4th Annual Sales Enablement Report.* CSO Insights: The Research Division of Miller Heiman Group. Littleton, CO.

Moncrief, W. C., Marshall, G. W., & Watkins, C. (2000). Tracking Academic Research in Selling and Sales Management: Authors, Authorships, Academic Institutions, and Journals. *Journal of Personal Selling & Sales Management, 20*(2), 99-108.

Mullins, R. R., Ahearne, M., Lam, S. K., Hall, Z. R., & Boichuk, J. P. (2014). Know Your Customer: How Salesperson Perceptions of Customer Relationship Quality Form and Influence Account Profitability. *Journal of Marketing, 78*(6), 38-58.

Ngai, E. W. T. (2005). Customer Relationship Management Research (1992-2002): An Academic Literature Review and Classification. *Marketing Intelligence & Planning. 23*(6), 582-605.

Nord, J.H. & Nord, G.D. (1995). MIS Research: Journal Status and Analysis. *Information & Management, 29*(1), 29-42.

Nygaard, A., & Dahlstrom, R. (2002). Role Stress and Effectiveness in Horizontal Alliances. *Journal of Marketing, 66*(2), 61-83.

Ohiomah, A., Benyoucef, M., & Andreev, P. (2020). A Multidimensional Perspective of Business-to-Business Sales Success: A Meta-Analytic Review. *Industrial Marketing Management, 90*(October), 435-452.

Onwuegbuzie, A. J., & Frels, R. (2016). *7 Steps to a Comprehensive Literature Review.* London: SAGE Publications Ltd.

Paesbrugghe, B., Rangarajan, D., Hochstein, B., & Sharma, A. (2020). Evaluation of Salespeople by the Purchasing Function: Implications for the Evolving Role of Salespeople. *Journal of Personal Selling & Sales Management, 40*(4), 289-305.

Palmatier, R. W., Houston, M. B., & Hulland, J. (2018). Review Articles: Purpose, Process, and Structure. *Journal of the Academy of Marketing Science, 46*(1), 1-5.

Parvinen, P., Aspara, J., Kajalo, S., & Hietanen, J. (2013). Sales Activity Systematization and Performance: Differences Between Product and Service Firms. *Journal of Business & Industrial Marketing, 28*(6), 494-505.

Patton, M. Q. (2015). *Qualitative Research and Evaluation Methods,* London: SAGE Publications Ltd.

Paulhus, D. L. (2002). Socially Desirable Responding: The Evolution of a Construct. In Braun, H. I., Jackson, D. N., & Wiley, D. E. (Eds.). *The Role of Constructs in Psychological and Educational Measurement* (pp 49-69). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-T., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology, 88*(5), 879-904.

Pullins, E. B., Timonen, H., Kaski, T. & Holopainen, M. (2017). An Investigation of the Theory Practice Gap in Professional Sales. *Journal of Marketing Theory and Practice, 25*(1), 17-38.

Rangarajan, D., Dugan, R., Rouziou, M., & Kunkle, M. (2020). People, Process, and Performance: Setting an Agenda for Sales Enablement Research. *Journal of Personal Selling & Sales Management, 40*(3), 213-220.

Rapp, A., Gabler, C., & Ogilvie, J. (2020). A Holistic Perspective of Sales Research: Areas of Consideration to Develop More Comprehensive Conceptual and Empirical Frameworks. *Journal of Personal Selling & Sales Management, 40*(4), 227-233.

Rauyruen, P., & Miller, K. E. (2007). Relationship Quality as a Predictor of B2B Customer Loyalty. *Journal of Business Research, 60*(1), 21-31,

Rich, G. A., Bommer, W. H., MacKenzie, S. B., Podsakoff, P. M., & Johnson, J. L. (1999). Methods in Sales Research: Apples and Apples or Apples and Oranges? A Meta-Analysis of Objective and Subjective Measures of Salesperson Performance. *Journal of Personal Selling & Sales Management, 19*(4), 41-52.

Richard, P. J., Devinney, T. M., Yip, G. S., & Johnson, G. (2009). Measuring Organizational Performance: Towards Methodological Best Practice. *Journal of Management, 35*(3), 718-804.

Richins, M. L. (2004). The Material Values Scale: Measurement Properties and Development of a Short Form. *Journal of Consumer Research, 31*(1), 209-219.

Robson, C., & McCartan, K. (2016). *Real World Research: A Resource for Users of Social Research Methods in Applied Settings.* West Sussex: John Wiley & Sons Ltd.

Schmitz, C., Lee, Y.-C., & Lilien, G. L. (2014). Cross-Selling Performance in Complex Selling Contexts: An Examination of Supervisory- and Compensation-Based Controls. *Journal of Marketing, 78*(3), 1-19.

Seale, C. (2009). Validity, Reliability and the Quality of Research. In Seale, C. (Eds.). *Researching Society and Culture* (pp 71-83). London: SAGE Publication Ltd.

Siguaw, J. A., Kimes, S. E., & Gassenheimer, J. B. (2003). B2B Sales Force Productivity: Applications of Revenue Management Strategies to Sales Management. *Industrial Marketing Management, 32*(7), 539-551.

Singh, R., & Abraham, K. (2010). Determinants of B2B Salespersons' Performance and Effectiveness: A Review and Synthesis of Literature. *Journal of Business & Industrial Marketing, 25*(7), 535-546.

Slagstad, R. (1995). *Positivisme og Vitenskapsteori: Et Essay om den Norske Positivismestriden.* Oslo: Universitetsforlaget.

Snyder, H. (2019). Literature Review as a Research Methodology: An Overview and Guidelines. *Journal of Business Research, 104*(November), 333-339.

Solberg, B. (2001). Vitenskap og Etikk. In Seglen, E. (Eds.) *Vitenskap, Teknologi og Samfunn.* Oslo: J. W. Cappelens Forlag a.s.

Song, F., Parekh, S., Hooper, L., Loke, Y. K., Ryder, J., Sutton, A. J., Hing, C., Kwok, C. S., Pang, C., & Harvey, I. (2010). Dissemination and Publication of Research Findings: An Updated Review of Related Biases. *Health Technology Assessment, 14*(8), 1-219.

Steenkamp, J.-B. E. M., De Jong, M. G., & Baumgartner, H. (2010). Socially Desirable Response Tendencies in Survey Research. *Journal of Marketing Research, 47*(2), 199-214.

Steneck, N. H. (2006). Fostering Integrity in Research: Definitions, Current Knowledge, and Future Directions. *Science and Engineering Ethics, 12*(1), 53-74.

Steward, M. D., Walker, B. A., Hutt, M. D., & Kumar, A. (2010). The Coordination Strategies of High-Performing Salespeople: Internal Working Relationships That Drive Success. *Journal of the Academy of Marketing Science*, *38*(5), 550-566.

Storbacka, K., Ryals, L., Davies, I. A., & Nenonen, S. (2009). The Changing Role of Sales: Viewing Sales as a Strategic, Cross-Functional Process. *European Journal of Marketing, 43*(7/8), 890-906.

Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining.* Essex: Pearson Education Limited.

Tourangeau, R., & Yan, T. (2007). Sensitive Questions in Surveys. *Psychological Bulletin, 133*(5), 859-883.

Töytäri, P., & Rajala, R. (2015). Value-Based Selling: An Organizational Capability Perspective. *Industrial Marketing Management 45*(February), 101-112.

Verbeke, W., Dietz, B., & Verwaal, E. (2011). Drivers of Sales Performance: A Contemporary Meta-Analysis–Have Salespeople Become Knowledge Brokers? *Journal of the Academy of Marketing Science, 39*(3), 407-428.

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant Validity Testing in Marketing: An Analysis, Causes for Concern, and Proposed Remedies. *Journal of the Academy of Marketing Science, 44*(1), 119-134.

Wang, G., Dou, W., & Zhou, N. (2012). The Interactive Effects of Sales Force Controls on Salespeople Behaviors and Customer Outcomes. *Journal of Personal Selling & Sales Management, 32*(2), 225-243.

Wang, C., Hoegg, J., & Dahl, D. W. (2018). The Impact of a Sales Team's Perceived Entitativity on Customer Satisfaction. *Journal of the Academy of Marketing Science, 46*(3), 190-211.

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association, 58*(301), 236-244.

Wildemuth, B. (1993). Post-Positivistic Research: Two Examples of Methodological Pluralism. *Library Quarterly 63*(4), 450-468.

Williams M., & Attaway, J. S. (1996). Exploring Salespersons' Customer Orientation as a Mediator of Organizational Culture's Influence of Buyer-Seller Relationships. *Journal of Personal Selling & Sales Management, 16*(4), 33-52.

Zallocco, R., Pullins, E. B., & Mallin, M. L. (2009). A Re-Examination of B2B Sales Performance. *Journal of Business & Industrial Marketing, 24*(8), 598-610

Zoltners, A. A., Sinha, P., & Lorimer, S. E. (2008). Sales Force Effectiveness: A Framework for Researchers and Practitioners. *Journal of Personal Selling & Sales Management, 27*(2), 115-131.

Zoltners, A. A., Sinha, P., & Lorimer, S. E. (2012). Breaking the Sales Force Incentive Addiction: A Balanced Approach to Sales Force Effectiveness. *Journal of Personal Selling & Sales Management, 32*(2), 171-186.

6. THE FOUR RESEARCH PAPERS

Paper 1:

Measures to assess B2B sales performance outcomes: A systematic review and future directions

Paper 2:

Desired outcomes from B2B selling: A systematic review and conceptualization

Paper 3:

Methods to assess outcomes from B2B selling: A systematic review, cross-journal examination, and guidelines

Paper 4:

Data sources to assess sales performance outcomes

Utgitt i ph.d. serie ved Handelshøgskolen:

Nr. 1 – 2003   Lars Øystein Widding
                Bygging av kunnskapsreservoarer i teknologibaserte nyetableringer

Nr. 2 – 2005   Pawan Adhikari
                Government Accounting in Nepal: Tracing the Past and the Present

Nr. 3 – 2005   Tor Korneliussen
                The Relationship between Initation, Barriers, Product Quality and Interna-
                tionalization

Nr. 4 – 2005   Bjørn Willy Åmo
                Employee innovation behavior

Nr. 5 – 2005   Odd Birger Hansen
                Regnskap og entreprenørskap. En fortolkende studie av hvordan to entre-
                prenører bruker regnskap

Nr. 6 – 2006   Espen John Isaksen
                Early Business Performance
                - Initial factors effecting new business outcomes

Nr. 7 – 2006   Konstantin Timoshenko
                Russian Government Accounting:
                Changes at the Central level and at a University

Nr. 8 – 2006   Einar Rasmussen
                Facilitating university spin-off ventures
                -an entrepreneurship process perspective

Nr. 9 – 2007   Gry Agnete Alsos
                Portfolio Entrepreneurship - general and farm contexts

Nr. 10 – 2007   Elsa Solstad
                Tre sykehus - to verdener - en fusjon.
                En studie av reorganisering i et helseforetak

Nr. 11 – 2007   Levi Gårseth-Nesbakk
                Experimentation with accrual accounting at the central government level
                in Norway - how a global phenomenon becomes a local practice

Nr. 12 – 2007   Tatiana Iakovleva
                Factors Associated with new venture performance:
                The context of St. Petersburg

This thesis investigates the assessment of sales performance outcomes in business-to-business (B2B) sales research. These outcomes range from economic outcomes, such as sales revenue, to outcomes associated with salespeople's customer interactions, such as customer satisfaction. Sales researchers frequently use these outcomes as dependent variables to identify antecedents to improved selling, and sales managers have an intense focus on optimizing these outcomes.

This thesis investigates the conceptualization, operationalization, data sources, and respondents used to assess the outcomes. A systematic literature review of 139 studies generated data to investigate these methodological issues in four research papers.

The first paper investigates the measures used to assess the outcomes, and the second paper develops a theoretical framework that conceptualizes outcomes from B2B selling. The third paper investigates the number of measures, the use of objective and subjective measures, and the respondent types used to assess the outcomes. Finally, the fourth paper investigates the data sources used to assess various outcomes.

In summary, this thesis reveals the large variety of quality and sophistication in the methods researchers used to assess outcomes from B2B selling. Moreover, this thesis reveals the widespread use of methods that, according to previous research, do not provide the most reliable and valid assessments—for example, the use of few revenue-focused measures, subjective measures, self-ratings, and single-source measures as well as a mismatch between data sources and collected measures. This finding should encourage many sales researchers to reevaluate their methods used to assess these outcomes. This thesis suggests theoretical frameworks, guidelines, and future research to help researchers and managers improve their assessments of outcomes from B2B selling.