2020

Supply Chain Leaders' Strategies for Selecting Performance Measures

Elizabeth Brush Nohe
Walden University

Walden University

College of Management and Technology

This is to certify that the doctoral dissertation by

Elizabeth B. Nohe

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Branford McAllister, Committee Chairperson, Management Faculty
Dr. Karina Kasztelnik, Committee Member, Management Faculty
Dr. Lisa Barrow, University Reviewer, Management Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2020

Abstract

Supply Chain Leaders' Strategies for Selecting Performance Measures

by

Elizabeth B. Nohe

MPhil, Walden University, 2019

MS, Loyola University Baltimore, 1994

BSChE, Rutgers University, 1986

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Management

Walden University

November 2020

Abstract

There is a performance gap between supply chain leaders versus laggards in critical areas

such as service and shareholder value. Although research indicates that the reasons for

the performance differences are choosing the wrong performance measures and setting

goals too low, there remains a lack of knowledge and understanding regarding the

decision-making strategies used by supply chain leaders in selecting key performance

measures and goals for their respective organizations. The purpose of this descriptive

phenomenological study was to improve the understanding of the decision-making

strategies used by supply chain leaders in their selection of key performance measures

and associated goals. The conceptual framework was decision-making theory. The

research question concerned the decision-making strategies used by supply chain leaders

in choosing the key performance measures and goals for their organizations. Fifteen

senior supply chain professionals from organizations recognized in the Gartner Top 25

Supply Chains for 2016 were interviewed. Major findings were that a majority of

participants employed multiple decision-making strategies both in choosing performance

measures and in determining goals. Naturalistic decision-making strategies were used by

all participants in performance measure selection although heuristics decision-making

strategies were used most frequently in setting goals. To create positive social change,

supply chain leaders should use multiple decision-making strategies with a focus on

naturalistic decision-making for performance measure selection and heuristics decision-

making strategies for goal setting. Doing so may increase stakeholder value benefitting

employees, the company, and the communities in which the company operates.

Supply Chain Leaders' Strategies for Selecting Performance Measures

by

Elizabeth B. Nohe

MPhil, Walden University, 2019

MS, Loyola University Baltimore, 1994

BSChE, Rutgers University, 1986

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Management

Walden University

November 2020

Dedication

This study is dedicated to my mother, Elizabeth Whinnen Evenson, who taught me so much through what she said and, more importantly, how she lived. Mom taught me to remain curious and open to new experiences; to read everything I could get my hands on; to write with proper grammar; and to be optimistic, grateful, self-reliant, and tenacious. She showed me how to overcome any obstacle and to believe I could accomplish anything I set my mind to. Most importantly, she taught me the love of a mother for a child and to live a life of dignity and grace. Although she did not live to see me finish my PhD, I can feel her love and pride in every fiber of my being.

Acknowledgments

There are so many people I wish to thank for helping me achieve my lifelong goal of earning my PhD. First, I want to thank my husband, Doug Nohe, who made this journey possible through his encouragement, love, prodding, and provision of the time and space I needed to make progress. And a big thank you to Jenny, Katie, and Kendall, my adult daughters, for letting me know how proud they were of me for working on my PhD, expressing interest in the topic, and cheering me on.

Thank you to my PhD cohort for their advice, guidance, and support, particularly Cheryl Douthitt, Lori Larson, and Putnam Texel, all of whom I met in class and who will remain lifelong friends. I also want to thank my many accountability partners who regularly checked in on and provided encouragement to me. I am especially grateful to the senior supply chain professionals who participated in my study. Thank you all for your time, patience, and candor. I learned so much from each of you.

And a special thanks to my committee: Dr. Branford McAllister, my committee chairperson, whose advice and guidance made me a better writer and researcher and who seemed to know just when I needed a push or some words of encouragement; Dr. Karina Kasztelnik, my second committee member, who provided support and guidance on conducting a descriptive phenomenological study; and Dr. Lisa Barrow, university reviewer, whose deep knowledge of APA form and style improved the readability of my study.

Table of Contents

iv

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Achieving performance goals is a benchmark that continues to distinguish leaders and laggards in the supply chain sector. According to Swink, Johnson, and Quinn (2012), there is a growing gap in organizational performance between companies considered as supply chain leaders versus laggards. Ellinger et al. (2012) found that companies recognized as leaders in supply chain performance have higher than average customer service levels and greater than average shareholder value when compared with their peers. Gilmore (2012b) surmised that the performance divide between leaders and laggards was due to laggards choosing the wrong performance metrics and/or setting their goals too low. This is despite the extensive research that has been conducted on performance measurement systems (PMSs) in recent years (Melnyk, Bititci, Platts, Tobias, & Andersen 2014).

The study was intended to extend the body of knowledge regarding supply chain PMSs by exploring the decision-making strategies used in the selection of performance measures and goals by supply chain leaders. The results of the study could provide new insights to help supply chain managers improve the design, implementation, and use of PMSs. These changes could lead to improved supply chain performance that could also contribute to improved organizational performance (see Ellinger et al., 2012).

In the remainder of Chapter 1, I provide the background for the study, the problem statement, research purpose, and research question. I give a synopsis of the conceptual framework which is decision-making theory. I then discuss the nature of the study; define key terms; and discuss the assumptions, scope, delimitations, and

limitations of the study. I describe the expected significance of the study for practice and theory and close with a brief summary of the chapter.

Background of the Study

For organizations, supply chain leadership is pivotal to achieving performance goals and unlocking shareholder value. Chan, Nayak, Raj, Chong, and Manoj (2014) stated that there is a positive correlation between supply chain performance and overall organizational performance. Cecere, Hart, Denman, and King (2016) found that publicly traded companies with supply chains characterized as leaders had higher shareholder value than those firms with lower-performing supply chains. Cecere et al. added that leaders also outperformed peer companies in measures of growth, operating margin, inventory turnover, and return on invested capital. It logically follows that knowledge that can be applied to improving supply chain performance may help overall firm performance.

Performance measurement, which has become more nuanced over the years, appears pivotal to organizations seeking to become supply chain leaders. Coe and Letza (2014) stated that early research on performance measurement was focused on financial measures. Coe and Letza criticized such measures as being too narrowly focused and providing only a short-term, and backward-looking, view of the business. Bititci, Garengo, Dörfler, and Nudurupati (2012) shared that other measurement frameworks evolved that (a) provided a more balanced view of the business, (b) enabled linkage between business strategy and operational execution, and (c) contained both leading and lagging indicators. Estampe, Lamouri, Paris, and Brahim-Djelloul (2013) provided an

overview of 16 of the most commonly referenced supply chain performance measurement models. And Hoque (2014) stated that the Kaplan and Norton Balanced Scorecard is the most commonly cited performance measurement framework in the literature.

Researchers have focused their efforts on studying how to design, implement, use, and refresh PMSs. Gutierrez, Scavarda, Fiorencio, and Martins (2015) described the essential elements to be considered in designing, implementing, and using PMSs. And Koufteros, Verghese, and Lucianetti (2014) shared potential barriers to success for each phase of the PMS lifecycle. There seems to be adequate research to guide supply chain leaders on how to create PMSs.

Researchers have also studied possible moderators and mediators between PMSs and performance. Franco-Santos, Lucianetti, and Bourne (2012) developed a framework showing the linkages between PMSs and (a) individual behavior, (b) organizational capabilities, and (c) organizational performance. Marginson, McAulay, Roush, and van Zijl (2014) showed a positive relationship between use of PMSs and managers' mental mindsets. Koufteros et al. (2014) found mixed results between maturity of PMSs and firm performance. Although there are some inconsistencies in research results, it seems logical that organizations with well-developed PMSs should have better performance than those with less developed or no measurement systems.

The literature indicates that there is a strong linkage between the design, implementation, use, and refresh processes of a PMS and organizational performance (Franco-Santos et al., 2012; Koufteros et al., 2014). Waal and Kourtit (2013) stated that

simply implementing a PMS does not guarantee improved firm performance. Melnyk et al. (2014) observed that although there exists a plethora of literature related to performance measurement, it appears managers still lack the tools needed to effectively lead their organizations. Specifically, there is a lack of research on the decision-making strategies actually used by supply chain leaders in selecting their key performance measures and goals. This knowledge could help supply chain managers improve the design of their respective PMSs and bolster overall organizational performance.

Problem Statement

There exists a copious amount of research related to supply chain-relevant performance metrics and building PMS models. Kocaolu, Gülsün, and Tanyas, (2013) suggested the use of the supply chain operations research model although Bhattacharya et al. (2013) recommended the Balanced Scorecard. Performance metrics have been (a) categorized as strategic, tactical, or operational (Estampe et al., 2013); (b) identified as financial or nonfinancial (Hoque, 2014); (c) determined to be leading or lagging indicators; and (d) classified as quantitative or qualitative (Coe & Letza, 2014). Based on my review of the literature, there appears to be a gap in the supply chain research related to how performance measures are chosen and how goals are determined. Thus, the research problem was that there is a lack of knowledge and understanding with respect to the decision-making strategies used by supply chain leaders in selecting the key performance measures and goals for their respective organizations.

Purpose of the Study

The purpose of this qualitative descriptive phenomenological study was to improve the understanding of the decision-making strategies used by supply chain leaders in their selection of key performance measures and associated goals. Mayoh and Onwuegbuzie (2015) said that although each respondent's experiences related to a phenomenon are unique, a descriptive phenomenology can be used when the researcher is seeking to find features of the experience that are common across respondents. Research that makes salient commonalities in the decision-making strategies used by senior managers of organizations recognized as supply chain leaders could help supply chain managers in other organizations improve the strategies they use to select key performance measures and goals. Because improvements in PMS design have been associated with improved overall organizational performance (Franco-Santos et al., 2012), research that aids PMS design could favorably impact firm performance.

Research Question

The central research question for this study was, What are the decision-making strategies used by supply chain leaders in choosing the key performance measures and goals for their organizations? The research question was a logical result of the problem statement that there is a lack of knowledge and understanding in the literature regarding how supply chain leaders actually select the key performance measures and goals to drive organizational performance. The problem statement arose from two general themes in the literature. The first theme was that there is a growing performance gap between organizations recognized as leaders from those recognized as laggards (Swink et al.,

2012). The second theme was that supply chain professionals have all the tools they need to design, implement, and apply performance measurement systems (Waal & Kourtit, 2013). Therefore, supply chain managers are either selecting the wrong measures or setting their goals too low (Gilmore, 2012a).

Conceptual Framework

I explored the decision-making strategies, or combination of strategies, that are used by supply chain leaders in deciding on the key performance measures and goals for their organizations. Prior researchers studying performance measurement relied mainly on performance measurement theory (Paulraj, Chen, & Lado, 2012). There has also been research on how performance measures impact organizational performance (Coe & Letza, 2014). I applied a new lens to the study of performance measurement: decision-making theory. In Chapter 2, I delve more deeply into the evolution of performance measurement theory and highlight the gap in the literature regarding understanding how supply chain managers actually choose the key performance measures and goals for their organizations.

The conceptual framework for the study was decision-making theory. Shaban (2012) stated that there are three distinct academic schools of thought on decision-making theory: (a) rational decision-making, (b) heuristics and biases, and (c) naturalistic decision-making, which is often called *going with your gut*. Shaban asserted that decision-making models are normative, descriptive, and/or prescriptive in nature. I briefly touch on each of these three academic approaches and then go more deeply into each in Chapter 2.

Franklin (2013) stated that rational decision-making theory has roots in economic theory and the concept of the rational man. Gigerenzer (2015) described the classical rational process as (a) defining the problem/task, (b) identifying all possible solutions/outcomes, (c) evaluating each outcome based on some predetermined criteria of value, (d) selecting the outcome with the highest value, and (d) implementing the selected decision. Seminal researchers in rational decision-making theory include Von Neumann and Morgenstern (1953), who were the authors of expected utility theory (EUT). EUT describes how people should use economic value in making a decision among alternatives with known probabilities. Leonard Savage (1954) developed the subjective utility model by adding a component to the EUT that captures the decision-maker's beliefs. Rational decision-making follows a well-defined and prescriptive process.

Artinger, Petersen, Gigerenzer, and Weibler (2015) asserted that the study of heuristics and biases in decision-making arose from the need to explain decision-making that seemed to violate rules of rational decision-making theory. Gigerenzer (2015) stated that heuristics are mental shortcuts that are used when all possible outcomes and probabilities are not known or not used. Although some research has shown the benefits of the fast and frugal method of heuristics-based decision-making (Hands, 2014), other research has shown such mental shortcuts can lead to poor decisions (Kahneman, 2011). The most frequently mentioned heuristics and biases in the literature include the representativeness heuristic (Tversky & Kahneman, 1974), the availability heuristic (Braga, Ferreira, & Sherman, 2015), and the anchoring and adjustment heuristic (Cheek

& Norem, 2016). Hands (2014) added that heuristic decision-making processes are often used under conditions of uncertainty and where time and data are limited.

According to Shan and Yang (2016), a third area of decision-making research involves the idea of intuition and experience in making decisions. Klein (2015) stated that the concept of naturalistic decision-making arose out of the need to understand decision-making in real life, complex, and high-stress situations such as those faced by members of the military, police officers, and medical practitioners. Gore, Flin, Stanton, and Wong (2015) identified Gary Klein and Judith Orasanu as seminal researchers of naturalistic decision-making. Klein explained that naturalistic decision-making models could be used prescriptively to help reduce errors and to help practitioners use their skills more effectively. Riegel, Dickson, and Topaz (2013) stated that Klein's recognition-primed decision-making model is the most often applied naturalistic decision-making model in the literature. Riegel et al. applied the recognition-primed decision-making model to study how adults dealing with chronic heart failure made ongoing future self-care decisions.

Researchers have applied decision-making theory to explore the strategies used by individuals to achieve personal and organizational objectives. This research includes how travelers make route choices (Ramos, Daamen, & Hoogendoorn, 2014), how hospital leaders establish standards of care in a hospital setting (Gigerenzer, 2015), and how student athletes make decisions that enable them to balance the dual roles of being a student and an athlete (Macquet & Skalej, 2015). I applied decision-making theory to guide the development of the research protocol and data analysis. The use of the theory

led to insights that may help advance knowledge for both the scholar and practitioner communities regarding improvements in the decision-making process employed by supply chain managers in designing their PMSs.

Nature of the Study

Based on a review of the literature, I determined that a qualitative approach would be the most appropriate methodology for my study. I applied Giorgi's (2012) descriptive phenomenological procedure to study the decision-making strategies used by supply chain leaders in choosing the performance measures and goals for their respective supply chains. Yin (2014) said that the most important criterion for determining what research method to use is the question being asked. If the intent is to answer a how, why, or a what question that is exploratory in nature, the researcher should use a qualitative approach. Mayoh and Onwuegbuzie (2015) explained that qualitative research involves data collection and analysis of nonnumerical data to better understand phenomenon such as concepts or experiences. It can be used to make salient insights into a problem or to generate new areas for further research.

Giorgi (2012) advised that phenomenology is the research method of choice when the goal is to generate an increased understanding of the subjective experiences of a group of people from their own perspectives. Similarly, Mayoh and Onwuegbuzie (2015) said that a descriptive phenomenological approach is used when the researcher wants to collect the common essences of an experience within a group to create an essential structure of the phenomenon. Mayoh and Onwuegbuzie added that an important distinction of the descriptive phenomenological approach from other phenomenological

perspectives is that there is one correct objective interpretation of the common essences of the phenomenon. The phenomenon for this study was the description of the decision-making strategies used by supply chain leaders in choosing the key performance measures and goals for their supply chain organizations.

Definitions

The following are terms along with their definitions that were used in this study.

Balanced Scorecard (BSC): A method developed by Kaplan and Norton (1996) to balance the lagging indicators of financial performance with leading indicators that include the financial view, the customer view, the internal process view, and the learning and growth view of organizational performance.

Heuristic: A mental shortcut that reduces decision-making time thus allowing people to solve problems and make judgments more quickly and efficiently (Gigerenzer, 2015).

Key performance indicator: Information, collected at regular intervals, that tracks the performance of a system (Wu, 2012). Indicators may be quantifiable (e.g., customer service) or nonquantifiable (e.g., customer satisfaction).

Meaning unit: A natural chunk of text that contains a single thought as determined by the researcher from a phenomenological psychological perspective (Brinkmann & Kvale, 2015). The primary purpose for meaning units is to help break down a description into more manageable pieces.

Naturalistic decision-making theory: A theory that is used to explore how decisions are made by experts in complex natural environments (Klein & Wright, 2016).

Performance measurement system: The interaction of principles and actions whose intent is to deliver on the objectives and strategies identified by the business through changing the behavior of individuals in an organization (Choong, 2013).

Rational decision-making theory: A decision-making theory that relies on the use of logic and/or statistics (Gigerenzer, 2015).

Supply chain management: The integration of business processes that span from the customer's customers through to the supplier's suppliers (Seuring, 2013).

Supply chain operations reference (SCOR) model: A model that initially referred to the supply chain management processes that were described by the plan, source, make, and deliver processes. Subsequent versions of the model included the return process (Estampe et al., 2013).

Assumptions

Myers (2013) stated that although assumptions in a study may be somewhat out of the researcher's control, if those assumptions went away, the study would no longer be relevant. Myers added that is imperative not only to state the assumptions, but also to demonstrate the steps that will be taken to ensure these assumptions are *likely* true. Following is a list of assumptions used in this study:

The participants in this study were assumed to be candid and transparent about their experiences regarding the decision-making strategies employed in choosing supply chain key performance measures and goals. In order to foster an environment of trust, so that candid conversation can occur, the identities of all respondents and their

organizations were kept confidential. In order to help reduce researcher bias, I clearly explained the purpose of the study in as neutral a way as possible.

The participants in the study either led or were a participant in the decision-making process of selecting the supply chain key performance measures and/or goals for their respective organizations. As part of the participant selection criteria, I chose supply chain leaders with the words executive vice president, vice president, senior director, director, or senior manager in their position titles. I asked potential respondents to describe their roles in PMS design for additional participant validation.

The research question in this study focused on describing the decision-making strategies used by supply chain leaders in selecting key performance measures and goals. In constructing questions as part of an interview protocol, Dikko (2016) recommended testing the questions prior to commencing with the main study. Therefore, I conducted a pilot study to allow for adjustment in the interview protocol prior to conducting the larger study.

A key element of phenomenological research, as described by Giorgi (2012), is that reality is created by each person as they assign meaning to their individual experiences related to a phenomenon. Giorgi cautioned that researchers could bring their perspectives to interpreting others' experiences. It was imperative that I remained aware of this potential bias and worked diligently to ensure that the phenomenon being studied represented the perspectives of the participants over my own. Sousa (2014) recommended that the researcher utilize member-checking to ensure that the results reflect the participants' experiences relative to the phenomenon being studied.

Scope and Delimitations

According to Myers (2013), the scope and delimitations define the boundaries of the study. These boundaries were within my control. The research problem is one of the first areas of boundary definition and is contained in the purpose statement section of this chapter.

A goal was to study the decision-making strategies used by supply chain leaders in choosing their key performance measures and goals. I took the purposive sample from companies that had been identified in prior research as having leading supply chain organizations. Cecere et al. (2016) conducted their third "Supply Chains to Admire" study, in which they researched 320 publicly traded companies across 31 industries to determine which companies had the best supply chain performance. Cecere et al. compared each company against its industry-specific peer group in terms of shareholder value, growth, operating margin, inventory turns, and Return on Invested Capital (ROIC) for the years 2009-2015. Cecere et al. looked for sustained and consistent improvement. Cecere et al. categorized the companies as winners, finalists, or underperformers.

Cecere et al. (2016) found that companies from 19 of 31 industries earned the winner or finalist designation. The 16 companies categorized as winners comprised five percent of the participant pool. There were 21 companies, representing seven percent of the participant pool, identified as finalists. The remaining 88% of companies made up the underperformers group. Cecere et al. looked at year-over-year trends and found that most companies were regressing on the supply chain metrics considered in the study. Cecere et al. noted that although some of the underperformers were making progress on single

metrics, overall most of the companies were failing to show progress across the balanced set of metrics. Some of the companies identified as underperformers in the Cecere et al. study have been highlighted in other research as having best-in-class supply chains (Aronow, Burkett, Nilles, & Romano, 2016).

I targeted senior level supply chain professionals from either the 16 companies categorized as winners of the "Supply Chains to Admire" study (Cecere et al., 2016) or from the Gartner Supply Chain Top 25 (Aronow et al., 2016) as the initial participant pool for the study. My preference was to talk to more than one senior level supply chain manager from each company as part of the validation process. If I had been unable to recruit the targeted 25 participants, I had planned to approach senior level supply chain professionals on the list of 21 finalist companies from the same study.

Limitations

Myers (2013) said that limitations are the potential weaknesses of a study and are often out of the researcher's control. However the researcher must determine how to minimize the impact of the limitations on the study. The following is a list of limitations for this study:

At the time of the interviews, I held the chief supply chain role where I worked. Therefore, participants might knowingly or unknowingly have modified their responses based on what they thought I wanted or expected to hear during the interview. Malterud (2012) cautioned that the researcher remain aware of such investigator effects and approach the engagement with authenticity, caring, and trustworthiness. I openly shared

my background and the reason for my interest in the research problem and then approached analyzing the data using bracketing to avoid as much bias as possible.

Malterud (2012) highlighted that a limitation of descriptive phenomenology could occur during the process of de-constructing the data in which the individual context may get lost. Malterud added that an important aspect of the method to mitigate this limitation occurs during the final step of re-contextualizing the data. It is imperative that the researcher validates the resulting output against the original transcripts.

According to Sousa (2014), one goal of qualitative research is to construct a framework of concepts that can be applied to a more general population. Sousa added that in order to ensure transferability of a study it is imperative that the researcher ensures trustworthiness of the data collection, analysis, and reporting processes. I provided sufficient detail for each step in the process so that someone reading the study would have an adequate understanding of both the context and the steps taken to analyze, categorize, reduce, and reconstruct the common experience of the participants in choosing the key performance measures and goals for their organizations. I also used member-checking at key steps in the process.

Significance of the Study

Companies with strong supply chain organizations provide better service to their customers (Ellinger et al., 2012) and better return to their shareholders (Cecere et al., 2016). Melnyk et al. (2014) stated that PMSs allow for the translation of strategy into language that makes the strategy operational, thus enhancing organizational performance. Franco-Santos et al. (2012) acknowledged that the design, implementation, and use of

PMSs have been shown to be problematic. Franco-Santos et al. added that research that addresses these problem areas could lead to improved organizational performance. Improvements in the design, implementation, and use of performance measures and goals could also lead to positive social change for the individual, the organization, and the company. In the succeeding subsections I explain the potential significance of the study in three areas: theory, practice, and social change.

Significance to Theory

The study was intended to fill a gap in the literature by exploring the strategies used by supply chain leaders in the development of their PMSs through the lens of decision-making theory. Based on a review of the literature, I applied a unique lens to the study of performance measures. Applying a new lens to a topic that has been heavily researched over the last several decades could provide some new and interesting insights. One goal of this qualitative descriptive phenomenological study was to make transparent the common experiences of supply chain leaders as they employed decision-making strategies to select their key performance measures and goals. A second goal of the study was the development of a framework that describes these common decision-making strategies.

Significance to Practice

The framework developed as part of the output for this study could act as a trigger for supply chain managers in other organizations to make salient the decision-making strategies they use in choosing their key performance measures and goals. Making the invisible, visible could allow these supply chain managers to reflect upon the decision-

making strategies they have used in developing their key performance measures and goals. The framework could also act as a decision-making model to help guide supply chain managers in designing or redesigning the PMSs for their organizations. The output of this study could act as a framework to be used by supply chain managers for communicating to other members in the organization the decision-making strategies used to develop the PMS, potentially increasing understanding of, alignment with, and commitment to the chosen performance measures and goals.

Significance to Social Change

The results of this study could lead to positive social change in several ways. Gutierrez et al. (2015) observed that PMS implementations were sometimes accompanied by fear and resulting subversive behaviors of employees and managers due to the transparency performance measures provided. Marginson et al. (2014) observed that well-designed PMSs allowed managers to change their mental models from an orientation of command and control to one supporting a learning and growth environment that further enhanced employee trust and commitment to the PMSs. Research that helps improve PMS design through the use of well-chosen decision-making strategies could increase employees' and managers' trust in how performance measures were chosen thereby improving the work environment and individual performance.

Franco-Santos et al. (2012) found that well-designed PMSs help senior executives focus the organization on the vital few strategies. Franco-Santos added that strong PMSs favorably impacted strategy processes through (a) manager engagement in strategy development and review, (b) translation of strategic plans into operational language, and

(c) through changing managers' mindsets such that strategy is seen as a continuous process rather than a one-time event. Research that helps improve the decision-making processes used in designing PMSs through the selection of appropriate key performance measures and goals could help develop executives' strategic planning and implementation skills leading to improved organizational performance.

According to Franco-Santos et al. (2012), prior research has shown a favorable relationship between the quality of PMSs and company performance. Saunila, Pekkola, and Ukko (2014) found that performance measurement favorably impacted the relationship between innovation capability and organizational performance. Marginson et al. (2014) observed that psychological empowerment mediated the relationship between PMSs and organizational performance. Abushaiba and Zainuddin (2012) demonstrated that favorable PMS design affects a firm's competitiveness thereby affecting company performance. It logically follows that research that helps improve the selection of key performance measures and goals could help improve company performance leading to positive social change for the various stakeholders of the company including the employees, the community, and other shareholders.

Summary and Transition

Chapter 1 was focused on establishing the need for the study: a widening performance gap between supply chain leaders and laggards that is adversely impacting overall company performance. The proposal to undertake a descriptive phenomenological study arose out of the gap in the literature and the lack of understanding of the actual decision-making strategies that supply chain leaders use to establish key performance

measures and goals for their respective supply chains. The overarching research question was presented. Decision-making theory as the conceptual framework was proposed. Justification for the descriptive phenomenological research methodology as the nature of the study was presented along with a recommended sample population, sample size, and high-level overview of the research process. Key terms were defined. Assumptions, scope, delimitations, and limitations of the study were articulated. The chapter closed with a discussion of the significance of the proposed study to theory, practice and positive social change.

In Chapter 2 I provide a more thorough exploration of the literature on PMSs in general and more specifically on supply chain metrics. Decision-making theory, the conceptual framework for the study, is discussed in greater detail. The selection of research methodology is explained and supported.

Chapter 2: Literature Review

There exists extensive research on various performance measurement frameworks that could be used for monitoring and improving supply chain performance (see Coe & Letza, 2014; Estampe et al., 2013; Okongwu et al., 2015). There is also a large corpus of literature on the design and implementation of PMSs (see Bourne et al., 2000; Franco-Santos et al., 2012; Gutierrez et al., 2015). In the literature review, I describe and critically compare the most important studies among this body of research.

The research problem was that there is a lack of knowledge and understanding with respect to the decision-making strategies used by supply chain leaders in selecting the key performance measures and goals that will guide the organization to improved performance. The purpose of this descriptive phenomenological study was to explore the decision-making strategies used by supply chain leaders in choosing the performance measures and goals for their respective supply chains.

According to Chan et al. (2014), there exists a logical link between performance of the supply chain and the overall performance of the organization. Deshpande (2012) stated that ever-increasing concerns around global uncertainty along with a rise in prosperity in developing geographies has made supply chain performance even more critical to the sustainability of the broader organization. Chan added that competition has expanded beyond organizational boundaries and is now between respective supply chains. Estampe et al. (2013) acknowledged that although measuring performance of supply chains is complex, it is imperative that such measurement is done and done well.

Chalyvidis, Ogden, and Johnson (2013) proclaimed that a new research lens is needed to overcome the barriers of inadequate PMSs in the supply chain.

Gopal and Thakkar (2012) found that the evolution of supply chain PMSs mirrored the development of PMSs in the accounting and general business management literature. Early researchers focused on the development of specific metrics for supply chain to augment the existing financial metrics (Gopal & Thakkar, 2012). Over time, various performance measurement frameworks such as the SCOR model and the BSC emerged. Gutierrez et al. (2015) added that research on methods to design and implement PMSs soon followed. The emergence of literature on what to measure and how to implement a PMS suggests that supply chain leaders had the tools necessary to select the performance measures and associated goals needed to improve the performance of their respective supply chains. Gilmore (2012b) wrote that supply chain leaders were reporting that they were meeting or exceeding their performance measurement targets despite the findings of Swink et al. (2012) who conducted a broad global survey and found a growing gap between what they coined as supply chain *leaders and laggards*. It appeared that practitioners, despite having what seemed to be the needed tools, were either choosing the wrong metrics or setting goals that were too low (Gilmore, 2012a).

I begin Chapter 2 by describing the literature search strategy. Then, I discuss decision-making theory, the conceptual framework for this study. I summarize and synthesize the literature related to the evolution of research on performance measurement. I delve into more detail on decision-making theory and its use in prior research. Then, I synthesize the research on the application of descriptive

phenomenology to the specific research question and why this method is a meaningful approach. Finally, I summarize the major themes from the literature including both what is known and what is not known related to the decision-making strategies used in selecting performance measurements and goals. I conclude the chapter by describing how the study contributes to the body of knowledge on performance measurement in the supply chain followed by a brief transition to Chapter 3.

Literature Search Strategy

I used multiple research databases for the literature review including EBSCOhost, ProQuest ABI/INFORM, Elsevier Science Direct, and SAGE Journals. For supply chain related topics, I used the Gartner and APICS (American Production and Inventory Control Society) databases. Also, I used Google Scholar. I used the following keywords in the search:

- *performance measures*,
- *metrics*,
- *supply chain management*,
- *supply chain metrics*,
- *supply chain performance*,
- *decision-making*,
- *rational decision-making*,
- *heuristics and biases*,
- *naturalistic decision-making*,
- *qualitative methodology*,

- *phenomenology*, and

- *case study*.

I also sought out specific performance measurement frameworks such as BSC and SCOR. I searched for specific decision-making theories including *expected utility theory, prospect theory, SP/A theory, satisficing heuristic, recognition heuristic, take-the-best heuristic,* and *recognition primed decision-making.* I looked for a combination of terms including (a) *supply chain* and *firm performance,* (b) *metrics* and *firm performance,* (c) *metrics* and *decision-making theory,* and (d) *supply chain* and *decision-making* theory. I used the process of citation chaining to find additional articles of interest. I continued the chaining process until I found no additional authors or articles on the specific search area. Additionally, I explored seminal authors on performance measurement including Kaplan, Norton, Ittner, Neely, Bourne, Bititci, Nudurupati, and Franco-Santos. Likewise, I surveyed seminal authors on decision-making theory including von Neumann, Morganstern, Tverskey, Kahneman, Gigerenzer, Grassmaier, Klein, and Orasanu.

To find literature on the research methodology for the study, I explored the various methods of qualitative research. Then, I searched specific methods under the broader heading of phenomenology and the two major approaches, *descriptive phenomenology* and *hermeneutic phenomenology*. I also surveyed seminal researchers of phenomenology including Husserl, Moustakas, and Giorgi.

Conceptual Framework

The phenomenon for the study was the decision-making strategies used by supply chain leaders in the selection of performance measures and associated goals to improve

the performance of their respective organizations. Prior researchers studying performance measurement relied mainly on performance measurement theory. There had been a substantial amount of focus on performance measurement taxonomies (Paulraj et al., 2012) and the development of many performance measurement frameworks such as Kaplan and Norton's BSC (Kaplan, 2012), the SCOR model (Ntabe, LeBel, Munson, & Santa-Eulalia, 2015) and the EFQM Excellence model (EFQM, 2015). In addition, there is abundant research on how to design, implement, use, and refresh PMSs (e.g., Agostino & Arnaboldi 2012). There has also been research on the linkages between performance measures and strategy and how performance measures can influence behavior thus affecting organizational performance (Coe & Letza, 2014).

Maxwell (2012) stated that one of the roles of theory is to provide new insights into a phenomenon. I applied a new lens to the study of the selection of performance measures and goals by focusing on the word *selection*. I drew upon decision-making theory as the conceptual framework for the study to explore the decision-making strategies used by supply chain leaders in selecting the performance measures and goals for their organizations.

According to Gigerenzer (2015), there are three general approaches to decision-making research: (a) rational decision-making, (b) heuristics and biases, and (c) naturalistic decision-making. Von Neumann and Morgenstern (1953) and Savage (1954) were seminal researchers in the development of normative rational models of decision-making under risk. These expected utility theory models were based on the notion that the decision-maker would consider all possible choices for a decision and then assess

each choice by assigning a probability value for each expected outcome of that choice (Savage, 1954; Von Neumann & Morgenstern, 1953). The decision-maker would then determine a maximum utility by a process of weighting and summing; ultimately choosing the option with the greatest value. Fennell and Baddeley (2012) added that expected utility models have acted as a foundation for the further development of rational decision-making theory including Bayes's theory, which should be used to update the probability of an event occurring in light of new information.

In the 1970s empirical data emerged indicating that decision-makers regularly violated the rules of rational decision-making (Gigerenzer, 2015). Tversky and Kahneman (1974) were leaders in the development of the heuristics and biases research program. Researchers in the heuristics and biases programs brought to light many models that described how people regularly and routinely demonstrated deficiencies in rational decision-making. Gigerenzer (2015) stated that some of the most commonly cited heuristics and biases included (a) the availability heuristic, (b) the anchoring and adjustment heuristic, (c) the recognition heuristic, and (d) the satisficing heuristic.

Hands (2014) offered a different perspective on the use of heuristics in the fast and frugal heuristics research program. Gigerenzer (2015) argued that the use of heuristics would not necessarily result in a trade-off between effort and accuracy. Gigerenzer added that researchers in the fast and frugal heuristics program demonstrated empirically that in many cases heuristics outperformed rational models in both the speed and the quality of decisions.

According to Artinger et al. (2015), much of the heuristics and bias decision-making research has been applied to laboratory settings intended to demonstrate empirically that people routinely violate the rules of rational-decision-making. Such research generated many descriptive models of decision-making. Gigerenzer (2015) shared research showing that individuals, including physicians, did not apply sound logic when it came to understanding statistics related to health screening tools for cancer resulting in considerable unnecessary health care costs. Gigerenzer argued that individuals needed more education in the concept of risk because, although screening tests can be beneficial for detecting health issues, there may also be negative consequences including adverse physical and psychological effects of false positive results. Gigerenzer stated that the recognition heuristic could be used to predict preferences such as the outcome of elections. But other researchers (Gore et al., 2015) argued that laboratory studies using college students failed to explain how experts incorporated experience into their decision-making processes.

Klein (2015) stated that research showed it was becoming increasingly clear how people did not make decisions. And decision-making quality, using rational models, had not improved over time. Klein added that scholars and practitioners were seeking models that provided insights into how people actually made decisions in real-life situations. Klein stated that naturalistic decision-making emerged in the late 1980s as a result of frustration by some researchers in applying traditional analytical models of decision-making to real and complex environments.

According to Gore et al. (2015), Gary Klein and Judith Orasanu, pioneers in naturalistic decision-making research, created an invitation only event in 1989, assembling other researchers that were studying decision-making in field settings. Gore et al. highlighted that Rasmussen's cognitive control model, Hammond's cognitive continuum theory, and most notably, Klein's recognition-primed decision-making model were some of the models that emerged as part of the naturalistic decision-making research programs. Although these models were developed independently, there were some common themes among the models. Decision-makers were relying on previous experiences to categorize new situations. Decision-makers were using some type of process to make these categorizations. And the categorization provided a framework for some form of action.

Klein (2015) stated that decision-makers, under the naturalistic decision-making approach, were portrayed as drawing upon prior experience and knowledge to actively manage their actions in dynamic environments rather than evaluating the potential outcomes from a list of possible options and then waiting to see if such predictions materialized. These naturalistic decision-making models contrasted expert from novice behavior in both the quality and the process of making decisions. Gore et al. (2015) shared that naturalistic decision-making research has been applied to many real-life decision-making environments including healthcare settings, firefighters, police officers, military personnel, elite athletes, and situations involving labor relations.

Because I was not able to find prior studies in which researchers used decision-making theory as the conceptual lens for performance measurement. And because the

research on decision-making fell under three main conceptual programs: (a) rational decision-making theory, (b) heuristics and biases theories, and (c) naturalistic decision-making, I took a comparative approach to exploring the phenomenon; the decision-making strategies supply chain leaders use to select the key performance measures and goals for their respective organizations. Yin (2014) stated that in taking a comparative structures approach, the researcher addresses the data multiple times, each time analyzing the data through a different conceptual lens.

A highly regarded example of the comparative structures approach is Allison's (1971) research on the Cuban Missile Crisis of 1962. Allison interpreted the thirteen most crucial days of the Cuban Missile Crisis through three lenses. The first lens Allison used was the *Rational Actor* model influenced heavily by economists, statesmen, and game theory. The model assumed that individuals behave in a rational manner with perfect knowledge of the situation. The decision-maker should consider all options and select the option with maximum utility.

The second lens Allison (1971) used was the *Organizational Process* model in which the decision-maker uses a satisficing heuristic to achieve a minimum goal while also minimizing risk. The third lens applied by Allison was the *Bureaucratic Politics* model in which the decision-making by the head-of-state is influenced by internal politics and negotiation with the leaders of the state, each of whom have different levels of influence with the decision-maker. Leaders might take actions that would not have been approved by the collective group.

Allison (1971) also used the *Organizational Process and Bureaucratic Politics* models to explain what appeared to be other irrational decisions in military history including the Japanese invasion of Pearl Harbor. He argued that assuming that humans would behave in a rational manner was dangerous and could lead to catastrophic results. Allison's research provided new insights into the ways leaders made decisions in times of crisis hopefully adding to the quality of future decisions. Likewise, the results of my study provided new insights into how supply chain leaders decide on key performance measures and goals thereby assisting other supply chain managers in improving the effectiveness of their respective PMSs.

<div align="center">

Literature Review

</div>

Overview of Performance Measurement Systems

Definition of PMSs. Franco-Santos et al. (2012) stated that researchers have not come to a common agreement on the definition of PMSs. Franco-Santos et al. argued that inconsistency in the definition of PMSs and related terms among researchers not only leads to confusion but also limits the generalizability of research. Choong (2013) added that lack of consistent definition also makes comparisons across studies onerous and creates difficulty in both drawing conclusions from and applying research to practice. I used Choong's definition that a PMS includes not only the measures and goals but also the performance of resources and processes within an organization.

Choong and Burgess (2014) pointed out that many of the preeminent researchers on PMSs use terms such as metric, measure, indicator, and performance measure interchangeably thereby adding to the research confusion. Trochim, Donnelly, and Arora

(2014) stated that the terms *measure, metric*, and *indicator* each have a specific

definition. A measure was defined as the numerical result of measurement. The number

of customer orders processed is an example of a measure.

Trochim et al. (2014) defined a metric as a ratio that indicates the extent to which

an objective is being met. An example of a performance metric for managers trying to

improve service to their customers could be number of cases shipped as a percent of

cases ordered. Choong (2013) said that a metric not only measures the extent that a

process meets a specific objective, the objective must be important to stakeholders.

Choong (2013) explained that an indicator allows one to understand how far a

project, process, or organization is from achieving a goal and whether or not there is

forward progress. A well-constructed indicator may be composed of both quantitative and

qualitative elements, is understandable, and is easy to measure. Indicators are valuable

tools for softer measurements. Customer satisfaction is an example of an indicator.

Evolution of Performance Measurement Systems

Pre-industrial accounting stage. Bititci et al. (2012) stated that performance

measurement has evolved over time to support and respond to global business trends.

Bititci et al. pointed out that accounting systems and, therefore, PMSs remained relatively

unchanged from the time of the introduction of the double entry book keeping during the

fifteenth century until the advent of the Industrial Revolution in the nineteenth century.

Financial measures and the Industrial Revolution stage. According to Kaplan

(1984), the rapid growth of the railroad, manufacturing, and distribution sectors required

a substantial evolution in cost accounting and management controls in order to monitor

and manage multi-functional and multi-divisional organizations. Cost accounting systems were needed to handle a large transactional volume and to provide summary level financial information that reflected the performance of many units spanning a broad geography. Kaplan stated that in the second stage of PMS evolution, metrics such as cost per ton-mile and operating income relative to sales, allowed industrial leaders like Andrew Carnegie and Pierre du Pont to focus on improving the cost positions of their respective financial empires. Kaplan added that in the latter years of the Industrial Revolution the focus turned to productivity with the advent of Taylor's Scientific Management and the desire to break work down into repeatable and measurable tasks. This resulted in the development of time and cost standards and the addition of a management layer to oversee the establishment and implementation of such standards.

According to Kaplan (1984), Pierre S. du Pont created the concept of return on investment (ROI). Du Pont preferred using ROI rather than profits as a percent of sales as a key performance metric because ROI included the effect of capital investment on business outcomes. Brown, a financial officer in the DuPont organization, further decomposed ROI into other metrics that allowed managers to assess the impact of their departments' performance in contributing to the overall return for the company.

Kaplan (1984) explained that after World War I, du Pont, Brown, and Alfred P. Sloan extended these principles to the struggling General Motors (GM) Company, of which the DuPont Corporation was a major stockholder. The improved GM accounting system provided a vehicle to cascade operating goals to each unit, allowed for reporting of performance variances to those goals, provided senior management the information to

allocate the appropriate levels of resources, and acted as the basis for handing out rewards to each unit based on performance. The GM accounting system that evolved in the early 1920s is the basis for the organizational reporting and performance evaluation system widely in use today. Kaplan argued that all accounting currently in use had already been developed by 1925.

Although Kaplan (1984) acknowledged the advances in management control enabled by the use of ROI and the profit center structure, he also warned of the problems associated with a short-term performance perspective. Kaplan commented that rather than using accounting practices for internal planning and control, savvy managers learned ways to game the system by taking advantage of the interpretation of external accounting rules. Mergers, acquisitions, divestitures, debit repurchases, and other such transactions could increase short-term earnings but didn't necessarily generate long-term economic value for the firm's stakeholders. Kaplan's opinion was that the most damaging behavior by managers was the myopic decision to reduce discretionary spending that supported intangibles such as product development, quality improvement, and human resource development in order to meet short-term financial objectives; in essence, mortgaging the future of the firm.

The quality movement stage. The third stage of PMS evolution was influenced by the beginning of globalization and the dawning of the quality movement. Bititci et al. (2012) stated that leaders from resource wealthy Western countries emerged from World War II with a renewed interest in maximizing throughput and output from their robust manufacturing base. In addition to reliance on financial metrics, managers and engineers

focused on productivity improvements. The result was mass production of a limited product assortment leading to large inventories of low cost goods. Bititci et al. pointed out that customer and employee satisfaction were often sacrificed in the name of productivity. Coe and Letza (2014) stated that performance measurement was heavily skewed toward revenue, efficiency, and ROI; measures and metrics with a singular backward looking focus on financial outcomes.

Sato (2016) noted that while the Western world was enjoying an era of plenty, the Japanese were attempting to re-build after the Second World War and were facing a plethora of quality issues. Sato stated that American scientist, W. Edwards Deming, introduced the Japanese to the use of statistics-based quality control methods. Deming believed that a focus on producing quality products would result in improvements in both cost (less waste) and cycle time. Deming was adamant that processes were to be designed with the customers' needs first and foremost. Through the influence of Deming, forward looking metrics related to quality and customer satisfaction joined financial measures in many Japanese companies providing a more balanced perspective of organizational health. Sato added that the Japanese companies that were part of the quality revolution enjoyed a competitive advantage in the growing global marketplace.

Dahlgaard-Park, Chen, Jang, and Dahlgaard (2013) shared that as the latter half of the 20th century unfolded, leaders of American businesses found that the productivity growth in the US had declined while Europe and Japan were enjoying increasing competitiveness. Dahlgaard-Park et al. stated that the situation caused leaders in Western-based companies to explore, adopt, and adapt quality techniques within their own

organizations to reduce costs and increase customer satisfaction. Dahlgaard-Park et al. found that as the quality movement emigrated from Japan there was also an evolution from purely a statistical tool to a broader management philosophy called Total Quality Management that incorporated elements of process management, human resource management, and strategic management. Dahlgaard-Park et al. went so far as to call this a management revolution.

Although the quality movement exposed managers to new means of measuring performance, such as quality, time, cost, and flexibility, Coe and Letza (2014) pointed out that most companies still focused mainly on financial metrics. Coe and Letza argued that this unbalanced focus had undermined the competitiveness of manufacturing companies in the United States. Coe and Letza added that academics criticized the historical, short-term, nonstrategic, and internal perspective of focusing solely on financial measures. Although ROI and profit center accounting assisted Andrew Carnegie in managing costs in his steel company over a hundred years ago, Hoque (2014) stated that in the 1980s researchers were concerned with the focus only on traditional financial measures. Hoque called for research on other performance measurement frameworks that could provide feedback and support to business leaders experiencing a climate of high uncertainty, excess capacity, and increased global competition. To summarize, leaders required performance measures that were balanced across financial and nonfinancial aspects of the business, measured both tangible and intangible assets, included leading and lagging indicators, and had a strategic focus rather than merely an operational and functional perspective.

Integrated PM frameworks and models stage. Taticchi, Balachandran, and Tonelli (2012) stated that over the past 20 years there has been a dramatic increase of interest in performance measurement. Bititci et al. (2012) called this the period of integrated performance measurement. This stage was characterized by the emergence of many PMS models and frameworks that relied on a mixture of financial and nonfinancial measures. Estampe et al. (2013) found this period of the performance measurement journey notable stating that the objectives of such models included helping management measure business performance, conduct analyses, and increase the efficiency of the firm by having information that improved decision-making.

According to Peng and Prybutok (2015), the US Congress, in an effort to improve the quality of US businesses, established the Malcolm Baldridge National Quality Award in 1987. Taticchi et al. (2012) stated the other models soon followed. These PMS models incorporated ideas such as the use of nonfinancial indicators, building PMSs from the view of the customer, and the introduction of the concept of balanced measures.

Estampe et al. (2013) listed 16 of the most well-known supply chain performance measurement models. Some examples were the ABC (Activity-based Costing) model, the FLR (Framework for Logistics Research) model, the BSC (Balanced Scorecard), the WCL (World Class Logistics) model, the ECR (Efficient Customer Response) model, the EFQM Excellence model, and the SPM (Strategic Profit Model). Estampe et al. included in their literature review (a) references from the literature, (b) model origin, (c) type of analyses included in the model, (d) conditions and constraints of each model, (e) degree of conceptualization, and (f) established indicators. The authors acknowledged a gap in

the literature with respect to evaluating the usefulness of these models beyond the specific industries for which they were developed. Estampe et al. created a framework for evaluating the application of these models more broadly in the supply chain.

Coe and Letza (2014) stated that the BSC has been touted as one of the top management ideas of the previous century. Hoque (2014) conducted a literature review in honor of the 20th anniversary of Kaplan and Norton's inaugural article on the BSC that appeared in a 1992 issue of the Harvard Business Review. Hoque stated that although other authors had previously provided balanced frameworks, in 1992, Kaplan and Norton introduced the Balanced Business Scorecard. This scorecard provided, in a single performance measurement tool, both the financial and nonfinancial perspectives of an organization. The four perspectives in the scorecard were *financial*, *customer*, *internal business*, and *innovation and learning* views of performance.

Hoque (204) researched the evolution of the Balanced Business Scorecard. He listed several changes including shortening of the name to the Balanced Scorecard, changing the internal business perspective to internal business processes and modifying the innovation and learning viewpoint, to learning and growth as a result of further field study by Kaplan and Norton. Other developments included linking each perspective to create a feed-forward control system with measures of learning and growth affecting internal business processes measures thereby affecting customer measures and ultimately leading to changes in financial measures. Bhattacharya et al. (2013) recommended that a sustainability perspective, for those organizations in which sustainability is a key business strategy, be added to the BSC. Kaplan (2012) stated that although he preferred

the shareholder to the stakeholder value approach, he did provide case study research on incorporating the triple bottom line view into the BSC framework.

Kaplan (2012) commented that although the original 1992 article on the BSC was focused on performance measurement, it soon became apparent to both Kaplan and Norton that the BSC could act as a framework for both strategy development and execution. Strategy execution subsequently became their research focus. They were adamant that the starting point for development of the BSC was formulation of the vital few strategies that would lead the organization forward. Kaplan explained that these strategies provided the framework for developing the objectives, measures, targets, and initiatives to deliver on each strategy. The result was called a strategy map.

Cheng and Humphreys (2012) stated that the BSC framework provided for double-loop learning through a management control loop to track progress and a strategic learning loop that made strategy refinement part of the continuous improvement process. Coe and Letza (2014) concurred, adding that the BSC enabled a feedback and learning process that allowed all stakeholders to understand their performance in meeting organizational objectives. The BSC also provided information to managers to determine if the plans that had been put in place were delivering the expected results. The strategies and/or plans could be modified as part of a course correction process. Kaplan (2012) summed up the strategic management system process by stating that the strategy map and the BSC were the vehicles senior managers should use to communicate the desired organizational outcomes to middle and lower level employees. This in turn would enable those managers and employees to come up with new ways to conduct business to drive

organization performance. Strategy formulation was intended to be top down while strategy implementation was to be bottoms-up.

Hoque (2014) found a burgeoning stream of literature indicating that higher BSC usage was associated with improved financial performance for the firm. Wu (2012) stated that the BSC appeared to be an adequate model for enabling organizational leaders to reach their goals through the prioritization of actions. Nørreklit, Nørreklit, Falconer, and Bjørnenak (2012) were more critical of the BSC, stating that sound logic for the four perspectives was missing and that research showing causality among these perspectives was inconclusive. Nørreklit et al. recommended that future research establish cause and effect relationships between measures and results.

Okongwu, Brulhart, and Moncef (2015) studied the BSC from a supply chain management perspective. Okongwu et al. supported other research showing that the financial performance of the firm was impacted directly by the second level components of the BSC and indirectly by third level components, demonstrating the interdependencies between the intangible and tangible assets of the firm. Hoque (2014) found that researchers have paid little attention to more current developments of the BSC in relation to strategy maps, strategic alignment, and execution of strategy. This finding was consistent with the rebuttal to critics from Kaplan (2012) who was surprised that most BSC research focused on the performance measurement aspect and very little on the 16 years' worth of work on strategy execution. Kaplan challenged academics to get out there and perform original research. Hoque (2014) concurred, recommending additional

study on the relationship between the BSC and organizational performance including field studies in real-life settings.

Hoque (2014) stated that the BSC enabled managers to continue to monitor financial results while simultaneously focusing on building future capabilities that would enable their organizations to compete and grow in a rapidly changing global environment. Hoque commented that since the BSC retained traditional financial measures while adding the perspectives of customer, internal business processes, and organizational learning and growth, the BSC should cause a shift in management focus from tangible assets to a proactive view of capitalizing on intangible assets with the goal of future value creation. Hoque summarized the current state of thinking about the BSC stating that the BSC provided a solid framework for strategy development, policy enablement, and as a mechanism for organizational accountability and control.

PMS development and implementation stage. Franco-Santos et al. (2012) observed that by the end of the 20th century there was a veritable cornucopia of PMS frameworks to aid managers in determining what to measure. Gutierrez et al. (2015) agreed, adding that most models focused on structure rather than on guidance related to developing or updating PMSs. There was a paucity of research that could assist managers in implementing PMSs within their respective organizations. The next stage of PMS research evolved from multidimensional model development to how such frameworks should be implemented. Franco-Santos et al. added that recent research has been focused on real-life studies.

Design, implement, use, and refresh phases of PMSs. Bourne, Mills, Wilcox,

Neely, and Platts (2000) proposed that PMS implementations should consist of the

general phases; *design*, *implement*, *use,* and *refresh*. Bourne et al. recognized that there is

blurring of the activities that occur in each phase. Agostino and Arnaboldi (2012)

acknowledged that both the design and use phases of PMSs have been widely researched.

Gutierrez et al. (2015) found a greater research focus on PMS design rather than on

implementing, using, or assessing PMSs. Searcy (2012) reviewed the literature on PMS

implementations specifically as it related to the addition of sustainability measures and

identified three key phases; design, implementation and use, and evolution. Hoque (2014)

conducted a literature review on BSC-related topics and found that a large body of

literature focused on the design, implementation, and use of the BSC. Gutierrez et al.

synthesized a framework for the design, implementation, use, and assessment of PMSs

based on prior research.

Design phase. Gutierrez et al. (2015) stated that the design phase was further

divided into sub-phases that included (a) deciding on the business objectives that will be

the subject of measurement, (b) design of the measures, and (c) development of the

metrics framework for final review and alignment. Objectives were to be established by

considering not only the requirements of the customer but also the needs of the

employees, shareholders, vendors, and other stakeholders in the community within which

the firm operates. Strategies were to be developed with the intent of changing stakeholder

behavior. Finally the metrics were to be designed to enable managers to monitor progress

and overall performance. Estampe et al. (2013) presented a list of frameworks to aid in the design of a PMS.

Gutierrez et al. (2015) suggested that business leaders understand and incorporate into PMS design the answers to questions such as:

- "What are our objectives?"

- "What should we measure?"

- "What performance measures shall we use and what attributes are important?"

- "What is the process to review the performance measures?"

Chalmeta, Palomero, and Matilla (2012) stated that a set of key performance indicators (KPIs) linked to strategy should be at the heart of the design of the PMS. The cause and effect relationship between each KPI should be established. Cause indicators should predict future value for the organization. Effect indicators are the actual results. Chan et al. (2014) corroborated this approach stating that there are two types of performance measures, those that measure results (e.g., sales, customer satisfaction, etc.) and those that are determinants of results (e.g., innovation, quality, etc.).

Although Franco-Santos et al. (2012) stated that there is a substantial body of literature on designing PMSs, Chan et al. (2014) unearthed very little literature on PMS design and performance measurement selection in the supply chain arena. Chan et al. found that due to the complex nature of the supply chain, decision-makers suffer from information overload. After a comprehensive review of the literature on designing performance measures, metrics, and indicators, I have found few studies that focus on selecting performance measures and none that explain the actual decision-making process

leaders use to decide on the critical few measures and goals that they will use to improve organizational performance.

Implementation phase. According to Gutierrez et al. (2015), the implementation phase is categorized as the data collection phase. This included the methods (manual or computer system-enabled) as well as the processes and procedures for the gathering of data, converting that data into information that is usable by management, and the dissemination of the resulting output. Chalmeta et al. (2012) stated that a key task of the implementation phase is the development of a communication plan designed to make employees aware of and knowledgeable about an upcoming PMS implementation. Chalmeta et al. added that employee engagement was key to the success of the PMS implementation. Additionally, Chalmeta et al. stressed the importance for leaders to communicate why the company was undertaking a PMS implementation. And Chalmeta et al. stated it was also imperative to train managers and employees on the meaning and importance of each indicator as well as how the indicator was to be used to drive the desired behavior change in the organization.

Use phase. Henri (2006) stated that there are four classifications for the use of PMSs: (a) monitoring, (b) focusing attention, (c) strategic decision-making, and (d) legitimization. Monitoring lets managers know how the organization is doing in achieving goals. Results were compared with targets and if necessary adjustments were made to get back on track. PMSs could also be used to detect problems thereby focusing the attention of the organization on issues and their resolution as well as creating alignment as to the direction the organization is heading. Henri added that senior leaders

relied on PMSs for strategic decision-making by using the information from the PMS to determine the best course of action from a variety of possible choices. Henri said PMSs might also reveal cause-and-effect relationships between internal actions and results of those actions. Henri stated that use of PMSs in strategic decision-making supports organizational learning and problem solving. PMSs can also be used to legitimize prior decisions made or actions taken. Linking such decisions and actions to actual outputs influences the credibility and authority of leaders with their stakeholders.

Koufteros et al. (2014) said that although it logically follows that firms with well-developed PMS systems should outperform those with under-developed or no PMS system, the literature showed mixed results related to PMS usage and firm performance. Koufteros et al. also noted a gap in the literature regarding how companies actually use PMSs to harness their resources and whether such uses do lead to improved operational and strategic capabilities over time. Koufteros et al. stated that managers use PMSs for *diagnostic* purposes and/or for *interactive* purposes. Diagnostic use including actions such as monitoring, focusing attention, and legitimizing can be viewed as the mechanistic use for PMSs. The authors referred to interactive use as a strategic management tool that (a) builds organizational capabilities through organizational dialogue, thereby stimulating the development of new ideas and new actions; (b) encourages discussion among various management levels; (c) results in continual debate over the results, assumptions made, and previous actions taken; and (d) promotes engagement and interest throughout the organization. And Koufteros et al. summarized that interactive use of PMSs expands

organizational learning and opportunity-seeking behaviors that can be a catalyst for organizational renewal.

Koufteros et al. (2014) added that applying both dynamic and interactive uses in a balanced way leads to dynamic tension. Such tension balances taking actions consistent with organizational goals while simultaneously giving employees sufficient freedom to make decisions. Additionally, Koufteros et al. demonstrated empirically that dynamic tension created unique organizational capabilities giving firms a competitive advantage. Firms that did not use PMS systems in a balanced way were slower at decision-making, had less efficient use of resources, and lower and less reliable performance over time. Finally, Koufteros et al. concluded that organizations should deploy PMSs both for diagnostic and interactive uses to maximize operational and strategic management capabilities.

Refresh phase. Gutierrez et al. (2015) stressed the importance of including a reflection and refresh process to regularly evaluate if the measures and goals were driving the appropriate behavior change of all stakeholders and to monitor that over time the strategies and measures did not become misaligned. Chalmeta et al. (2012) stated that the PMS must be evaluated on a regular schedule so that the PMS can be optimized. Gutierrez et al. added that while new measures are regularly added, old measures often aren't retired leading to unnecessary PMS complexity. Gutierrez et al. concluded that an effective PMS system must include a refresh process for reviewing metrics, measures, indicators, and goals as business conditions change.

Bititci et al. (2012) stated that performance measurement has evolved in response to a more complex and volatile global business environment. Gutierrez et al. (2015) argued that if an organization is to survive in a rapidly changing environment, leaders must continuously assess their goals in order to operate in a more effective and efficient manner. Such an environment requires more rapid response to changes thereby putting increased pressure to update the targets for existing measures, to remove those measures that no longer bring value, and to replace them with new measures. Bititci et al. questioned how the shift to a growing reliance on knowledge workers and a service-based economy could impact PMS design, implementation, and use. Then Bititci et al. summarized the research challenges stating that scholars need to understand performance measurement as both a social system and a learning system.

Barriers to success with PMSs. Gutierrez et al. (2015) stated that problems encountered during the various phases of the PMS lifecycle could impede the usefulness of such a system. Gutierrez et al. stated that issues can be categorized as cultural, systems, and/or process-related concerns. Bititci et al. (2012) found that a command-and-control culture may lead to dysfunctional behaviors that demotivate employees ultimately resulting in hidden costs. Gutierrez et al. (2015) added that fear, as a result of increased transparency of performance, along with associated subversive behaviors could undermine a PMS implementation. They added that top management commitment and a focus on creating a learning environment could help alleviate those effects. Several authors (Bititci et al., 2012; Chalmeta et al., 2012; Koufteros et al., 2014) encouraged managers to apply good change management practices to PMS implementations

including employee involvement, establishment of common objectives, training, a focus on learning, and positive reinforcement by managers to create an open, nonthreatening, and performance driven culture. Bititci et al. (2012) referred to performance measurement and management as a social process that is influenced by the feelings and belief systems of all affected stakeholders.

Chalmeta et al. (2012) found that one of the top barriers to the success of a PMS implementation was due to difficulties in gathering complete, accurate, and pertinent data from information systems. Chalmeta et al. found a direct relationship between data collection automation and use of a metric. Gutierrez et al. (2015) agreed, adding that lack of data quality and reliability as well as fragmented system infrastructure could undermine PMS efforts. Gutierrez et al. advised including IT early in the PMS design process thereby allowing enough time for system and process development.

Gutierrez et al. (2015) conducted an action research project to re-implement a PMS for the logistics department in the supply chain function of a large broadcasting company. The research team began the project at the refresh phase by assessing the current PMS system and using gaps in the assessment as input into the design phase. Additionally, Gutierrez et al. considered the cultural, systems, and process barriers identified in the literature as they led the project from design to implementation, use, and renewal. Gutierrez et al. and their client team encountered many issues along the way. Ultimately Gutierrez et al. corroborated the findings of Koufteros et al. (2014) that by balancing both the diagnostic and interactive aspects of a PMS, the logistics department was able to demonstrate continuously improving performance in service and cost. And

Gutierrez et al. implied that such supply chain improvements helped overall firm performance.

Contemporary PMSs stage. According to Bititci et al. (2012), a performance measurement revolution is currently underway. Both scholars and practitioners have recognized that traditional PMS frameworks are no longer adequate. Franco-Santos et al. (2012) stated that there isn't a common definition of a contemporary performance measurement system. They recommended focusing on the minimum necessary and sufficient conditions in defining contemporary PMSs. A contemporary PMS must contain both financial and nonfinancial measures that are used to translate strategy to operational activities. Contemporary PMSs must provide the framework for performance evaluation. A contemporary PMS must have a supporting infrastructure. Lastly, a contemporary PMS must contain a specific set of processes that incorporate the design of measures, the gathering of data, the resulting performance review, and development of action plans.

Linkages Among Performance Measurement Systems, Behavior, and Performance

Franco-Santos et al. (2012) conducted a literature review related to the consequences of implementing PMSs. Franco-Santos et al. developed a conceptual framework that linked PMSs to consequences for (a) individual behavior, (b) organizational capabilities, and (c) organizational performance. In the subsequent sections, I dive more deeply into the research on each of these relationships.

Linkage between PMSs and individual behavior. Franco-Santos et al. (2012) made salient several ways in which PMSs affected people's behavior including (a) strategic focus; (b) cooperation, coordination, and participation; (c) motivation; (d)

understanding of and satisfaction with job role; (e) leadership and culture; and (f) subjectivity, justice, and trust. In the category of strategic focus, Franco-Santos et al. found that PMSs helped increase executive focus on the most important strategies for the organization. PMSs use also increased executive dialogue about strategy. Francos-Santos et al. found that although the authors of several qualitative studies argued that strategic focus was a key driver in impacting performance, there is a lack of research to measure strategic focus empirically, noting this as an area of future research.

Franco-Santos et al. (2012) found evidence in the literature that PMSs facilitated cooperation, coordination, and participation both inside and between members of organizations. And Franco-Santos et al. stated that PMSs enabled cooperation in supply chain relationships that enhanced the buyer-supplier dyad that in turn, led to increases in learning and problem solving behaviors. Additionally, Franco-Santos et al. also noted studies that found increased cooperation among employees in different units of an organization when performance measures were aligned versus the more commonly found situation of misaligned and competing measures among divisions. Franco-Santos et al. observed that the recursive and inclusive processes required for PMS implementation created an environment that increased employee involvement, cooperation, and participation.

Franco-Santos et al. (2012) found that although there was a consensus in the research on the impact of employee involvement both in the performance measurement and performance management processes, the literature showed mixed results when it came to motivation effects. Franco-Santos et al. also found a positive motivational effect

when employees participated in the performance measurement process and a negative

effect when performance measurement was tied to bonus payments. Additionally,

Franco-Santos et al. summarized the research stating that motivational effects are as

much about the process of developing and using PMSs as they are about the actual

results. Marginson et al. (2014) found that interactive use of PMSs positively affected

managers' thinking and motivation by influencing managers' psychological

empowerment and reducing role ambiguity. This research supported speculation by

Franco-Santos et al. (2012) that feedback mechanisms inherent in the use of PMSs may

aid managers in improving their own performance that in turn could help managers

develop their individual mental models. Franco-Santos et al. recommended that leaders

maximize (a) managers' psychological empowerment, (b) employees' participation in the

process, and (c) everyone's commitment to the goals. And Franco-Santos et al. added that

the motivational effects of bonus payments tied to performance measurement on

individual performance remains an unresolved issue in the literature.

Franco-Santos et al. (2012) also found that adoption of PMSs facilitated the

individual's role understanding and job satisfaction. Although the owners and designers

of PMSs are generally satisfied, Franco-Santos et al. posited that this effect might be due

to ownership bias. Gutierrez et al. (2015) commented that if individuals are given many

competing goals, the situation might tax the employee's cognitive limits leading to

focusing on one goal at the expense of other goals. Waal and Kourtit (2013) added that

setting goals and budgets allowed individuals to clarify what is expected of them. Franco-

Santos et al. also found that use of performance measurement increased job satisfaction if

individuals trusted their supervisors and believed that the performance evaluation process was transparent and fair.

According to Franco-Santos et al. (2012), PMSs could act as catalysts for new ways of leading and working. The resulting cultural change might include an increase in participative and consultative management styles along with higher frequency and quality of manager-employee dialogue. Over time new cultural routines, values, and beliefs could evolve and become established. Franco-Santos et al. also asserted that although PMSs influence culture, likewise culture can become boundary setting thereby moderating the effects of PMSs on individual behavior.

Franco-Santos et al. (2012) found that the implementation and use of nonfinancial measures inherent to PMSs have shown mixed results related to perceptions of justice and trust. Some studies uncovered that nonfinancial measures might be perceived as too subjective leading to concerns about the validity of the PMS. Marginson et al. (2014) shared that other researchers found that subjective measures could be positive because such measures allowed organizations to be adaptable and flexible. Franco-Santos et al. summarized the situation by stating that nonfinancial measures could be perceived as positive or negative depending on the characteristics of the organization. They called for more research in this area. Although research results were mixed relative to some of the effects of PMSs on individual behavior, generally it is acknowledged in the literature that the design and use of PMSs have a strong and lasting impact on individual behavior. It logically follows from the literature that selecting the right PMS could contribute to driving the desired individual behavior.

Linkage between PMSs and organizational capabilities. Franco-Santos et al. (2012) highlighted several ways in which PMSs impact organizational capabilities: (a) impact on strategy process, (b) effect on communication processes, (c) building strategic capabilities, (d) influence on management processes, and (e) effects on corporate control. PMSs impact strategy processes through the engagement of managers in the development and review of strategic plans, the translation of those plans into operational language, and by changing the manager's mindset to one in which strategy is seen as a continuous process rather than a series of one-time events. Although in general the research showed a positive influence of PMSs on the strategy process, Franco-Santos et al. found that the design of the PMS can have a boundary-setting effect based on the culture of the organization and the cognitive limitations of the managers.

Koufteros et al. (2014) demonstrated empirically that diagnostic use of PMSs (monitoring, focusing attention, and legitimization) positively impacted organizational capabilities as measured by strategic management, operational, and external stakeholder-relations capabilities. This was in direct contradiction of earlier work by Henri (2006) who found that diagnostic use of PMSs could lead to dysfunctional behaviors thus acting as a negative factor. Koufteros et al. found that the interactive use of PMSs encouraged dialogue and stimulated new ideas and actions that favorably impacted strategic management, operational, and external stakeholder-relations capabilities.

Koufteros et al. (2014) stated that the greatest increase in organizational capability was found when organizations practiced strong dynamic tension, described as balancing the complementary levers of diagnostic and interactive use of PMSs. Organizations with

high dynamic tension in the use of their PMSs were able to make decisions more quickly and had better resource utilization leading to unique organizational capabilities. This result is consistent with research by Marginson et al. (2014) who found that both diagnostic and interactive uses of nonfinancial measures reduced role ambiguity while interactive use increased psychological empowerment in the organization.

Linkage between PMSs and organizational performance. Franco-Santos et al. (2012) found that although the general consensus in the literature is that PMS use improves organizational performance, other studies show little to no relationship between PMS use and performance. And Franco-Santos et al. discovered that for studies that used reported performance (typically via annual reports) some studies report a positive relationship between the existence of PMSs and organizational performance, others show little to no relationship, and still other studies show mixed results. For studies in which managers self-reported perceptions of organizational performance, there is also mixed results. Results from quantitative studies generally show positive results between PMS implementation and organizational performance whereas results from qualitative studies show mixed results.

Waal and Kourtit (2013) observed that although there is evidence that PMSs have been implemented in about 70 percent of mid- to large-sized firms in the US and Europe, Koufteros et al. (2014) found that prior research has shown mixed results related to the use of PMSs and organizational performance. It seems that implementation of PMSs does not automatically improve organizational performance.

Saunila, Pekkola, and Ukko (2014) found that performance measurement moderated the relationship between innovation capability and organizational performance in small- and medium-sized firms. Franco-Santos et al. (2012) stated that research has shown that the design, development, and use of PMSs are keys to organizational performance improvement. Some of the moderating factors Franco-Santos et al. cited included (a) strategic orientation, (b) organizational structure, (c) culture, (d) management style, (e) quality of information systems, and (f) external factors such as competitive landscape and environmental uncertainty. Organizational size, market position, and product life cycle were not found to influence PMS use or effectiveness. Marginson et al. (2014) stated that psychological empowerment mediates the relationship between PMSs and organizational performance. Abushaiba and Zainuddin (2012) proposed a framework in which PMS design indirectly impacts firm performance by affecting the firm's competitive advantage.

Franco-Santos et al. (2012) stated that there is a limited amount of research on the use of PMSs and team performance. The research indicated a positive correlation when team members are included in setting performance goals and when teamwork is encouraged via the inclusion of team specific measures and goals in the performance management and bonus payment processes. Franco-Santos et al. found that team performance improvement appeared to be linked to the design and use of PMSs but called for more research on this topic.

Although research suggested that PMS use affects individual behavior, building of organizational capabilities, and organizational performance, it appears that the extent

of performance improvement is highly dependent upon how the PMSs were designed, implemented, and used along with the internal and external context for use (Franco-Santos et al., 2012; Koufteros et al., 2014; Marginson et al., 2014). PMSs are costly to implement and use from a resource perspective and may foster impressions of unfairness and bias within an organization (Franco-Santos et al., 2012; Gutierrez et al., 2015). Simply developing a PMS may not lead to organizational improvement (Waal & Kourtit, 2013). It logically follows that research that could improve the effectiveness of the design, implementation, use, and renewal of PMSs could be of value to both scholars and practitioners.

This literature review detailed the evolution of PMSs with highlights on the current and future challenges in designing PMSs that are relevant and useful to managers. Melnyk et al. (2014) summarized the current state of literature on PMS by declaring that despite the vast amount of research and development related to performance measurement, business leaders still do not seem to have the tools they need to effectively steer their organizations. One major gap in the research I found that may help to better understand the situation is an absence in the literature of studies on how supply chain leaders sift through the myriad of information to determine the performance measures that matter. My study took a new approach to studying PMSs by exploring how supply chain leaders actually chose the measures and goals for their respective supply chains through the lens of decision-making theory.

Overview of Decision-Making Research

The phenomenon in the study was the decision-making strategies supply chain leaders use to select the key performance measures and goals they used to measure the performance of their respective organizations. Gigerenzer (2015) stated that errors in decision-making have had catastrophic impacts on organizations, citing the financial crisis of 2008 as an example. Franklin (2013) commented that expertise in decision-making is a core competency of an effective manager.

Gigerenzer (2015) said that decisions are made using rational or nonrational strategies. Rational strategies rely on the use of logic and/or statistics. Nonrational strategies could include the use of heuristics, biases, expertise, and intuition. There has been debate among scholars on the strategies that should be used and are used in making decisions.

Gonzalez, Meyer, Klein, Yates, and Roth (2013) stated that a focus on decision-making research should be on decision behavior that is exceptionally challenging and can have major repercussions if not implemented well. I suggested that one approach to adding to the body of knowledge on performance measurement was by exploring the decision-making strategies actually used by supply chain leaders in selecting their performance measures because of the potential impact of PMS selection on organizational performance. In the following sections of Chapter 2, I review rational and nonrational decision-making approaches and explain how these approaches apply to the research.

Rational decision-making. According to Glimcher (2015), early research on decision-making was rooted in mathematical and economic theory. These early explorers were seeking to understand why people made the choices they did with the goal of building models that would predict how rational decisions should be made. Szpiro (2013) wrote an article about the exchange of letters in the 18th century between the Swiss mathematician Nikolaus Bernoulli and Pierre Remond de Montmort, a French nobleman and fellow mathematician, about a relatively simple gamble that came to be called the St. Petersburg paradox. After much dialogue, Bernouilli realized that the value of a choice in gambles is not based merely on its price, but also on the satisfaction the choice provides. Szpiro postulated that this discussion that took place over 300 years ago might have been the genesis of rational decision-making theory.

Logical decision-making. Franklin (2013) stated that the general approach to rational, often called classical decision-making involves four steps. These steps included (a) describing a problem or decision to be made, (b) generating an exhaustive list of possible solutions/outcomes, (c) assessing the value of each option, and (d) selecting the best choice based on the highest value. Franklin said that rational decision-making is process oriented and the process does not change with the type of decision being made.

Although rational decision-making provides a formal structure, Franklin (2013) observed that such rigidity has drawbacks including that the process (a) is linear without feedback loops, (b) mandates all alternatives be identified, (c) requires the decision-maker to understand the linkage between alternatives and outcomes, (d) requires perfect information to develop the likelihood of each outcome, and (e) assumes a single unified

purpose for the decision. Shaban (2012) said that these normative theories of decision-making took a positivist perspective in evaluating how good decisions should be made. Such theories did not consider how decisions were actually made in real-world conditions.

Expected utility theory. Szpiro (2013) stated that Bernouilli's theory went largely ignored until the mid-20th century when mathematician John von Neumann and economist Oskar Morgenstern developed a mathematical framework, expected utility theory (EUT), that measured the concept of utility or value to the decision-maker. Ramos et al. (2014) stated that a foundational assumption of EUT is that people are utility maximizers. The axioms upon which EUT was founded include *completeness, transivity, continuity,* and *independence.* These axioms paved the way to mathematically calculate the rational decision-maker's preference among various prospects or outcomes.

 According to Ramos et al. (2014), the axioms of completeness and transivity together provided the basis for determining the ordering of preferences namely

- completeness*:* for all outcomes x and y; x > y, or y > x, or x = y and
- transivity*:* for all outcomes x, y, and z; if $x \geq y$ and $y \geq z$, then $x \geq z$.

Completeness implies that there is a clear preference among prospects while transivity provides a direction for the preference.

Ramos et al. (2014) stated that the continuity axiom insured that an intermediate outcome exists that has the same utility to a decision-maker as a set of outcomes with a specific likelihood of occurrence.

- Continuity: for all outcomes x, y, and z, if $x < y < z$, then there is a probability

  (p) between 0 and 1 such that the $p{\cdot}x + (1 - p){\cdot}z$ has the same value to the

  decision-maker as outcome y alone.

Taken together the first three axioms implied that mathematically there exists a value

function V that describes the decision-maker's preferences.

Ramos et al. (2014) added that the independence axiom indicates that the

preference for the selection of outcomes is independent of the common components of

the outcome. If outcome x is preferred over y, when a third outcome is introduced, the

decision-maker would still prefer x over y.

- Independence: for all outcomes x, y, and z; if $x > y$, then $p{\cdot}x + (1 - p){\cdot}z > p{\cdot}y +$

  $(1 - p){\cdot}z$ for all z and p is an element of the set (0,1).

Ramos et al. (2014) add that the independence axiom is the one most often challenged by

researchers.

Ramos et al. stated that collectively the axioms above allow for the development

of a formula to calculate expected utility by multiplying the value of each outcome with

its probability of occurring and then summing the results. The rational decision-maker

will prefer outcome x over outcome y if the expected utility for outcome x is greater than

the expected utility value for y. Ramos et al. added that the EUT model can be used

prescriptively by asking a person to make a determination about utilities and the

associated probabilities. By using the EUT formula it should be possible to predict the

person's decision. This weighting and adding process is a key element of utility-based

decision-making models.

According to Ramos et al. (2014), economists used EUT not only as a model of rational decision-making but also used EUT as a way to describe how individuals make decisions under conditions of risk. The presumption was that individuals are rational in nature and make decisions in logical and systematic ways. Fennell and Baddeley (2012) shared that normative EUT models have been used to explain why people gamble and why they buy insurance policies. Buchholz and Schymura (2012) stated that EUT has been applied not only to individual choices but also to societal decisions of catastrophic risks with high cost but low probability like pandemics or the earth colliding with a giant asteroid.

Kahneman (2011) proclaimed that EUT is arguably the most important social science theory by describing, through the use of axiomatic logic, how decisions should be made. Fennell and Baddeley (2012) asserted that it is the axiomatic nature of utility theory that has allowed the model to have advantages over other less rigorously developed models. Fennell and Baddeley explained that the economic theory- based EUT model has been useful as the foundation for other rational decision-making models.

Bayesian theory. According to Fennel and Baddeley (2012), many models have been developed to explain what appears to be other than rational behavior. Some of these models are based on Bayesian theory. Bowers and Davis (2012) stated that Bayes' theorem explains how a person should update the probability of an outcome by incorporating new information. An example from Bowers and Davis was the use of Bayesian theory to determine the likelihood that a patient had lung cancer. Based on prior research, the probability that a 30-year old man has lung cancer is 0.5%; a very low

probability. Assume this patient is given a blood test for lung cancer. This test has a 99%

accuracy rate when cancer exists and a 1% false-alarm rate. Entering these probabilities

into Bayes' rule resulted in a 33% probability that if the 30-year old man tested positive,

that he actually has cancer. Bowers and Davis admitted that this result seems

counterintuitive. Bowers and Davis added that it is the low prior probability of cancer in

someone so young that heavily influences the posterior probability. If the patient had

been a 90 year-old smoker, the prior probability would rise to 33% and the posterior

probability would increase to a 98% likelihood of cancer with a positive test result.

Shaban (2012) explained that Bayes' Theorem is used to account for the impact of

uncertainty and stress in decision-making. He added that failure in decision-making is

generally caused by (a) a lack of decision-making skills, (b) failure to apply decision-

making rules, and (c) lack of knowledge or information about the likelihood of an

outcome occurring. Bowers and Davis (2012) cited studies demonstrating that doctors

performed poorly in determining posterior probabilities given large variances in prior

probabilities thereby supporting Shaban's perspective on decision-making failures.

Fennell and Baddeley (2012) stated that the rational approach would be to

overweight low prior probabilities and underweight high prior probabilities. Fennell and

Baddeley argued that if the prior probability had been encountered in the past then the

knowledge gained from the prior encounter should be used. If the situation is new or

markedly different from prior encounters then a *prior of ignorance* should be used.

Additionally, Fennel and Baddeley acknowledged that it could be difficult to ascertain if

a current situation is sufficiently like a previous situation to apply the prior probability.

Fennell and Baddeley listed four conditions that could erode the certainty of a situation: (a) the environment is continually changing, (b) probabilities are based on finite levels of experience, (c) memories are flawed, and (d) people aren't always truthful. And Fennel and Baddeley warned that improper application of priors could lead to very poor decisions that might yield disastrous results.

Many other rational decision-making models. Chai, Liu, and Ngai (2013) conducted a literature review of articles published between 2008-2012 on the use of decision-making techniques applied to a complex supply chain challenge, that of supplier selection. Chai et al. came up with three groupings for the models: (a) multi-attribute decision-making techniques, (b) mathematical programming techniques, and (c) artificial intelligence. They found that the majority of researchers used multiple techniques. Chai et al. also found increased usage of models based on *fuzzy set theory* because of the difficulty in applying clear black and white decision criteria.

Criticisms of rational decision-making models. Fennell and Baddeley (2012) pointed out that the classic EUT model assumes no uncertainty in the probability factor and that such an assumption is naïve at best. Fennell and Baddeley added that rational decision-making models require complete knowledge of all possible alternatives and their probabilities along with a predictable and stable environment; conditions that are impractical in the real-world. Franklin (2013) added that von Neumann based this theory around games of chance in which both gains and probabilities were well known. Franklin stated that for many decisions faced by leaders, both value and probability are often difficult if not impossible to determine with any certainty. Gigerenzer (2014) stated that

applying the small-world assumptions used in the EUT model to real larger worlds could lead to disastrous results.

Fennell and Baddeley (2012) acknowledged that researchers have observed that actual behavior often deviates from what one would expect from applying the axioms of EUT to situations. Almy and Krueger (2013) stated that one goal of determining utility is to understand what is psychologically important to each decision-maker. Almy and Krueger encouraged individuals to consider von Neumann and Morgenstern's axioms as part of the decision-making process.

Nonrational decision-making. According to Spaniel (2014), the independence assumption of the EUT model was called into question with an experiment conducted in the 1950s by a French mathematician named Allais in which he first offered respondents the choice between two gambles; Lottery A with an 11% chance of winning $1million and an 89% chance of winning nothing, or Lottery B with a 10% chance of winning $5 million and a 90% chance of winning nothing. Allais then offered the same respondents another gamble; Lottery C with a 100% payout of $1 million, or Lottery D with a 10% chance of winning $5 million, an 89% chance of winning $1 million, and a 1% chance of winning nothing. Spaniel explained that if one were to follow the rules of rational decision-making, the respondent would select Lotteries A and C or B and D or be indifferent between A and B or C and D because these gambles have the same expected utility. Choosing Lotteries B and C or A and D, which many respondents did, were nonrational decisions according to the rules of EUT.

Although Allais believed that such experiments called into question the relevance of the independence axiom (Allais & Hagen, 1979), Spaniel (2014) argued that for mundane decisions the independence axiom is problematic but the frequency of violation of the axiom is less clear for important decisions. Spaniel pointed out that the respondents in the Allais' experiment could have been acting rationally when deciding if they should expend the cognitive energy to figure out the probabilistic value of the fictitious gamble or if the respondents should answer as quickly as possible because they were paid for their participation regardless of the quality of the answer. Spaniel postulated that if respondents were going to actually receive money from the gambles the researcher would see an increase in the number of participants following the independence preference. Spaniel added that it might be possible to calculate a utility function for the respondents.

Huck and Müller (2012) demonstrated that more respondents showed rates of EUT violations that were greater for higher payouts versus lower payout fictitious gambles. Additionally, Huck and Müller were surprised to find slightly higher rates of EUT violations for real versus fictitious small gambles. Huck and Müller also found a strong inverse correlation between levels of education and violations of EUT. Artinger et al. (2015) acknowledged that although the EUT framework provided the basic elements of sound decision-making as well as a fairly solid description of intentional decision-making, economic examples like that explained in the Allais' Paradox paved the way for descriptive models of decision-making that portrayed how people actually make decisions.

Prospect theory. Barberis (2012) stated that for nearly 30 years economists skirted the issue of nonlinearity of EUT although behavioral psychologists placed the blame for rational decision-making failure on subjects rather than on flaws in the EUT model. Ramos et al. (2014) added that in spite of wide spread acceptance of EUT, experiments by behavioral psychologists showed that people regularly violated the axioms of rationality. This led to the development of descriptive nonutility theories of decision-making.

Kahneman and Tversky (1979) published an article that showed empirically that individuals in laboratory settings systematically violated the tenets of EUT. Kahneman and Tversky offered a new descriptive model of decision-making under risk called *prospect theory*. Tversky and Kahneman (1992) subsequently published a refined model, *cumulative prospect theory* (CPT) that extended the model to situations of uncertainty under risk. Glöckner and Pachur (2012) said that the most often cited nonutility model is Kahneman and Tversky's prospect theory. Barberis (2012) acknowledged that CPT is still considered by most to be the best description of how individuals evaluate risk in experimental settings. He qualified that claim by adding that CPT has not been widely used during the last three decades in practical economic applications.

Tversky and Kahneman (1992) said that there are two general phases in the CPT model. The first involves editing and framing. In this phase the decision-maker considers the events, contingencies, and outcomes of an alternative or prospect along with determining a reference point against which to evaluate prospects. A subjective value is determined that is based on perceived gains or losses rather than a calculated objective

utility value as in EUT. Glöckner and Pachur (2012) stated that the reference point is the state to which the decision-maker has adapted to or aspires to. Additionally Glöckner and Pachur found that although the most common reference point is the status quo, the reference point could also be influenced by the individual's hopes and goals.

Tversky and Kahneman (1992) called the second phase in making a decision the evaluation phase. The decision-maker evaluates each prospect against the reference point previously established in a two-step process. First, the individual transforms the objective values of the consequences of each outcome into a personal subjective value. Fredrickson (2013) stated that research has determined that losses are weighted more than twice as heavily as gains. Segregated gains or losses have a greater impact than a single gain or loss. Gains and losses also have diminishing marginal returns.

Tversky and Kahneman (1992) stated that the second step of the evaluation phase is the decision-weighting step. In this step the decision-maker converts statistical probabilities into individual decision weights. At the lower end of probability, the decision-maker is highly sensitive to even small probability changes as the impossible becomes somewhat possible. According to Fredrickson (2013), this hypersensitivity explains why individuals do not behave rationally to low probability risk. This behavior explains the purchase of a lottery ticket and the decision to spend vast amounts of resources to remove the final traces of hazardous materials from a location.

At the other end of the spectrum, Tversky and Kahneman (1992) observed the *certainty effect* where high probability approaches the sure thing. In the middle probability ranges, a large change in probability yields a minor change in decision

weighting. These observations explain the Allais Paradox in which a change in probability from 99% to 100% has a much greater impact on decisions than does a probability change from 10% to 11%.

Glöckner and Pachur (2012) said that transformation from strict probability to the individual's decision weighting helps explain violations of human behavior from that expected in EUT. Fredrickson (2013) called this transformation *mental accounting*. Mental accounting and loss aversion, two key elements of prospect theory, were identified by Fredrickson as key factors to explain research findings that respondents were more likely to reduce spending when the value of respondents' assets declined than to increase spending when respondents had an increase in the value of their assets.

Becker-Peth and Thonemann (2016) applied prospect theory to procurement contracts in supply chain by studying the effects of the existence of reference points on the inventory decisions for study participants. Becker-Peth and Thonemann found that procurement professionals violated rational utility maximization and demonstrated that reference point-valuation affected the decision-making of procurement professionals when it came to contraction design. Finally, Becker-Peth and Thonemann showed the impact that different revenue sharing parameters could have on the profitability of the manufacturer versus the retailer. The study highlighted the importance of considering behavioral factors in contract design to aid in delivering on the contract designer's objectives.

Hoffmann, Henry, and Kalogeras (2013) stated that more recent reference-dependent models of risky decision-making include multiple reference points. Over time

these reference points may change as decision-makers consider additional factors including the result of prior decisions. Hoffman et al. did not suggest that descriptive models were better or worse than normative models in utility maximization but only that descriptive models seemed to better reflect how individuals chose to behave.

   Regret theory. Loomes and Sugden (1982) agreed with other researchers that individuals behave in ways that are inconsistent with the concepts of rational EUT. They offered an alternative model to CPT that they called *regret theory* (RT). The RT model offered a different approach to modifying utility. Loomes and Sugden proposed that there exists a choiceless utility function which is the value an individual receives by comparing the amount of pleasure or displeasure that she will feel as a result of the alternative chosen compared to the amount of pleasure or displeasure she thinks she would feel about the outcome of the alternative not selected. If the chosen outcome is more pleasurable than the individual feels the alternate outcome would have provided, the individual feels rejoicing. If the chosen outcome is perceived to be less pleasurable than the outcome not chosen then the individual feels regret. Choiceless utility is the amount of regret or rejoicing experienced between what is and what might have been. The maximum utility is thus modified by the choiceless utility function.

   According to Loomes and Sugden (1982), many people feel the emotions of regret and rejoicing. Such psychological experiences are beyond the explanation of rational theory. In making decisions under uncertainty individuals anticipate feeling these emotions and factor these feelings into the decision-making process. Loomes and Sugden argued that although maximizing this modified expected utility function does not follow

the axioms of rationality, including such sensations into the process should not be considered irrational. Loomes and Sugden also argued that if the individual feels neither regret nor rejoicing when choosing among alternatives the utility function is not modified and this special case matches the result from expected utility theory.

Subsequent to Loomes' and Sugden's (1982) introduction of RT, other authors have improved upon the theory. Chorus (2012) developed the *random regret minimization* (RRM) model that extended the single factor concepts in RT. Chorus postulated that individuals make decisions by minimizing the potential regret one might feel by not choosing other alternatives from a group of possible choices. The RRM model takes into account multiple attributes and multiple outcomes for each attribute thereby increasing the applicability of regret theory to more complex and real situations.

Hensher, Greene, and Chorus (2013) were the first researchers to apply the RRM model to durable goods. Hensher et al. studied individual's selections in purchasing vehicles powered by different types of fuels: gas, diesel, and hybrid. Hensher et al. applied both the RRM model and the EUT model and found that the RRM model was slightly more predictive of buyer's behavior.

Summary of EUT, CPT, and RT. Ramos et al. (2014) provided a summary for the basic assumptions behind expected utility, cumulative prospect, and regret theories. EUTs assume that people behave rationally and try to maximize the value of their choice. Under CPT, people evaluate a decision against a reference point. Gains are perceived values above a reference point and losses are perceived values below the reference point. In RT, individuals are trying to minimize the regret associated with their choice.

Preferences are based on comparing the value of the decision chosen with decisions not taken. And although EUT is much more straight-forward to apply, CPT theory and RT models seemed to be better at describing behavior that departs from the tenets of rational decision-making.

Although I did not find research explicitly using expected utility or nonexpected utility theories as frameworks for research on the decision-making strategies supply chain leaders use when deciding on their key performance measures and goals, I did find a body of transportation-related literature on using such theories to describe and to predict travelers' behavior when choosing routes and modes of transportation. Ramos et al. (2014) conducted a literature review of recent studies and found that although EUT was by far the most widely used model to study travelers' decision-making behavior, there was emerging literature using prospect theories and regret theories to study the traveler's decision-making. Although Ramos et al. found that theoretically EUT and CPT were equally suited, RT seemed to be slightly more suited to modeling traveler's behavior. Ramos et al. suggested that with further evolution, nonexpected utility theories could lead to even better models. Ramos et al. called for more research to validate the use of such theories.

Other than rational decision-making theories. Artinger et al. (2015) pointed out that although rational decision-making models have dominated research in management, operations, and economics, a noteworthy amount of empirical research demonstrated that managers regularly violated the principles of rationality in their decision-making procedures. One line of research led by Tversky and Kahneman (1974),

was called the *heuristics and biases* research program. The view of researchers following the heuristics and biases tradition is that although the mental shortcuts used in applying heuristics to decision-making may simplify the process, deviations from the laws of logic and probability resulted in inferior strategies and flawed results. A competing perspective led by Gigerenzer (2015) under the *fast and frugal heuristics* research program focused on decision-making under poorly defined conditions with high uncertainty. Under these conditions heuristics performed exceedingly well and could be considered a simply rational approach. In the next few sections I will go into more detail on the research emanating from these two programs.

Tversky and Kahneman (1974) observed that people regularly made decisions that violated the rules of rational decision-making. Tversky and Kahneman postulated that humans had difficulty determining probabilities and calculating values that were required of rational decision-making models. Tversky and Kahneman found that individuals frequently used heuristics to simplify decision-making. Although Tversky and Kahneman noted that heuristics could aid in simplifying decision-making, heuristics resulted in biases that could contribute to errors in judgment. Braga et al. (2015) stated that the two most widely studied heuristics are the representativeness heuristic and the availability heuristic.

Representative heuristic. Tversky and Kahneman (1974) demonstrated that people routinely violated the rules of probability when determining the likelihood that an object belongs to one group over another. In one study, participants were given the description of a person named Steve who was shy, helpful, yet withdrawn. Participants

were given a list of possible occupations for Steve. Overwhelmingly, respondents thought

that Steve was a librarian even though the number of librarians from the total population

was much lower than for other occupations. Tversky and Kahneman called this heuristic

the representative heuristic, in which people overlooked statistical probabilities to place

Steve according to a stereotype they had of librarians. Tversky and Kahneman found that

respondents did not change their choices in light of new information that indicated

another occupation was much more likely. This manifestation of the representative bias

was in direct conflict with Bayes' rule in which a rational decision-maker would use this

new probability information to modify the decision.

According to Kahneman (2011) other biases in decision-making, categorized

under the broader notion of the representative heuristic included (a) lack of sensitivity to

sample size; (b) misinterpretations of the role of chance; (c) insensitivity to predictability;

(d) misplaced belief that the more an object is representative of the stereotype, the more

certain respondents are of their decisions; and (e) misunderstanding of future

performance as it relates to regression toward the mean. The last bias is particularly

salient to the topic of performance measurement. Kahneman explained that if the value of

a data point improved everyone celebrates the success. The same group is subsequently

disappointed, when after a later measurement, the results are poorer. Members vow to

focus more effort on the result. Kahneman added that the group's leader could

erroneously conclude that rewarding the team caused the group to take their focus off the

effort and that only with constant vigilance by the leader does performance remain high.

Kahneman stated that it is more likely that the result was due to random chance with the performance moving toward the mean of the capability for that particular process.

Availability heuristic. Tversky and Kahneman (1974) observed another group of biased-based behaviors that they called the availability heuristic. Tversky and Kahneman found respondents' predictions of the occurrence of events increased the more easily that previous occurrences of similar events came to mind. Braga et al. (2015) added that the more specific and concrete an event was, the more likely an individual was to call on the availability heuristic.

Kahneman (2011) elaborated that biases could affect the cognitive availability of information including (a) how easily an individual could retrieve similar events or objects, (b) the ease of a search set such as the number of words that begin with the letter *r* versus the number of words with *r* in the third position, and (c) how the construction of a coherent narrative facilitated retrieval. Kahneman observed that a decision-maker might use the availability heuristic to estimate the number of instances of an event, the likelihood of an event occurring, or the probability of a co-occurrence based on the ease of retrieval, construction, and association.

Braga et al. (2015) stated that there is often confusion regarding the use of the representative heuristic versus the availability heuristic. Braga et al. clarified the situation by saying that when a decision is reliant upon the degree to which an event matches a stereotypical representation of the target, the individual would use the representative heuristic. If the decision relied on the ease with which specific instances came to mind then the individual would draw upon the availability heuristic. Braga et al. drew upon

prior work on construal theory that demonstrated that decision-making depends on whether the decision is made at a more abstract versus a more concrete level. Braga et al. demonstrated that higher levels of abstraction were correlated with the use of the representative heuristic whereas lower levels of construal were favorably correlated with the use of the availability heuristic.

Adjustment and anchoring heuristic. A third heuristic highlighted by Tversky and Kahneman (1974) was the adjustment and anchoring heuristic in which an individual makes a prediction based on a starting reference value and then adjusts up or down from that reference point. Tversky and Kahneman showed that respondents estimated the number of countries in Africa to be greater after first answering if the number of countries was greater or less than 65 than respondents did if they first anchored on whether there were more or less than 10 countries in Africa. Kahneman (2011) said that hundreds of studies have shown the powerful effect on judgment from irrelevant anchors. Such experiments highlighted the level of irrationality demonstrated by respondents in how reference points were determined. Kahneman added that he and Tversky found biases both in establishing the original anchor point and in determining the size and direction of the subsequent adjustment.

Cheek and Norem (2016) stated that anchoring effects occur in real world situations and that experts (eg. doctors and judges) are not immune to the influence of anchors. Cheek and Norem found that respondents with an analytical thinking style were more susceptible to anchoring effects than respondents with a more holistic style of thinking. Bahník and Strack (2016) found that whether a reference point serves as an

anchor depends on the compatibility of that anchor with the topic of the decision being made. Bahník and Strack added that plausible anchors have more influence than implausible anchors.

Malhotra, Zhu, and Reus (2015) studied the impact of anchors on setting the premiums paid for mergers and acquisitions. Malhotra et al. found that decisions related to prices offered for future acquisitions were strongly influenced by the premiums paid for the most recent deals. Malhotra et al. recommended that executives and their financial advisors make the possible effects of anchors a transparent part of the acquisition offer process.

Heap (2013) stated that although the most commonly used anchors include *the status quo* as well as *the recent past*, such anchors can lead to overly optimistic results. Tamir and Mitchell (2013) described *anchoring on ourselves,* in which the decision-makers anchor on their own values and beliefs and then adjust for differences in the values and beliefs of others. Decision-makers also adjust for ways in which they believe their situation was unique. Tilburg and Igou (2014) extended this line of research and found that respondents strongly prefer to continue down the same path rather than incorporating new strategies into their decision-making process adding further credence to the notion that cognitive limitations play a key role in other-than-rational decision-making.

Kahneman (2011) pointed out that there is evidence for a nearly overwhelming variety of priming effects on anchors. Decision-makers' thoughts and subsequent decisions are highly influenced by the external environment. Kahneman suggested that

the decision-maker be wary of any anchoring value offered externally and that the decision-maker works diligently to explore and remove from the process such biases, especially for important decisions.

The fast and frugal research program. Hafenbrädl, Waeger, Marewski, and Gigerenzer (2016) conveyed a more positive perspective on the use of heuristics in decision-making than did researchers from the heuristics and biases program. Hafenbrädl et al. labeled many of these rules as fast and frugal heuristics because research had shown that such strategies approached or even outperformed outcomes using rational rules of decision-making. The heuristics were considered frugal because a smaller solution set was considered which allowed the individual to come to a decision more efficiently.

Artinger et al. (2015) stated that there are three main differences between the heuristics and biases tradition and the fast and frugal heuristics program:

- *Descriptive versus prescriptive models.* The heuristics and biases tradition was deemed more descriptive in nature relying on labels such as in the availability heuristic. In the fast and frugal heuristics program the objective was to apply a normative perspective by making salient the cognitive processes used in making decisions. In this way simple prescriptive models could be developed.

- *Concept of ecological rationality.* The study of ecological rationality is normative in nature and is intended to help determine the conditions under which a specific heuristic was more successful. In the heuristics and biases tradition a heuristic was considered less successful than the use of logic and statistics. The fast and frugal heuristics perspective was that a heuristic could

be considered a normative choice if it led to as good as or better performance than that provided by a more complex model.

- *Less may be more.* Since the heuristics and biases program did not account for risk versus uncertainty, followers believed a heuristic could never provide a better outcome than the result from a weighting and adding utility model. The only benefit from the use of a heuristics under the heuristics and biases mindset was that it required less cognitive effort. Gigerenzer (2015) found that in situations with high complexity and variability an inverse U-shaped relationship existed between accuracy and amount of information stating that at some point, more information could actually be detrimental to the quality of the decision. The fast and frugal perspective was that simpler models were not only easier cognitively but could produce superior results.

Gigerenzer (2015) found there to be a plethora of descriptive heuristics in the literature. He advised that it would be beneficial to develop more general theoretical frameworks composed of building blocks and proposed the following approach.

1. Define the search rule used to determine the population of possible results.

2. Define the rules used to decide when to stop a search.

3. Define the rules used to determine how to reach a final decision.

Artinger et al. (2015) used the building blocks above to develop five general classes of heuristics.

Satisficing. Simon (1955) published seminal work on bounded rationality. Although he acknowledge that utility maximization might result in the optimal solution,

this was not practical in a world filled with uncertainty. He stated that satisficing, defined as selecting the first option that meets a minimum aspiration level, was often a good enough strategy. Gigerenzer (2015) operationalized the satisficing heuristic as follows.

1. Search rule: Establish an aspiration level. Search through possible options.

2. Stopping rule: Stop the search with the first option that meets the aspiration level.

3. Decision rule: Choose this option.

Artinger et al. (2015) noted that establishing a fixed aspiration level reduced the error attributed to sample variance while adjusting the aspiration level during the process would reduce bias. Artinger et al. found that used car dealers who applied rational theory by adjusting prices with changes in the market had lower profit than dealers who reduced prices at a fixed time interval. Berg (2014) found that successful entrepreneurs in the Dallas-Fort Worth, Texas area used satisficing behavior when choosing the locations for their businesses. The aspiration level was set at a minimum acceptable financial return and a very small population of locations was in the search set. In the majority of the cases between one and three locations were considered. This research had important implications for local economic policy because incentives such as tax reduction were not part of the search criteria used by the entrepreneurs in location selection.

Tallying and 1/N. Hafenbrädl et al. (2016) stated that random fluctuations could act as noise to a decision-maker attempting to find the true signal when determining the relations among variables. Hafenbrädl et al. stated that a simple strategy called tallying,

also known as 1/N, performed better than more complex linear regression models under certain conditions. Gigerenzer (2015) defined tallying through the following process.

1. *Search rule:* Search through all possible outcomes for an option in any order, adding positive outcomes and subtracting negative outcomes to the resulting sum.

2. *Stopping rule:* Stop after N options, where N is a subset of all possible choices.

3. *Decision rule:* Choose the alternative with the greatest result. If there is a tie, then guess.

DeMiguel, Garlappi, and Uppal (2009) compared the predictive capabilities of using the 1/N fast and frugal heuristic with 14 other investments models. DeMiguel et al. found that investing equally in the shares of a stock portfolio not only simplified the investment process but resulted in as-good-as or better returns than more complex strategies. Albar and Jetter (2013) found that tallying performed as well as a regression model in evaluating new product ideas. Hafenbrädl et al. concluded that in noisy environments simple heuristics could eliminate the effects of variance leading to more accurate predictions.

Lexicographic strategies. Artinger et al. (2015) described a group of heuristics that involved sequential cognitive processes in the decision-making strategy. The rules for lexicographic-type heuristics follow.

1. *Search rule:* Order the attributes of an alternative by level of importance.

2. *Stopping rule:* Stop when the decision-maker finds the first attribute that differentiates between the alternatives.

3. *Decision rule:* Choose the alternative that contains the more highly valued

attribute level.

Gigerenzer (2015) placed the *take-the-best* (TTB) heuristic in the lexicographic category. Gigerenzer stated that application of the TTB heuristic required the decision-maker to order the attributes by decreasing importance. The decision-maker must evaluate each possible choice against those attributes winnowing out the less favorable options through pair-wise comparisons.

Manley, Orr, and Cheng (2015) stated that TTB is the most widely applied heuristic and has been studied in a variety of situations including medical decisions, deciding on investment options, court judgments, and election strategies. Manley et al. applied TTB to route-choice decisions made by minicab drivers in London. Manley et al. developed a framework that included five attributes and the ordering sequence of these attributes by collecting more than 500 thousand cab ride observations. The framework described the rapid good enough decisions made by cabbies when encountering uncertainties in route choice.

Tappeiner, Howorth, Achleitner, and Schraml (2012) found that reputation and voluntary accountability efforts were the most important attributes considered by a group of investors in deciding among social ventures in which to invest. Tappeiner et al. were surprised at the small number of attributes that were considered given the weighty financial nature of the decisions. Mousavi and Gigerenzer (2014) explained that use of TTB is ecologically rational in conditions where there are many interrelated cues that are highly redundant. Mousavi and Gigerenzer stated that use of TTB allows the decision-

maker to ignore some cues and associated dependencies between cues during the ranking process. Mousavi and Gigerenzer added that TTB often yields better results than more complex linear regression models or decision-tree algorithms.

Similarity. Hafenbrädl et al. (2016) said that some heuristics, such as the similarity heuristic, exploit the workings of the human brain with regards to memory and cognition. Artinger et al. (2015) stated that the similarity heuristic is regularly used in decision-making when managers encounter novel situations. A key element of this process included identifying an important characteristic of the decision and then finding a prior experience that shared this same characteristic. Artinger et al. stated that priors could be found by searching one's own experience set, by asking other experts, and by research such as that found through case studies. Gigerenzer (2015) defined the rules for the similarity heuristic as follows:

1. *Search rule:* Search for the experience that is closer in similarity to the new experience than other prior experiences.

2. *Stopping rule:* Stop the search when a more similar experience is found.

3. *Decision rule:* Conclude that the alternative chosen has a higher similarity value on that characteristic than others in the search set.

Artinger et al. (2015) stated that the success of this heuristic is based on both the selection of the relevant characteristics for a decision as well as on finding an applicable prior experience. Lovallo, Clarke, and Camerer (2012) found that in many cases, leaders only considered a single analogous case when making strategy decisions. Lovallo et al. demonstrated that creating a larger reference class led to higher quality decisions in three

ways: (a) considering more analogous references increased the options available to decision-makers; (b) analyzing the various analogies could lead to a deeper, more fact-based assessment; and (c) paying particular attention to the most and least similar cases increased the quality of the analysis. Gary, Wood, and Pillinger (2012) stated that although research has shown senior executives make decisions by drawing on analogous prior experiences, other research has shown that decision-makers struggle to find similar analogies to use in making effective decisions. Gary et al. found that exposure of executives to a wide variety of experiences improved the ability of the decision-maker to seek out and find appropriate prior experiences to use as the basis for making quality decisions.

Recognition. According to Artinger et al. (2015), there has been a plethora of research on the impact of recognition on decision-making. Gigerenzer (2015) stated that the recognition heuristic was originally considered as a building block for other heuristics such as the take-the-best heuristic. Gigerenzer added that the recognition heuristic is called into play when a decision-maker does not know or cannot make a logic-based decision with the criterion presented. Gigerenzer described recognition-based decisions as those in which simply recognizing an object is a strong predictor of that object figuring strongly in the final decision. For example, when respondents were asked which city has more residents, Tokyo or Busan, greater recognition of Tokyo caused decision-makers to infer that Tokyo was also more populous than Busan.

Gigerenzer (2015) stated that three conditions needed to exist in order for the use of the recognition heuristic to be considered as ecologically rational: (a) recognition

validity is greater than chance would indicate, (b) inferences are made from memory rather than through the aide of visual cues, and (c) recognition comes from the decision-maker's natural environment rather than from a lab setting. Artinger et al. (2015) described the process for applying the recognition heuristic:

1. *Search rule:* Search for an object that is recognizable.

2. *Stopping rule:* Stop once a recognized object is found.

3. *Decision rule:* Conclude that the recognized object has a greater value for a specific attribute than the value of other choices in the search set.

Michalkiewicz and Erdfelder (2016) stated that the recognition heuristic has been shown to perform well across many conditions. Other research has shown that in situations where the recognition validity for a choice is not substantial, decision-makers may invoke a recognition-plus-evaluation process in which respondents may overrule the recognition heuristic and look for additional cues (Gigerenzer, 2014).

Thoma and Williams, 2013) studied the impact of brand recognition on consumers' choices. Respondents were presented pairs of consumer items. Under each item was the name of the manufacturer and the number of ratings stars. Respondents were asked to choose between brands (one famous and one relatively unknown) of a variety of consumer items such as computers, earphones, and mobile phones. Response times were recorded. Respondents chose the recognized brands more often, even in the presence of negative ratings. Positive ratings had limited effect on the percent of recognized brands chosen. There was also evidence that in some cases, respondents used another cue, the number of stars in rating the product, on a limited basis. Although the

results indicated that consumers do use cues other than brand in making choices, brand recognition was a factor in that choice.

Current state of heuristics research. Gigerenzer (2015) highlighted three major shifts over time in research on heuristics. First there has been a move from heuristics as general labels, like the availability heuristic, to predictive models with rules, such as the similarity heuristic. Second there has been a shift from studying heuristics from the perspective of their preferences to their inferences. Initially accuracy of decision-making was measured against the rational process of weighting and summing. Heuristics measured in this way could only be seen as more frugal, not more accurate. By having clear criteria for inferences, it has been possible to demonstrate that the decision-maker used simple heuristics and delivered higher decision accuracy than with weighting and adding rules. Research has also moved from assuming that in all situations logical or statistical rules are normative in all circumstances, to a lens of ecological rationality in which the environment determines the preferred approach; heuristic or otherwise.

Hafenbrädl et al. (2016) offered a contemporary example of the value of heuristics. Hafenbrädl et al. described the decision-making process used when pilots Chesley Sullenberger and Jeffrey Skiled decided, after their US Airways plane was hit by a flock of geese, to land the aircraft on the Hudson River rather than returning to LaGuardia airport. When asked how they made the decision, the pilots responded that they applied the *gaze heuristic*. If after fixing one's eyes on the landing spot, the spot rises in the line-of-sight, a pilot knows he won't be able to make it to the landing spot. This simple heuristic allowed the pilots to make a more rapid decision than calculating

altitude and speed from the complex gauges on the aircraft resulting in the lives of all 155 passengers being saved.

Although there are a myriad of studies on the application of heuristics to a wide variety of decisions, I was not able to find any studies related to the use of heuristics by leaders in selecting performance measures and related goals. To my knowledge, no one has ever applied the heuristics and bias lens in exploring how supply chain leaders choose the measures that matter. I now turn to the literature on the third school of thought in decision-making theory, naturalistic decision-making, defined as research in how decisions are made in complex natural environments.

Naturalistic decision-making. According to Shan and Yang (2016), applied scientists found that the prior research on judgment and decision-making did not address the problems they were encountering when studying experts making decisions in actual work environments.

Klein and Wright (2016) highlighted a conference held in 1989, led by Judith Orasanu, of the Army Research Institute for Behavioral and Social Sciences and attended by 30 researchers, as the catalyst for the naturalistic decision-making program. A year earlier the highly experienced crew of the Vincennes, a US Navy destroyer, shot down an Iranian passenger plane, mistaking it for an attacking F-14 fighter plane (Klein, 2015). Klein and Wright summarized the objective of the members of this initial conference as a group of researchers who were interested in how experts made decisions under complex and changing conditions with ill-defined goals, organizational constraints, and high risks such as that found in military, emergency response, and medical situations. Subsequently,

the *Cognitive Engineering and Decision-making* technical working group was formed, a

dozen conferences have been held and the *Journal of Cognitive Engineering and*

Decision-making has been created to advance research in naturalistic decision-making.

Klein and Wright (2016) asserted that unlike other forms of decision-making

research, naturalistic decision-making is more focused on creating useful models rather

than testing hypotheses and is more interested in descriptive models that can be applied to

messy problems with vague goals and for which there are no correct decisions. Klein and

Wright added that these conditions describe those situations in which most people work

and make decisions. Klein (2015) stated that research has moved beyond the belief that

the only way to make a quality decision is to generate a variety of options and select the

best one. Klein added that research on learning has evolved beyond believing expertise is

developed by learning policies and procedures. Researchers have demonstrated that

gathering more information may not reduce uncertainty. In fact, more information can

actually harm performance. Klein added that although researchers in the rational and the

heuristics and biases programs focus on cognitive limitations and search for methods to

overcome such limitations to reduce mistakes, naturalistic decision-making researchers

focus on human capabilities. Naturalistic decision-making researchers also recognize

good performance is about much more than preventing mistakes; it is about discovering

the importance and contribution of experience to making good decisions in a variety of

complex situations.

Recognition-primed decision model. According to Riegel et al. (2013), although

many naturalistic decision-making models have been developed, Klein's recognition-

primed decision-making model is considered the prototypical model for how people make decisions in real-world situations. Klein (2015) described the recognition-primed decision-making model as the combination of intuition and analysis. Boyes and Potter (2015) added that the recognition-primed decision-making model contains four main elements: (a) situation recognition, (b) situation comprehension, (c) mental simulation, and (d) sequential option evaluation.

Boyes and Potter (2015) stated that in the situation recognition step of the recognition-primed decision-making model, the decision-maker attempts to match the current situation to prior similar experiences. Klein (2015) said that when decision-makers need to act quickly, they match patterns between the current and previous situations. Macquet and Skalej (2015) added that if the current situation matches a typical situation in the decision-maker's memory, then the decision-maker carries out the action associated with that typical situation and the decision-maker has a clear path forward. Klein cautioned that relying only on intuition is risky because occasionally pattern matching could generate flawed options.

Macquet and Skalej (2015) explained that if the situation encountered by the decision-maker is not similar to prior experiences, the decision-maker must focus on understanding the situation. Boyes and Potter (2015) said the decision-maker accesses four types of information: (a) plausible goals, (b) relevant cues, (c) expectancies and (d) typical actions to aid in understanding. Boyes and Potter used the recognition-primed decision-making model to describe the decision-making of expert guides leading outdoor learning expeditions. A plausible goal for the group leader encountering an unexpected

lighting storm was to reduce the level of risk for the group's members. Relevant cues for the situation included observing anvil shaped cumulonimbus clouds and hearing thunder in the distance. Expectancies provided a perspective on what is likely to happen under these conditions that included the potential of a storm with high winds and lightning. Typical actions were the result of generating options such as moving to a safer location.

According to Macquet and Skalej (2015), situational understanding allows the decision-maker to draw upon prior somewhat similar experiences and apply an adaptation to a typical action that is appropriate for the current situation at hand. Okoli, Weller, Watt, and Wong (2013) added that the decision-maker would invoke mental simulation to assess if the first action that came to mind would yield an acceptable outcome. If so the process would move forward as for a simple match. If not, the decision-maker would continue the simulation with a modification of the action and reassess. If this modified option still failed to lead to an acceptable outcome the decision-maker would exert more mental energy to search for another possible action. Klein (2015) summarized that the recognition-primed decision-making model describes the balance of relying on intuition for speed where possible and calling upon a more deliberate and analytical strategy when the situation requires. Gore et al. (2015) said researchers have demonstrated empirically that experienced decision-makers generally accept the first option they considered as satisfactory. Boyes and Potter (2015) reiterated that greater experience leads to an expanded repertoire that improves the speed and quality of expert-based decision-making in the field.

Research applying the recognition-primed decision-making model. Riegel et al. (2013) applied the recognition-primed decision-making model to explore how adults dealing with chronic heart failure made decisions regarding self -care. Riegel et al. found that respondents applied a naturalistic decision-making process by applying aspects of situation understanding along with mental simulation to generate and evaluate plausible courses of action for their individual care. Riegel et al. also found that the conditions that influenced the decision-making process included (a) prior experience; (b) environmental situations including ambiguity, uncertainty, high stakes, urgency, illness, and the involvement of others in the decision; and (c) the existence of personal goals that helped inform the decision. These conditions are consistent with those cited by Klein and Wright (2016) as hallmarks of the naturalistic decision-making environment.

Cristancho, Vanstone, Lingard, LeBel, and Ott (2013) expanded upon prior research on how surgeons assess and react to challenges encountered in the operating room. Cristancho et al. used a grounded theory approach to interview surgeons who conducted a variety of elective and emergency surgeries. Cristancho et al. created a decision-making model that incorporated the three major themes from the research: (a) assessing the situation, (b) a reconciliation process, and (c) implementing the plan.

There were elements in the Cristancho et al. (2013) model that were similar to Klein's (2015) recognition-primed decision-making model developed by observing firefighters in the field. Some similarities included assessing the situation, understanding what potential solutions might look like, and evaluating solutions against a goal. Klein's recognition-primed decision-making model focused on the decision process rather than

what triggered the decisions. The Cristancho et al. model looked not only at processes but also considered surgery-specific contextual characteristics including (a) generation of a pre-operation plan, (b) time pressure that vacillated between moderate and extreme levels, and (c) familiarity with the operating environment and with the team members going into the situation. The reconciliation process in the Cristancho et al. model went much deeper into detail than did the mental simulation process of Klein's model.

Macquet and Skalej (2015) applied the recognition-primed decision-making model to study how elite student athletes balanced their time between studies and athletic training. Macquet and Skalej applied both sense-making theory and the recognition-primed-decision-making model to explore how these student athletes made sense of the situations they experienced that helped the respondents decide how to handle the demands of being students and athletes. Macquet and Skalej created a recognition-primed decision-making-based model that described the decision-making processes for both the simple match situations and for those situations in which the student athletes needed to further understand and diagnose the situation they were experiencing. Macquet and Skalej found that experienced third-year students were more likely to actively manage their training and recovery schedules with their coaches than were less experienced first year students providing further empirical support both for the recognition-primed decision-making model and that experience improves decision-making quality.

According to Gore et al. (2015), an emerging area of decision-making research is the combination of naturalistic decision-making methods with rational as well as heuristics and biases traditions. Shan and Yang (2016) argued that the naturalistic

decision-making and the fast and frugal heuristics programs share some of the same belief structures, including (a) concept of the *proficient decision-maker* working within the constraints of bounded rationality, (b) a process orientation around decision-making, and (c) decision rules that include decision-action matching. Gore et al. acknowledged that one area where the programs differ is that fast and frugal heuristics is focused on the development of formal decision-making models whereas naturalistic decision-making researchers challenged whether the importance of considering the decision-environment and the complexity required of real-world decision-making could be reduced to generic context-free models.

Shan and Yang (2016) highlighted areas of potential research that could benefit from synergies between the fast and frugal heuristics and naturalistic decision-making programs. Examples included enhancing applicability of fast and frugal heuristics models in the field aided by using naturalistic decision-making research techniques as well as applying fast and frugal heuristics to map the contextual boundary conditions in naturalistic decision-making studies. Shan and Yang recommended that researchers consider the essential elements of each decision-making program and select the decision-making approach most appropriate for the phenomenon to be studied.

Although there were many examples in the literature of applying naturalistic decision-making methodologies to studies in fields such as medicine, military field operations, aviation, and emergency response situations, as was the case with rational and heuristics and biases theories, I did not find any research applying naturalistic decision-

making to exploring the decision-making processes used to develop performance measures and goals either in general or specifically to the supply chain field.

Gap in the Literature

In the research literature, I found a plethora of articles on performance measurement, supply chain performance measurement, and decision-making theory. There was a paucity of articles on the intersection of *decision-making theory* and *supply chain performance*. Most of the literature on supply chain performance used quantitative methodologies. Of the articles, I did find on decision-making and supply chain performance, most were related to decision support models for supplier selection or transportation mode optimization. None specifically addressed the decision-making strategies used to select performance measures thereby supporting the existence of a gap in the literature regarding the decision-making strategies actually used by supply chain leaders in choosing their performance measures and goals. This gap validated the need for the study.

Research Approach

I did not find studies that made salient the actual lived experiences of how supply chain leaders chose the performance measures and goals for their organizations. Yin (2014) said that a key criterion for determining what research method to use depends on the questions being asked. If the intent is to answer a how, why, or a what question that is exploratory in nature, the researcher uses a qualitative approach. If the question can be further described by how many then a quantitative method is more appropriate. The research question was, what are the decision-making strategies used by supply chain

leaders in choosing their key performance measures and goals? It logically followed that insights gained from a qualitative study of the phenomenon of choosing performance measures and goals through the lens of decision-making theory might be beneficial to both the academic and practitioner communities.

Phenomenological Research

Giorgi (2012) advised that phenomenology is the research method of choice when the goal is to generate an increased understanding of the subjective experiences of a group of people from their own perspectives. Englander (2012) stated that the phenomenological approach attempts to answer a research question framed as, "What is it like?" Englander added that the heart of phenomenological study is targeted at uncovering the general or invariant structure inherent in a phenomenon. Anosike, Ehrich, and Ahmed (2012) observed that there is a dirth of research in management studies employing phenomenology as a methodology. Anosike argued that is is imperative to apply nonpositivist methods such as phenomenology in order to bridge the gap between management theory and practice.

Descriptive Phenomenology

Englander (2012) found that students often mix various qualitative methods failing to understand the underlying philosophical premises of each method. Anosike et al. (2012) added that there are two major schools of thought in phenomenological research; the hermeneutic approach as espoused by van Manen and the descriptive approach under the leadership of Giorgi. The focus for hermeneutic research is to provide deep insights into human experience or phenomenon. The result is a narrative by the

researcher that interprets the meaning of the phenomonen. The focus for descriptive research is to provide an accurate description of the subjects' common experience of the phenomenon. The result is a general statement of the phenomenon including a description of the essential structures of the experience studied. Mayoh and Onwuegbuzie (2015) agreed that a descriptive phenomenological approach is used when the researcher wants to collect the common essences of an experience within a group to create an essential structure of the phenomenon. Mayoh and Onwuegbuzie added that important distinctions of the descriptive phenomenological approach from other phenomenological perspectives are that humans are free agents who are responsible for the impact they make on the world and that there is one correct objective interpretation of the common essences of the phenomenon.

Knight et al. (2012) applied a descriptive phenomenological approach to gain a better understanding into the factors that made the 10 nursing students determine if they could remain in their nursing program along with exploring the actions taken to help ensure that these students could successfully graduate. Knight et al. found that common obstacles encountered included (a) work life balance, (b) commitment to study, (c) financial hardship, (d) family crises, and (e) being older. Knight et al. discovered that students' desire to achieve their goal of graduating was critical to program completion. Knight et al. also found an overarching theme for completion was support with subthemes that included support from family, support from classmates, and support from program staff and professors. Knight et al. recommended that the School of Nursing add pastoral care aspects to augment previous student support systems in order to increase

rate of successful graduation from the nursing program.

Fawcett, Fawcett, Watson, and Magnan (2012) acknowledged that research has shown that identifying and linking complementary capabilities among business partners via collaboration leads to improved supply chain performance. Fawcett et al. found a gap in the knowledge regarding how supply chain leaders construct collaborative capability. Fawcett et al. applied an extreme case sampling strategy and interviewed decision-makers from 15 companies recognized in the trade press for their excellence in supply chain collaboration.

Although the Fawcett et al. study was a multi-case study resulting in the creation of a model that described the construction of collaborative capabilities and not a descriptive phenomenology, this was one of the few qualitative studies I found related to supply chain. Fawcett et al. focused on the use of best-in-class respondents in order to ferret out process commonalities among the organizations. This approach was similar to my plan to use leaders from highly recognized supply chain organizations to uncover common decision-making strategies used in selecting key performance measures and goals to drive performance. Fawcett et al. stated that respondents' expressed frustration at their inability to demonstrate performance improvement. Fawcett et al. called for research related to designing PMSs and defining success providing further support for the need of the proposed study on selection of performance measures and goals.

Signori, Flint, and Golicic (2015) stated that prior research confirms the importance of creating sustainable supply chains. Signori et al. uncovered the need for further understanding of how executives perceive sustainability and how these

perspectives affect supply chain management. Signori et al. took what they called a combination research approach using elements of grounded theory and phenomenology. Signori et al. interviewed 112 managers in 88 organizations in the wine industry during a period of five years. The grounded theory aspect allowed Signori et al. to develop the concept of sustainable supply chain orientation that depicted (a) how a leader embraced sustainability, (b) how the leader aligns motivations with sustainability, and (c) how likely the leader is to partner with the supply chain. Signori et al. then applied a hermeneutic phenomenological approach to develop 10 leadership profiles that characterized sustainable supply chain organizations. To quote Signori et al., "We sought the best interpretations of what the managers were trying to describe of their own experiences." (p. 541). Signori et al. explained that the creation of the 10 profiles could provide insights into organizational obstacles and a path forward for reaching the desired sustainable supply chain level.

Signori et al. (2015) used interpretive rather than descriptive phenomenology to provide a view into the many different ways individual leaders in the wine industry experience sustainability along with the tension between sustainability and the need to run an efficient supply chain. This approach was the inverse of my research. My goal was to reach transcendental subjectivity of the common aspects actually experienced by supply chain leaders in deciding on key performance measures and goals for their organizations. This goal was the impetus for my descriptive phenomenological study on how supply chain leaders choose the measures that matter.

Summary and Conclusions

A review of the literature demonstrated that there is a linkage between the establishment of performance measures and goals, individual behavior and performance, and ultimately, organizational performance. There is a large body of literature on various performance measurement frameworks that capture both financial and nonfinancial measures, metrics, and indicators with the most prevalent being Kaplan and Norton's Balanced Scorecard. There is a large amount of research on how leaders should design, implementation, use, and refresh PMSs. There is a gap in the literature on how supply chain leaders actually chose the performance measures and goals that will drive organizational performance.

It seemed logical that one approach to closing this research gap was to apply decision-making theory to exploring the phenomenon of performance measure and goal selection. The literature on decision-making theory was categorized into three general schools of thought: rational decision-making theory, heuristics and biases theory, and naturalistic decision-making. I proposed the use of a descriptive phenomenological approach to make salient the decision-making processes used by supply chain leaders as they retrospectively explained how they chose the performance measures and goals to drive their organizations' performance. Rendering the decision-making processes explicit could help improve the quality of the decision-making processes in choosing performance measures and goals in the future thereby assisting in improving the overall performance of the organization and contributing to positive social change.

In Chapter 3, I justify the rationale for selecting the descriptive phenomenology as the appropriate research method for the study. I describe in detail how I proposed to conduct the study, the process I planned to use for selecting appropriate participants, how I planned to protect those participants, how I planned to collect the data, and how I planned to ensure that the research results were valid and trustworthy. I then tie the problem statement with the proposed research methodology and describe how the approach may help close the gap in the knowledge and understanding of how supply chain leaders chose the key performance measures and goals for their respective organizations.

Chapter 3: Research Method

According to Franco-Santos et al. (2012), improvements in performance measurement system design have been shown to improve overall organizational performance. I conducted a qualitative descriptive phenomenological study to increase the understanding of the decision-making strategies used by supply chain leaders in their selection of key performance measures and associated goals. Mayoh and Onwuegbuzie (2015) said that although each respondent's experiences related to a phenomenon are unique, a descriptive phenomenology can be used when the researcher is seeking to find features of the experience that are common across respondents. Research that makes salient commonalities in the decision-making strategies used in companies recognized as supply chain leaders could help supply chain managers in other organizations improve the processes they use to select the key performance measures and goals for their own organizations.

In this chapter, I share the research question and the research design, tradition, and methodology, along with the rationale for using the research approach to address the research question. I discuss the other research methods I considered and explain why I opted against using those methods. I make explicit my role as researcher in the study. In addition, I describe specific aspects of the research including (a) the logic used for participant selection; (b) the instrument used for data collection; (c) details of the pilot study; (d) procedures for recruitment, participation, and data collection; and (e) the data analysis plan. I also describe how I ensured that the study was trustworthy by addressing

issues of credibility, transferability, dependability, confirmability, and ethical procedures.

I then provide a summary of Chapter 3.

Research Design and Rationale

The central research question for the study was, What are the decision-making strategies used by supply chain leaders in choosing the key performance measures and goals for their organizations? The phenomenon for the study was the decision-making strategies used by supply chain leaders in choosing the key performance measures and goals for their organizations. Yin (2014) stated that the most important criterion for determining what research method to use is the question being asked. Yin added that if the intent is to answer a how, why, or a what question that is exploratory in nature, the researcher should use a qualitative approach. Although there was a plethora of research in the literature on performance measurement frameworks and on the design, implementation, and use of performance measurement systems (e.g., Agostino & Arnaboldi 2012; Paulraj et al., 2012), there was a gap in the research on how supply chain leaders actually choose their measures and goals. I took a different perspective on the topic by exploring the decision-making strategies used by supply chain leaders in selecting their key performance measures and goals through the lens of decision-making theory. I followed the advice of Mayoh and Onwuegbuzie (2015) that a qualitative approach is the more appropriate method to answer a research question that is exploratory in nature.

Golicic and Davis (2012) observed that most researchers studying supply chains have relied on quantitative research methods. The authors stated that issues in supply

chain are becoming increasingly complex warranting a need for diverse approaches to research. Golicic and Davis noted that researchers in the supply chain arena have been criticized for their lack of willingness to deploy research methods that are more appropriate for generating theory and for exploring complex issues. Anosike et al. (2012) argued for the use of nonpositivist research methods in general, and more specifically, for the use of phenomenology, to increase the understanding of the complex nature of management practices within an organization.

Giorgi (2012) recommended using the phenomenological method when the researcher wants to gain an increased understanding of the subjective experiences of a group of people from their own perspectives. Anosike et al. (2012) specified the use of Giorgi's descriptive phenomenological analysis (DPA) approach to draw out the meanings and essences of managers' organizational experiences. Mayoh and Onwuegbuzie (2015) said that a DPA approach is used when the researcher wants to explicate the collective common essences of an experience within a group to create an essential structure of the phenomenon. Anosike et al. stated that a key element of DPA is the development of a single objective interpretation of the common essences of a phenomenon that remains intact regardless of context.

There are several other qualitative approaches I considered and discarded. The first was the narrative methodology. Wang, Kim Koh, and Song (2015) stated that narrative research entails personal storytelling. Through such stories the researcher uncovers the desires, needs, and goals of the participant within that person's individual cultural framework (Wang et al., 2015). I did not choose narrative research because I was

seeking to find commonality in the decision-making process used in selecting measures and goals among the respondents in the study.

The second approach I considered was ethnography. Cunliffe and Karunanayake (2013) said that ethnographic research results in rich and detailed accounts of the daily life of a community to better understand the cultural workings of the population under study. Although there may be cultural factors that influence the decision-making strategies of the respondents, understanding culture was not the central focus of the study. Thus, ethnography was not appropriate for this study.

The third approach considered was grounded theory. Timmermans and Tavory (2012) stated that a grounded theory approach is most appropriate when the researcher is looking to create new theory. I did not choose a grounded theory approach because there existed a substantial amount of research on decision-making theory (e.g., Paulraj et al., 2012). Rather than develop a theory, I intended to discover what decision-making strategies were actually used by supply chain leaders in choosing their measures and targets.

The fourth approach I considered was the case study. Yin (2014) explained that a case study methodology is used when the researcher is seeking a deep understanding of the how or why nature of the phenomenon of interest. I did not choose the case study because I wanted to find commonalities in the decision-making process used by supply chain leaders in choosing their key performance measures and goals across several organizations rather than deeply exploring a single case or two.

I also considered conducting a hermeneutic phenomenological study. According

to Smith and Osborn (2015), hermeneutic or interpretive phenomenology is based on the ontological view that there are many realities that can exist for the same phenomenon. Because my goal was to study the common essences of the decision-making strategies used by supply chain leaders in choosing the key performance measures and goals for their organizations, I discarded the hermeneutic phenomenological method.

Mayoh and Onwuegbuzie (2015) explained that a mixed-methods approach is a combination of qualitative and quantitative methods in the same study. A sequential mixed-methods approach using a qualitative followed by a quantitative method could have been applicable to the study. Mayoh and Onweuegbuzie cautioned that mixed-methods studies require extensive resources and can take a long time to complete. For these reasons, I opted against using a mixed-methods approach. Subsequent researchers could follow my research by determining potential correlational relationships between the decision-making strategies used by supply chain leaders in choosing the performance measurements for their supply chain organizations and the performance of those respective organizations.

Gill (2014) stated that in determining the type of phenomenological approach to use, the researcher must factor in one's own epistemological and ontological orientations along with the nature of the research question. Gill added that if the researcher is looking to describe the experience of several participants as one common essence, then Giorgi's descriptive phenomenology is the appropriate method. After reviewing a variety of research approaches and based on the worldview of descriptive phenomenological

research, I determined that a DPA methodology was the most appropriate method to answer my research question.

Role of the Researcher

Bevan (2014) stated that the phenomenological researcher is fully immersed in the process from the start, unlike quantitative research in which researcher becomes immersed at the point of data analysis. Gill (2014) added that in the phenomenological methodology, the researcher is the instrument whose role it is to make explicit the lifeworld experience related to the phenomenon under study as described by the experiencer. Giorgi (2012) added that the researcher must put aside personal knowledge, beliefs, and preconceived notions by adopting an attitude referred to as bracketing or epoché. Gill said that adopting an epoché approach allows the researcher to take a transcendental perspective and look upon the phenomenon with a critical eye, taking nothing for granted.

Bevan (2014) said that an important role of the researcher occurs during the interview process. Bevan emphasized that the researcher must participate in the interview by being an active listener; asking questions to probe more deeply into the participant's experience or to provide clarification. Bevan shared that interviewing is a craft and that by adopting an orientation of deliberate naiveté, the researcher can achieve the phenomenological reduction that is at the heart of the method.

Giorgi (2012) stressed the importance for the researcher, as a social scientist, to adopt the appropriate disciplinary attitude within the context of the phenomenological method. In addition to bracketing, the researcher is actively engaged in eidetic reduction

by first decontextualizing the output from each interview into meaning units and then categories. Giorgi said the researcher then applies imaginative variation to reconstruct the data into the essential invariant collective structures of the participants' experience. Bevan (2014) cautioned that researchers must remain vigilant and self-aware of their own natural attitude relative to the phenomenon under study.

In addition to being a PhD candidate, I am also a supply chain professional with over 30 years of experience in the consumer goods and industrial verticals. I am the Chief Supply Chain Officer for the company where I work. I spoke at or attended many conferences targeting supply chain professionals so it was likely that I might have had some type of relationship with some of the potential respondents in my study. I made sure that I was transparent about my experience and any relationships I might have had with other senior leaders that work in the same company as the potential participants in the study. During the interview process I followed the advice of Bevan (2014) and put my own self-interests aside while remaining focused on the person I was interviewing. In the analysis phase I followed the disciplined *structured text condensation approach* recommended by Malterud (2012) along with the bracketing perspective so critical to Giorgi's (2012) descriptive phenomenological approach.

Methodology

Participant Selection Logic

Englander (2012) explained that one key difference between quantitative and qualitative research as it relates to the selection of participants is that the quantitative researcher is seeking to select representative participants to ensure the generalizability of

the study. Therefore, the participant selection process must make sure that each subject belongs to the study population. Englander stated that with the phenomenological approach, the researcher does not have any insight into representativeness until later in the analysis phase when the general framework of the phenomenon is developed. Englander added that it is imperative that the phenomenological researcher finds subjects that have experienced the phenomenon under study.

Elo et al. (2014) stated that a purposive sampling strategy is generally used when the researcher wants to select participants with the best knowledge about the phenomenon being studied. The sampling universe for the study included senior supply chain managers from companies identified as having leading supply chain organizations. There were a number of surveys designed to identify the top global performing supply chain organizations. The two most often cited in the trade literature were the Gartner Supply Chain Top 25 (Aronow, Burkett, Nilles, & Romano, 2016), which has been published every year for the past 12 years, and the Supply Chains to Admire study (Cecere et al., 2016), in its third year of existence. Although there was minor overlap with some companies appearing on both lists, the methodologies for the two studies are different. The Gartner study compared all companies against each other without consideration of industry. The selection heuristic for the Gartner study included 50% weighting on financial performance, 25% on Gartner analyst opinion, and 25% on peer-company member opinion. The companies in the Cecere et al. study, were initially selected using financial metrics and then further winnowed down based on demonstrating whether or

not those organizations were driving improved value over time against their peer group as measured by a calculated supply chain index developed by the authors.

I took the purposive sample from supply chain leaders in organizations listed in the Gartner Supply Chain Top 25 (Aronow et al., 2016) or from the 16 companies identified in the Supply Chains to Admire study as *Winners* (Cecere et al., 2016). Potential research participants were from the pool of supply chain leaders with titles such as vice president, director, or senior manager with the expectation that people with these titles will have played a role in the decision-making process for choosing key performance measures and targets. Although Patton (2002) acknowledged there are no hard and fast rules for sample size in qualitative research, others (e.g., Englander, 2012) have recommended a minimum sample size of 20 for phenomenological studies. I planned for 25 respondents from at least five companies to allow for the likelihood that some organizations and/or respondents might not complete the study. Appendix A provides the letter of cooperation signed by the participating companies and Appendix B is the letter of invitation I sent to participants.

Instrumentation

According to Englander (2012), the interview is the predominant data collection protocol used in phenomenological research to draw out the rich lived-experiences and meanings of the phenomenon under study. Englander added that once the researcher has developed a pool of individuals who have experienced the phenomenon of interest the next step is to interview each participant. Brinkmann and Kvale (2015) recommended the

use of a semi-structured interview protocol with enough structure to guide the interview but not so rigid as to influence the participants' responses.

Giorgi (2009) emphasized that the questions contained in the interview protocol must be in alignment with the descriptive phenomenological approach. Giorgi shared that the interviewer asks the participant for a description of a situation in which the participant experienced the phenomenon under study. Giorgi stressed the importance of making salient not only the experience and associated meanings but also the context within which the phenomenon was experienced. Giorgi explained that meanings are critical to understanding the phenomenon and meanings are context-dependent.

Englander (2012) said that the first interview question should be in the format of, "Can you please describe, in as much detail as possible, a situation in which you experienced [name the phenomenon] (p. 26)?" Englander added that additional questions should be guided by the response of the participant with the researcher keeping the focus of the interview on the phenomenon being studied. Englander explained that in the phenomenological method, the researcher is less focused on the subject–subject relationship (participant–researcher relationship) and more focused on the subject–phenomenon relationship (participant–phenomenon relationship).

I developed the interview protocol for the study (Appendix C) based on the guidance of Jacob and Furgerson (2012) who advised including both the interview questions and a procedural script to get the interview off to a good start and to bring the interview to a successful conclusion. Jacob and Furgerson suggested including in the beginning of the script prompts to (a) tell a bit about myself, (b) provide an overview of

the study, (c) give an explanation of informed consent, and (d) provide wording to alleviate any confidentiality concerns. Jacob and Furgerson recommended concluding the script with prompts to (a) thank each respondent, (b) provide my contact information, (c) describe the next steps in the research process, (d) schedule a follow up interview as a placeholder to ask any additional and/or clarifying questions, and (e) ask participants if they would like a summary of the research findings.

Jacob and Furgerson (2012) recommended that the questions in the interview are open-ended and expansive. Cohen, Kahn, and Steeves (2000) said that interviews are interactive, resembling conversations. Jacob and Furgerson suggested that the researcher begin with easier questions and move to more complex questions. Jacobs and Furgerson advised incorporating bulleted words or phrases under each question to act as prompts both to ensure the gathering of rich data and to get the interview back on track if needed.

Jacob and Furgerson (2012) pointed out that one of the benefits of qualitative research is the emergent nature of the method and that the researcher must be willing to make in-situ revisions to the protocol. Piloting the interview protocol was strongly recommended by Jacob and Furgerson as a way to help strengthen both the script and the interview questions. I conducted a pilot study with two participants and used feedback from the pilot study to further refine the interview protocol.

Pilot Study

According to Kim (2011), a pilot study is used as a way to test the research methodology the researcher intends to apply to the main study but conducted on a smaller scale. Kim noted that there are several benefits to conducting a pilot study including

validation of the research protocol, the data collection methods, and the recruitment strategy. Cohen et al. (2000) suggested using a pilot study to help the researcher develop appropriate interview prompts to ensure that informants fully describe their experience. These prompts could then be incorporated into the protocol for the main study. Kim stated that a pilot study could validate that the research methodology is aligned to the overarching research question. Kim added that undertaking a pilot study is invaluable to a novice researcher in assessing readiness, skills, and commitment to a phenomenological study.

I reflected on the main concerns I had about the study. I used these concerns to inform the goals of the pilot study. I wanted to test the recruitment strategy. I wanted to understand if the interview questions acted as appropriate prompts for participants in exploring how they chose the key performance measures and goals for their organizations. I wanted to be sure that I was able to follow the interview protocol. I wanted to validate that following the protocol enabled me to create an environment where I captured the deep and rich experiences of supply chain leaders as they openly shared their decision-making strategies in their own words. And I wanted to test the data collection plan. To test the recruitment strategy, interview protocol, and data collection plan, I followed the process outlined in the Recruitment and Data Collection Process Diagram (Appendix D) and described in the section titled Procedures for Recruitment, Participation, and Data collection. The only deviations were that I asked the participants of the pilot study the following additional questions:

- What did you like about the recruiting process?

- What steps could I take to improve the recruiting process?

- Did you feel that the questions allowed you to share your unique experiences in choosing the key performance measures and goals for your organization?

- Was there anything that you thought I should have asked but did not?

- Do you have any other feedback about the recruitment and interviewing process or content?

Kim (2011) conducted a pilot study with two participants. Upon reviewing the participants' responses, Kim was able to see researcher biases and preconceptions regarding the phenomenon showing up in the interview questions. Kim realized the questions needed to be broadened for the main study to allow the participants to share their own experiences. Cohen et al. (2000) cautioned the researcher not to allow the prompts to direct the interview in an overly restrictive way. I too reflected on the pilot study to determine how I should have modified the recruitment strategy, the interview protocol, and/or the approach as a phenomenological researcher. I used these insights to inform the main study.

Procedures for Recruitment, Participation, and Data Collection

Englander (2012) stated that the participant is important to the research process because the participant has actually experienced the phenomenon under study. To draw out the rich details of the experience, Englander explained that the interviewer must remain present and continuously shift focus between the phenomenon and the participant. Englander stressed that the researcher must remain flexible as to the line of questions while still keeping focus first and foremost on the phenomenon. Cohen et al. (2000)

stated that the follow up interview is useful not only for reviewing the transcript, but as an opportunity for an informant, who might have reflected on the interview, to offer additional thoughts in the second interview that could lead to gathering additional data.

Cohen et al. (2000) described the rich layers of data that could be collected not only from the participant's interview but also via field notes taken by the researcher. Cohen et al. stated that field notes could also increase the trustworthiness of the study by making explicit any researcher bias as well as making salient decisions made during the study and the logic behind those decisions. Cohen et al. added that field notes are useful for documenting elements of an interview that would not be detected in a transcript such as intonation, body language, and other aspects of the environment. Cohen et al. said field notes allow the researcher to continue to document the details of an interview once the recorder has been turned off, such as during a follow up interview. Cohen et al. suggested that field notes be used as a way for the researcher to self-evaluate and to reflect on the interview. Cohen et al. added that field notes could be used to record ideas, clues, hunches, and insights related to the unfolding research.

In addition to the use of a hard copy of the interview protocol, I also recorded the interview using a recording and transcription service (NoNotes.com, 2017). NoNotes allowed a verbatim recording of questions to the participant and the participant's responses. A transcription of the interview as a Microsoft Word Document was available from NoNotes. I also planned to record the interview using an iPhone as a backup in case something happened with the NoNotes recorded interview.

Brinkmann and Kvale (2015) observed that there is little information in the research literature regarding the quality of the transcription process. Brinkman and Kvale pointed out that to transcribe means to transform and that the transcription process by its very nature begins to decontextualize the conversation that began with the live interview. Brinkman and Kvale stated that the first loss of context occurs during the audio recording in which body language disappears. There is further degradation during the transcription with the loss of intonation, voice, and breathing. Brinkman and Kvale's observations support the position of Cohen et al. (2000) regarding the value of field notes to capture contextual details of the phenomenon as a key element of the research process.

Brinkmann and Kvale (2015) shared that in many research projects the transcription is done by an administrative assistant. Brinkman and Kvale stated that if using an outside transcription resource it is important that the researcher is clear regarding the transcription rules by choosing between verbatim transcription that includes the participant peppering the interview with ohs, uhms, and pauses or a transcription written in a more formal grammatical style. Although using additional resources for transcription can help speed up the pace of the research, Brinkman and Kvale shared that researchers who do their own transcription can learn a lot about the effectiveness of the interview style. I used NoNotes outside transcription service so that I could continue conducting interviews while waiting for the raw transcripts from previous interviews. I asked NoNotes to transcribe the interviews verbatim.

Cohen et al. (2000) stated that it is imperative that the researcher document field notes as soon as feasible after the interview. One suggestion offered by Cohen et al. was

to record field notes immediately and then transcribe and elaborate as soon as possible thereafter. I reserved time at the conclusion of each interview to reflect on the interview and record my field notes.

I used the output from the pilot study to make needed modifications to the recruitment, participation, and data collection process. Since I expected to receive additional potential participants as an output from the pilot study, I began the recruitment and data collection from this potential pool. I outlined the recruitment, participation, and data collection process in the Recruitment and Data Collection Process Diagram (Appendix D), summarized as follows.

1. Select point-of-contact and obtain potential participants from that organization.

2. Enlist participants, obtain consent, and schedule initial interview.

3. Conduct and record initial interview. Schedule follow up member-checking interview.

4. Receive raw interview transcript from NoNotes.

5. Sanitize transcript, send to participant, and conduct member-check.

6. Create concept map from final member-checked transcript.

7. Repeat steps above until I have collected 25 edited transcripts and 25 concept maps.

Data Analysis Plan

Giorgi's descriptive analysis methodology. Since I conducted a descriptive phenomenological study, I followed the descriptive analysis (DA) process as outlined by Giorgi (2009) who is considered a seminal researcher of this methodology. Giorgi said

that the purpose of DA is to understand the meaning of the phenomenon based solely on the data presented by the informant. Giorgi reminded the researcher that DA does not try to clarify any ambiguity in the data, merely to describe the ambiguity. Any data gaps are not filled in via theoretical speculation but rather through collecting more data. Giorgi said that the ultimate result of DA is the generation of a second-order decontextualized description of the phenomenon.

Giorgi (2012) stated that there are five steps in the DA method. In the first step, the researcher reads the entire transcript to get a sense of the whole experience in the words of the informant. In the second step the researcher begins to determine the meaning units contained within the text. Giorgi (2009) advised the researcher to read through the interview again, making a notation on the transcript each time the researcher notices a shift in meaning. Giorgi stated that determination of a meaning unit is somewhat arbitrary and different researchers reading the same transcript will develop different meaning units.

Giorgi (2009) added that it is the transformation of the meaning units that is at the heart of the DA method and is the third step in the analysis. Giorgi shared that in the transformation step, the informants' first-order lived expressions are transformed into second-order thematic statements. Giorgi said that the researcher again returns to the description that has been broken down into meaning units. The researcher then interrogates each meaning unit to draw out the invariant essences in order to develop higher level categories that contain the same psychological meaning. Giorgi added that this transformation is the first step in the development of a general framework. The

framework enables the integration of data from multiple informants into a single structure that, in spite of specific differences in the experience, shares a common psychological meaning.

Giorgi (2009) cautioned the researcher that step 3 is very time consuming. The researcher must become immersed and spend time with the data in an imaginative way including looking at the data from a polar opposite perspective. Giorgi added that the researcher might need to re-write many versions of the transformation before finding the one most suitable expression. The researcher must then test the invariant expression by returning to the meaning units and validating that the category description in fact captures all aspects of the data.

Giorgi (2009) advised the researcher to take on the perspective of the *critical other* that sits on the researcher's shoulder to ensure that the transformation makes sense to the broader community of stakeholders in the research. Giorgi added that researchers must make explicit the process they are going through while transforming the meaning units. Brinkmann and Kvale (2015) summarized the output of the third step as a set of transformed meaning units written as simple thematic statements that convey the informants' viewpoint as understood by the researcher.

Giorgi (2009) stated that in the fourth step of the DA, the researcher again interrogates the meaning units, this time from the lens of the purpose of the study. Since the study is to explore the decision-making strategies used by supply chain leaders in choosing their key performance measures and goals, in step 4 of the analysis I interrogated the meaning units through the general lens of decision-making theory and

the more specific lenses of rational decision-making, heuristics and biases decision-making, and naturalistic decision-making.

Brinkmann and Kvale (2015) stated that there are no predetermined number of transformations required of steps 3 and 4. Giorgi (2009) said that first the researcher must re-write the meaning units exactly as transcribed but this time in the third-person voice. Giorgi said the researcher must document in a workbook or journal how the transformations are worked through. It is through the transformation process that the researcher adds psychological knowledge. Giorgi added that the output from steps 3 and 4 will yield three to five general themes that will become the basis for writing the general structure of the phenomenon.

Giorgi (2012) stated that step 5 is writing the final structure. The final structure generally consists of several general constituent meanings along with the relationships among the meanings. Giorgi recommended that the researcher pressure tests the structure by removing a key constituent to see if the structure remains or if it collapses as a result of the removal. Giorgi said that the structure could include characteristics of the experience that the informants made salient, as well as aspects of which the informants were either unaware or chose not to directly share. Giorgi explained that during the study the informants will share what each of them lived relative to the phenomenon although the researcher must focus on how the informants lived that experience. Ideally the resulting structure is general enough that it would be applicable to people beyond the sample set in the study. Brinkmann and Kvale (2015) neatly summed up step 5 as the

process of tying together the essential unique themes from all the interviews in a simple yet comprehensive descriptive statement.

Giorgi (2009) stated that by following the DA method, the researcher has taken a complex set of informant interviews, broken the interviews down into a set of natural meaning units, made salient the main themes, subjected these themes to extensive interpretation and theoretical analysis, and synthesized the informants' common experience into a general descriptive structure of the phenomenon that is independent of the context. Giorgi stressed that the description of the structure is not the final step of the process. Since the structure is intended to provide deeper insights into the phenomenon, Giorgi coached the researcher to return to the transcripts and again interrogate the data, seeking to better understand the variations within and between the informants' responses. Giorgi encouraged the researcher to provide specific examples of the essential insights in order to provide color and clarity to the more abstract elements of the structure.

Brinkmann and Kvale (2015) strongly recommended that the researcher keep the final report top of mind, including how to present the results of the analysis, during all phases of the study. I followed the suggestions of Giorgi (2009) that in addition to the descriptive structure, the researcher consider creating a visual diagram showing the relationships among the constituents of the phenomenon as well as a table showing how each informant experienced each of the key aspects of the final structure.

Coding. Saldaña (2013) explained that coding is the assignment of a word or short phrase to a meaning unit that captures the essence of that piece of data. Saldaña added that coding allows the researcher to transition from data collection to data analysis.

Coding allows the researcher to find patterns in the data that show (a) similarity, (b) dissimilarity, (c) frequency of occurrence, (d) sequences, (e) cause and effect, and (f) occurrences relative to other activities or events. Saldaña observed that coding is an iterative process as the researcher applies varies analytic lenses to the data. Coding goes through two or more cycles. Saldaña added that as recoding progresses the output moves from being more specific to more general. Saldaña noted that in each subsequent cycle the researcher organizes, groups, focuses, and highlights the data in such a way that categories, themes, and sometimes theory, emerge from the process.

Saldaña (2013) said that in addition to coding the informants' transcripts, the researcher should also consider coding field notes along with any other documents or artifacts collected along the way. Saldaña recommended that coding occur as data is collected. Saldaña shared that during the coding process certain ideas and patterns may begin to emerge. Saldaña recommended the writing of analytic memos, defined as thoughts that occur during data gathering and analysis. Saldaña described analytic memos as a way of recording the mental dialogue the researcher has while going through the process of making connections in the data, raising questions, solving problems, building strategy, as well as general reflection. Saldaña added that analytic memos are fertile ground for coding. Saldaña advised the novice researcher to code anything and everything so as not to miss what later could be determined as a crucial piece in the research puzzle.

Saldaña (2013) offered a variety of coding methodologies for First Cycle Coding. Saldaña added that the choice of coding methodology depends on the nature of the

qualitative study. Saldaña suggested that novice researchers consider basic coding techniques such as Attribute Coding, Structural Coding, and/or Descriptive Coding for First Cycle Coding. Since I conducted a descriptive phenomenological study it seemed logical to apply Attribute Coding to allow for the capture of descriptive information including particular characteristics and demographics. For example I might have found that there were different decision-making strategies employed by supply chain leaders depending on gender or time in role factors.

Saldaña (2013) explained that Structural Coding is broadly used across qualitative studies that employ semi-structured interview protocols across multiple participants in order to examine commonalities and relationships in the body of data. Since the study was intended to explore whether there are similarities and differences in decision-making strategies across supply chain leaders, Structural Coding seemed applicable. Saldaña also recommended Descriptive Coding as particularly appropriate for novice researchers across all qualitative methodologies. Descriptive Coding results in the categorization of the data corpus, is particularly applicable to the analysis of field notes, and can lay the foundation for Second Cycle coding. Since the conceptual framework for the study was decision-making theory, I used pre-determined broad codes that followed the three main decision-making schools of thought; rational decision-making, heuristics and biases based decision-making, and naturalistic decision-making as a way to initially categorize respondents' transcripts. Saldana also cautioned the researcher not to get too wedded initially to a coding process.

Saldaña (2013) suggested that novice researchers consider using Eclectic Coding in which the researcher first applies codes to the data without regard for a specific coding technique. Then the researcher identifies each code to a coding methodology. At the conclusion of the First Cycle coding the researcher may find that one or two coding techniques have been used for the majority of the data which could then inform the Second Cycle coding technique(s) to be applied to re-coding the data in subsequent coding cycles.

Saldaña (2013) recommended trying out coding techniques on a smaller sample size, such as in a pilot study. Saldaña added that the researcher could then assess the results to determine the most suitable coding techniques for the study. I followed Saldaña's suggestions of using Eclectic Coding (which may include Attribute, Structural, and/or Descriptive Coding) to the pilot study to inform the approach to the main study.

Saldaña (2013) explained that moving from First Cycle to Second Cycle coding should be considered a process of re-coding. Saldaña said that as a result of First Cycle coding and simultaneous writing of analytic memos, the researcher begins to see insights, discoveries, and patterns emerging from the data. Saldaña added that these outputs help frame the deep thinking required by the researcher to determine how best to re-cycle through the data to provide additional focus and deeper insights, connections, and processes. Saldaña summarized the goal of the Second Cycle stage as the development of categories, themes, concepts, and theories. In each cycle, the number of codes decreases as the researcher develops broader categories, concepts, theories, and statements.

Two Second Cycle coding techniques offered by Saldaña (2013) that seemed relevant for the study were Pattern Coding and Focused Coding. Saldaña stated Pattern Coding is particularly useful for creating meta-categories that group together the codes generated from the First Cycle. Saldaña added that Super Coding is a variation of Pattern Coding that aids researchers in finding relationships among or between codes. Saldaña explained that Pattern Codes are often expressed as metaphors.

According to Saldaña (2013), Focused Coding is appropriate for essentially all qualitative studies and usually follows Initial Coding or Process Coding. Focused Coding allows for the generation of higher-level categories with subcategories beneath each category in a tree-like structure. Although Pattern Coding and Focus Coding seemed to be potential coding methodologies for the study, I followed the advice of Saldaña and let the results of the First Cycle coding influence the Second Cycle coding methodologies that I applied.

Software. Saldaña (2013) stated that there are a variety of tools that the researcher could use to code and analyze qualitative data. Saldaña recommended that the researcher initially consider manually coding small-scale studies on hard copy printouts. I followed this recommendation for the pilot study interviews. Saldaña also recommended that those who will be handling large volumes of data as a result of multiple participant interviews consider using computer-assisted qualitative data analysis software (CAQDAS) to help with data storage, organization, management, and initial analysis. I used NVivo.

I created concept maps of each informant's interview. To do so I used Visio Pro (Microsoft, 2017) because of its ability to visually organize complex concepts. I also created a concept map of the final descriptive structure to show pictorially the elements of the structure and the interrelationships among those elements.

Discrepant cases. There might have been places in the process where I could have encountered what might seem to be a discrepant case. Brinkmann and Kvale (2015) cautioned the researcher to remain diligent in the potential for subjective bias to sneak into a study. Brinkmann and Kvale stated that a sloppy researcher is one that only notices evidence that supports the researcher's opinions or preconceived notions failing to note any counter evidence. It was imperative that I determine if a case truly was an anomaly or if I missed something in the analysis. Brinkmann and Kvale suggested that when a discrepant cases is encountered this could be a trigger to interrogate the data again using a different perspective and asking different questions of the same data in order to explicate different interpretations of meaning. Should a case be truly unique and different, I attempted to make this condition transparent as I described how I attempted to transform the informant's text. By applying perspectival subjectivity to interrogation of the text throughout the process as well as providing transparency, I improved both the rigor and the richness of the study.

Issues of Trustworthiness

Lincoln and Guba (1985), seminal authors on qualitative inquiry, stated that it is difficult to ensure that qualitative research is correct. They proposed four criteria to ensure the trustworthiness of qualitative research: *credibility, transferability,*

dependability, and *confirmability*. Giorgi (2012) stated that these four criteria must be applied consistently with respect to the ontological, epistemological, and methodological framework of the study. Sousa (2014) added that although there exist consistency and validity elements that are common to all qualitative methodologies, the researcher must apply quality control measures that are specific to the chosen research method.

Credibility

According to Sousa (2014), credibility establishes that the research results are believable. Sousa offered some methods to ensure credibility including triangulation, prolonged contact with participants, peer debriefing, reflexivity, saturation, and member-checking. I planned to apply member-checking, concept maps, saturation, reflexivity, and peer debriefing because these methods are recommended by Giorgi (2012) for phenomenological studies.

Houghton, Casey, Shaw, and Murphy (2013) described member-checking as allowing the participant to read a verbatim transcription of the interview and to acknowledge that the participant's words were used. Houghton et al. highlighted a potential challenge in phenomenological research. Houghton et al. said that if participants are asked for feedback on the data after the researcher has synthesized and reconstructed the results participants likely will not be able to recognize their own responses once analysis and eidetic reduction has occurred. Sousa (2014) agreed with Houghton's et al. perspective on member-checking in descriptive phenomenological studies stating that the researcher may readdress the participants for clarification or for more detail but this must occur *before* beginning the analysis. As previously described in the section on

recruitment, participant selection, and data collection, I e-mailed each participant a copy of their transcribed interview along with a letter asking each participant to either confirm that the transcript reflected their experience in their own words or to provide changes. I also conducted one or more follow-up calls to ensure that each participant was in agreement that I adequately captured how they experienced the phenomenon.

I followed the advice of Shosha (2012) and, as a proxy for peer-debriefing, reviewed the data analysis process with my committee to ensure that I applied a rigorous approach to turning raw data into meaning units, themes, and categories. Shosha recommended the use of a journal to help aid in bracketing. A journal allowed me to make transparent any pre-conceived biases I had and to put these biases aside. I also followed the advice of Saldaña (2013), as described in the subsection on coding, and adopted the practice of writing analytics memos during the data analysis process. The journaling and use of analytic memos acted as an audit trail to describe any key decisions I made along the data analysis journey.

According to Sousa (2014), saturation is another critical component to ensuring credibility of a study. The plan was to initially interview five senior supply chain leaders who agreed to participate in the study. I took the result of those interviews to conduct the initial coding. I conducted the next five interviews, adding new data to the coding. I used three criteria to determine if saturation has occurred: (a) I was obtaining no new information, (b) I was adding no new coding, and (c) I had gathered enough information that would allow someone else to replicate the study. If I did not meet one of the above three criteria, I planned to continue to gather respondents until all three criteria were met

or I met the Walden University minimum of 20 respondents. Once I met the criteria I stopped recruiting new participants.

Transferability

Houghton et al. (2013) stated that transferability determines if the research results can be transferred to another similar situation without altering the meanings from the original study. Transferability is determined by the reader of the research using the thick descriptions of the specific context, the protocol, and the raw data provided by the researcher of the original study. Elo et al. (2014) stressed the importance of reporting high quality results and applying a thorough analysis process. Elo et al. added it is important to include clear descriptions of the characteristics of the participants along with the context within which the participants experienced the phenomenon to aid the reader in evaluating the transferability of the study.

Sousa (2014) stated that *analytical generalizability* is achieved from applying both internal and external consistency methods. Sousa said internal consistency includes constancy of process in participant selection, interview questions, data collection, and data analysis. Sousa added that external consistency includes the linkage between the application of concepts/theory, theory generation, and writing the resulting report. To aid in the transferability of the study, I provided detailed descriptions of the context and examples of raw data to make salient possible alternate interpretations of the results. I also provided direct quotes from the participants to add to the richness of the interpretations of the phenomenon. The selection of the participants also added to the

transferability of the study because I collected the experiences of supply chain leaders from several companies within a variety of industries.

Dependability

According to Houghton et al. (2013), dependability ensures that the data are gathered in a repeatable and reliable manner whereas confirmability demonstrates how the research findings were developed from the data. Houghton et al. recommended the use of an audit trail and a reflexivity journal, as important tools to ensure the study results are dependable and confirmable. The audit trail for the study came from two sources. I kept a journal of detailed notes describing the contextual background of the data collection process and I used analytic memos to capture the reasoning and rationale for all methodological decisions related to applying the descriptive phenomenological method as described by Giorgi (2012). The second source was from the NVivo software I used. Houghton et al. stated that NVivo could provide a complete trail of decisions made from data collection through to analysis. For example, the use of the NVivo query functionality may help to ensure a researcher does not place undue emphasis on minor findings thus helping to reduce researcher bias (Houghton et al., 2013).

Confirmability

Berger (2015) stated that reflexivity is the both the recognition of and the continual internal self-evaluation made that the researcher's position can have in affecting the results of a study. Berger added that factors that influence the researcher's positioning might include gender, age, ethnicity, professional experience, political orientation, etc. Berger shared three major areas that may be impacted by the researcher's

posture and that must be addressed via a reflexive approach to study: (a) the ability of the researcher to access and engage the participants, (b) the impact of the researcher-participant relationship, and (c) the worldview of the researcher and the impact of that worldview on how questions are posed and results are interpreted.

Houghton et al., (2013) recommended the use of a reflective diary where the researcher could capture what brought the researcher to study the phenomenon and how that perspective might affect the research. Saldaña (2013) recommended the use of analytic memos to make design decision rationale, emotions, and challenges transparent. Berger described reflexivity as self-supervision and suggested the research employ strategies such as member-checking, creating an audit trail, and the use of journaling. I used a self-supervision journal and analytic memos to capture the reasoning, decisions, and emotional reactions related to the study.

Ethical Procedures

I gained approval from Walden University's Institutional Review Board (IRB) to proceed with the research. Securing IRB approval meant that I obtained a certificate from the National Institute of Health showing I completed the Human Research Protection Training. According to Harris (2017), a researcher must demonstrate to the IRB that beneficence, justice, and respect for persons has been adequately addressed to receive IRB approval. Beneficence is described by Walden University as maximizing the possible benefits of the study while minimizing possible harm to the study participants. Justice includes fairly distributing the benefits and burdens of a study. And respect for persons means that the researcher acknowledges that the participants are able to

participate or withdraw of their own free will and the researcher protects any participants who have diminished autonomy. If I had found during the pilot study that I needed to change any of the participant recruitment, data collection, or data analysis procedures, I would have filed an IRB Change Request form and waited for approval before making the changes and continuing on with the study.

To make sure participants understood the study, what their involvement was to be, how they could withdraw, how they would receive results, and that they would not be compensated for participating in the study I made sure that I obtained an electronically signed copy of the consent form for each participant in accordance with Walden University (Harris, 2017) IRB requirements. Since the study was of a general business management and not a personal nature there was little likelihood of personal risk to participants. I ensured that a senior person in the organization approved employees participating in the study and also that there was no pressure to participate. I ensured that participants knew they could withdraw from the study at any time. If someone did had withdrawn I would have deleted all of the data related to that participant and recorded on the participant id number cross-reference file that the participant withdrew.

I previously described in detail in the methodology section the steps I took to protect the identity of each participant and their organization. These steps included use of a participant id number cross-reference file, sanitizing the raw transcriptions to remove any participant or organization information, deleting raw transcription data once the sanitized transcripts had been approved by the participant, saving each electronic document as a pass-word protected file on an external hard drive and saving each hard

copy file in a locked safe for a period of at least five years, in accordance with Walden University IRB policy. After five years I will delete the electronic files and shred the hard copy files.

Summary

In Chapter 3 I presented the approach for selecting the descriptive phenomenological methodology as the most appropriate approach for addressing the research question which was approved by the Walden University IRB (IRB approval number 01-29-18-0262213). I explained the role of the researcher in a descriptive phenomenological study. I drew from the literature to develop the methodology and interview protocol. I described in detail how I planned to conduct the pilot study and how I planned to use what I learned from the pilot study to inform the main study. I described how I planned to select the participants for the study, how I planned to conduct the data collection process along with a description of how I planned to apply Giorgi's (2009) descriptive analysis process for data analysis. I shared the technology I expected to use to gather and analyze the data from the participant interviews. I explained how I planned to ensure that the study was trustworthy, that I had a solid ethical procedures plan to protect both respondents and their organizations and that I planned to meet the requirements of the Walden University Internal Review Board. I believe I provided the necessary detail so that another researcher could replicate the study.

Chapter 4: Results

The purpose of this qualitative descriptive phenomenological study was to improve the understanding of the decision-making strategies used by supply chain leaders in their selection of key performance measures and associated goals. The central research question for the study was, What are the decision-making strategies key performance measures and goals for their organizations? In addition to the central research question, I developed seven interview questions for the interview protocol (Appendix C).

In the remainder of this chapter I share the impact of the pilot study on the main study. I explain how aspects of the research setting may have influenced participants' responses or interpretation of study results. After presenting participant demographics and characteristics, I describe the data collection process including changes in the data collection plan versus what I outlined in the proposal. I share the process I used for data analysis including how I moved from the original transcripts to meaning units, to categories, and finally to themes. Then, I explain the coding logic I used to ultimately develop the major themes. To demonstrate that the study is trustworthy, I present evidence of credibility, transferability, dependability, and confirmability.

Then, I present the results of the study. I share the analysis in response to the central research question by presenting the overarching themes along with specific examples in the words of the participants that emerged through the data analysis. I close by summarizing the chapter. In Chapter 5 I present the interpretation of the results.

Pilot Study

I conducted a pilot study to test the research methodology I decided to employ and more specifically the interview protocol. I also wanted to validate the data collection and member-checking process as outlined in Chapter 3. I chose the first two participants who agreed to participate in the main study for the pilot study. In addition to the questions and prompts for the main study, I asked the pilot study participants the following questions:

- Did you feel that the questions I asked you allowed you to share your unique experiences? In essence did I get from you what you would have wanted to convey in an interview on this subject?

- What did you like about the process I used to recruit you to the study? What did you dislike about the process?

- Is there any other feedback about either the recruiting or the interviewing process or content that you want to make sure I know about?

The first participant (P1) stated that the initial question, "What is the *single* most important key performance measure or indicator for your supply chain organization?," made P1 feel a bit boxed in. P1 said, "You cornered me a little bit. I think I would encourage you to broaden that topic. I think I could've shared with you a little bit more on things we're measuring." P1 added that being allowed to share the process on two or three KPIs might have enabled more information sharing. I thanked P1 for the feedback and said the intent was not to constrain thinking or responses in any way. During the interview I added an ad hoc seventh question, "Was there anything I should have asked

and you wanted to be sure that I know?" P1 responded that I did not make P1 feel

uncomfortable at all stating, "You went through what I would've asked, actually." P1

also added, "You left enough room for me to maneuver. You didn't take me in any

specific direction, not at all."

For the second pilot interview (P2), I slightly modified the first question by

removing the word *single* from the question, "What is the most important key

performance measure or indicator for your supply chain organization?" At the end of the

interview I asked P2 the same additional pilot interview questions, and I received

favorable responses both on the recruiting process and the interview ("I'm a willing

participant. If I can share information and help with a research project and hopefully get

information myself . . . then that collaboration is always beneficial."). I asked if the

questions allowed P2 to share unique experiences around KPIs and goals; P2 responded,

"Thank you, the questions helped to facilitate linking those dots (between process

documentation and establishing KPIs)." P2 was also complimentary of the sampling

strategy, calling it a "Very good methodical, valid, and reproducible process."

The pilot study confirmed that the interview protocol (with the modifications

noted) seemed to allow the pilot participants to share their unique experiences regarding

the selection of KPIs and goals for their respective organizations. I used NoNotes.com to

transcribe the recorded phone interviews. I then shared the sanitized transcripts with P1

and P2. Working iteratively with each participant, I was able to generate sanitized

transcripts that the two pilot study participants confirmed represented their respective

experiences with the topic. This showed that the proposed member-checking process was solid.

After loading each sanitized transcript into the NVivo12 qualitative research software, I read through each one and created a memo to capture impressions of each interview and resulting transcript. I found that I had collected a substantial amount of rich raw data from P1 and P2 that, along with the conceptual framework from the literature review in Chapter 2, formed the basis for the initial codebook structure. I first manually developed the meaning units from each transcript by underlining key words and/or phrases with a pencil. Then, I coded these meaning units into NVivo12. Refinement of the codes, nodes, and subnodes continued throughout the remainder of the data collection and analysis phases of the main study.

Research Setting

I conducted all interviews via iPhone. They were recorded using the NoNotes.com recording app. In advance of the interviews, I e-mailed participants the general questions that would be asked. All but three participants had read the questions in advance and were prepared with their responses. The three participants who had not prepared apologized and said this was due to their busy schedules. I assured them that was fine and what I was really interested in was hearing about their experiences with supply chain measures and goals.

All interviews were conducted during normal business hours. All but one participant took the interview from their offices. The one participant did not work from

home on certain days and chose to schedule our interview for one of those days. In no case did I sense that any participant was rushed for time or under any other form of stress.

After I accessed each raw transcript from the NoNotes.com server, I converted it to a sanitized transcript following the process outlined in the proposal. There were two occasions where I had some technical/connectivity issues with NoNotes.com. I reinitiated the call, and the participants and I had a laugh about the challenges of technology. In those situations, I ended up with two recordings for one interview which I subsequently I combined but I kept in the transcript the dialogue around the fact that the call dropped and that we were able to pick up where we left off.

Although conducting the interviews via phone was efficient, as I had participants from all over the United States and one in the United Kingdom, I found that I missed having the face-to-face interaction. I tried to use intonation of the voice on the phone to help guide the interview, but I am sure that I missed some nuances. However, after listening to the recordings of the interviews, I do not feel that it had much, if any, effect on the quality of the interviews, process, or output.

Demographics

Three community partners, all recognized as having exemplar supply chains, agreed to participate in the study (see the sample letter of cooperation in Appendix A). From these companies, 15 senior supply chain leaders completed the interview and member-checking processes. I conducted the demographic analysis, using SPSS v.25 on the data I collected from these 15 participants during their respective interviews.

Of the 15 respondents who completed the interview process, 11 were male (73%), and four were female (27%). One of the selection criteria was that the participants had to hold senior roles in their organizations. Three of the participants were senior managers, three were directors, three were senior directors, five were vice presidents (VPs), and one was a senior vice president (SVP). Of the female respondents one was a SVP, one was a VP, one was a senior director, and one was a director. All 15 respondents qualified for the study based on their level in their respective organizations.

Another criterion for selection was that participants were either responsible or accountable for selecting the KPIs and measures for their respective organizations. Using the RACI model; Responsible, Accountable, Consulted, or Informed (Blokdjik, 2008), 14 of the participants stated that they were *responsible* for the key performance measures and goals for their respective areas. Eight participants said they were *accountable*, and one participant said they were also *informed*. All 15 participants were *accountable* and/or *responsible* for the KPIs and goals and were, therefore, qualified for the study based on their role in their organizations.

I computed descriptive statistics for two demographic variables: (a) years as a supply chain professional and (b) years in current role. The mean number of years as a supply chain professional for the respondents was 19.87 years ($SD = 9.441$). The mode was 21 years. The participants' years as a supply chain professional were from 4 to 35 years for a range of 31 years. The two participants with fewer than 10 years of supply chain experience were seasoned professionals coming from outside the supply chain organization. The mean for the number of years each participant had been in their current

role was 3.77 (SD = 1.915). The mode was 3 years with participants reporting from 1 to 7 years in their current role for a range of 6 years.

Data Collection

I followed the recruiting and data collection process as described in Chapter 3 and Appendix D. I was able to gain consent from and complete the data collection process from 15 senior supply chain leaders from the three consenting companies. Although this was less than the 20 participants I had originally targeted, I met the saturation criteria I had outlined in the proposal. The first criterion was that I was obtaining no new information. The first four interview questions were designed to get at the essence of the research question. After the first eight participants shared their KPIs and I categorized the KPIs as service, cost, quality, or corporate social responsibility (CSR), the next seven responses fit neatly into one of those categories. When asked about the decision-making processes used to select KPIs, by the tenth participant everyone had used naturalistic decision-making and this carried through to the fifteenth participant. And all but one participant (P13) had used two or more decision-making theories in their responses.

The second criterion for saturation was, I was adding no new coding. Because I applied structural coding using the conceptual framework of decision-making, inherent in the design was a form of saturation. The only questions where I used open coding were in asking how participants used their KPIs and what they would change in the design or use. Within the first eight participants I had created a coding structure that fit the replies of the seven remaining respondents.

The third criterion for saturation was that I had gathered enough information that would allow another researcher to replicate my study. This was the most subjective of the criteria. After writing the results, I re-read them through the lens of a future potential researcher who would want to determine if the study was trustworthy and whether there was enough detail in the research process as well as the participants' responses to conduct a similar study with another population. If I felt certain areas were lacking, I more fully developed those areas. In spite of conducting only 75% of the targeted number of interviews, I am confident that I reached saturation in the study.

After conducting, recording, and transcribing each phone interview using the NoNotes.com app, I sanitized the data as described in Chapter 3. Each participant then identified, either via a subsequent phone meeting or an e-mail, changes to be made to the sanitized transcript. Once all changes were made to the satisfaction of the participant, I received a verbal or e-mail confirmation that the sanitized interview reflected the participant's intended responses to the interview questions. Since there were several steps involved in (a) recruiting participants, (b) scheduling initial interviews, (c) conducting those interviews, (d) transcribing, (e) sanitizing, (f) scheduling, and (g) member-checking the sanitized transcripts with each participant, I created an Excel-based audit trail showing the status in the process for each participant (Appendix E).

Participants were very senior supply chain professionals from large companies and, therefore, had busy schedules. It took a long time to get an initial 45 to 60 minute interview with each participant and then to get another 30 minutes of time for the member-checking phone meeting during which I reviewed the sanitized transcript with

each participant. I scheduled the member-checking meeting at the conclusion of the initial interview. In several cases the member-checking interview was subsequently rescheduled due to conflicts that arose on some participants' schedules. The first interview (P1) was conducted on June 19, 2018 and the final sanitized member-checking call with P15 took place on September 13, 2019 for a total duration of 15 months. I loaded the raw data from the 15 sanitized interviews, as they were completed, into NVivo12 for subsequent qualitative analysis.

Data Analysis

Saldaña (2013) advised that the central research question will guide coding choices made by a researcher. Initially, I created a node for each interview question. I then updated the style of each interview question to *heading 1*. This allowed for use of the NVivo12 auto code functionality to gather all 15 participants' responses under each interview question to facilitate further coding.

For the first interview question (Q1) I asked the participants to name one of their most important KPIs and the second interview question (Q2) provided the opportunity to provide a definition. As I looked through the responses to these two questions I decided that the best approach to analyze these questions was to create a table listing the KPI names by participant number and then assign a general theme to each KPI to see if certain types of KPIs were more common across participants' responses.

The third (Q3) and fourth (Q4) interview questions were intended to address the central research question regarding the decision-making strategies used by supply chain leaders in choosing the KPIs and goals for their organizations. I decided that for initial

coding I would use structural coding around the three main schools of decision-making theory: (a) rational decision-making, (b) heuristics and biases, and (c) naturalistic decision-making. Saldaña (2013) stated that structural coding is applying a conceptual-based inquiry to data related to a specific research question. I created nodes for each main school of decision-making theory, the conceptual framework. Under each node I created subnodes that described attributes of that particular decision-making theory. For example, under the node *naturalistic decision-making,* I created subnodes (a) *mental simulation,* (b) *sequential option evaluation,* (c) *situation comprehension,* and (d) *situation recognition*. Following a deductive process, I used the conceptual framework codebook to code responses to Q3 (Please tell me, with as much detail as possible, about the process you used to decide on [insert name of KPI] as a key performance measure for your organization.) and Q4 (Have you determined a goal or target for [insert name of KPI]? If yes, then please describe, with as much detail as possible, how you came to decide on the goal of [insert goal].).

I began to analyze the responses to Q3 by highlighting key words and/or phrases of each participant's response to that question and coding those meaning units to the various nodes and subnodes of each decision-making theory. For example, P11 commented, "You need to stretch yourselves and imagine what it's like to be in the shoes of your customer . . . and improve that because companies that do that, do better in the marketplace and so that's why we're trying to do that." I coded this sentence to *mental simulation*, a subnode under the parent node *naturalistic decision-making*. Early on in the analysis, I decided to break out *fast and frugal heuristics* as a separate node from

heuristics and biases to capture the nuances between these two lines of study on the use of heuristics in decision-making. I moved the various subnodes to the appropriate parent node.

I was curious to see what decision-making theories each participant used. I was also curious how frequently each participant used language that could be associated with each decision-making theory. I used NVivo12's matrix query feature to address these questions. I selected the 15 cases as the rows and the decision-making theories as the columns, with each cell showing the frequency a decision-making theory was used. I exported this query to an Excel file that can be found in Appendix F.

I applied the same structural coding to the participants' responses to Q4 as I did to Q3; coding meaning units to the various subnodes and nodes of decision-making theory using NVivo12. For example, P10 responded, "I think at the moment the target is to be better than the best in the corporation. . . . the short-term target is to become better, the long-term target is to sustain on a higher level than the average corporate level." I coded this meaning unit to the subnode *anchor and adjust* that rolls up to the parent node *heuristics and biases decision-making*. Once I completed the structural coding for Q4 for all 15 participants, I used a similar process to analyze Q4 data as I did for Q3; using NVivo12 to create a matrix query showing by participant the number of coding references for each decision-making theory. I exported this query to an Excel file that can be found in Appendix G.

After asking participants about their KPIs, how they selected those KPIs, and how they set their targets, I asked participants (Q5) how the KPIs were used in their respective

organizations. As I read through the first several responses to Q5, I began to see the beginnings of a process flow emerging from the participants' language. I drew a flowchart using P1's responses:

> So in my organization specifically, we use that measurement and we take it all the way from our monthly reviews with business units. It's one of the key numbers we review with our carriers as well. So we take it to the carrier level, we take it to the [transportation] lane level, we use it to analyze high level plans . . . where [the] service level is going up or down . . . . We do go all the way down to comparing what the customer is telling us versus what our system is telling us . . . to have a constructive dialogue with the customer so we can breach the gap between reality and perception. Once that gap is closed, now we can start talking about putting a plan together.

I modified the initial chart incorporating P10's response:

> I compare it versus previous experience. So, I see how we evolved compared to the previous measurement and I compare that against the best in the organization. That's the first step . . . it allows us to say, "Yes, we will continue to do that," or we change direction.

I continued updating the flow chart by combining steps and refining the flow of information until I had developed a simple process flow that described not only how KPI results flowed through the organization but also how the organization used those results to drive improvements in performance. Once I developed the initial model, I reviewed the

model against responses from the other participants to make sure the model held together. For example, P13 explained that they use the KPIs in the functions and then those ladder up to the product value streams so I made sure the process model showed a review at various levels of the organization.

The sixth interview question (Q6) was intended to explore what each participant would change if they could with respect to the KPIs they chose. I read through each participant's responses and then performed open coding with the overarching theme of *anything you would change*. For example, P5 said that it is the availability of the data that has been the biggest road block. I coded that meaning unit to *get data faster and easier*.

During the analysis, I created a code *unintended behavioral consequences of KPI* and another code *understanding consequences of behavior on KPI*. The first node was language describing how the KPI changed people's behavior and the second node was the need to teach people how the work they do impacts the KPI results. I made these two nodes subnodes under the parent node *KPI affects behavior*. P11 commented that one barrier with OTIF (on time and in full) is that people were concerned if this metric would be the basis for their bonuses. And P7 asked, "So, how do we use these KPIs to push our work even harder?"

Additionally, I found that the nodes (a) *getting data faster and easier*, (b) *leading vs lagging*; and (c) *no change* could become subnodes under the parent node *better KPIs*. But I felt that the node *experimenting with new metrics* was unique enough to stand alone as a node. For example, P12 offered, with regards to their KPIs, that they were experimenting and trying to figure out how to scale the KPIs.

After I completed the coding, including creating nodes and subnodes, I ended up with four nodes: (a) *better KPIs*, (b) *experimenting with new metrics*, (c) *external view*, and (d) *KPI affects behavior*.

For the final interview question (Q7) I asked, "Was there anything I did not ask you but you wanted to be sure that I know?" I created a node called *what else* under which I captured participants' responses. Much like with Q6, I read through each response creating subnodes under *what else*. I initially had 10 nodes for 15 participants. After combining some nodes, I ended up with six subnodes: (a) *change culture*, (b) *external collaboration*, (c) *KPI structure*, (d) *organization structure*, and (e*) nothing to add*.

Evidence of Trustworthiness

In Chapter 3, I shared that Lincoln and Guba (1985) acknowledged the difficulty in making sure that qualitative research is correct. They proposed four criteria to ensure the trustworthiness of qualitative research: credibility, transferability, dependability, and confirmability. Sousa (2014) stated that credibility establishes that research results are believable. Houghton et al. (2013) commented that transferability determines if the research results can be transferred to another similar situation without altering the meanings from the original study. They added that dependability ensures that the data were gathered in a repeatable and reliable manner and confirmability demonstrates how the research findings were developed from the data.

Credibility

Giorgi (2012) recommended application of methods such as member-checking, concept maps, reflexivity, and peer debriefing to ensure the credibility for a phenomenological study. Member-checking was completed for all 15 participants. As stated previously, I sent each participant a written sanitized interview and then followed up with a 30-minute call to review any questions I had and/or to make any changes the participant desired. Each participant either gave a verbal agreement or sent an e-mail stating that the final version of the sanitized transcript reflected their experience with the research topic.

I kept a journal throughout the research. In the journal, I kept track of the status of the research and wrote down what I thought the next steps were. Sometimes I had ideas that I thought would be helpful in later steps of research and recorded them in the journal for future reference. An excerpt of my reflection on the research from April 18, 2020 is the following:

> I conducted analysis of Q3. For details see memo called "Analysis of Q3."
> I have enough information to write up how I did this analysis such that
> others could follow my thought logic. I can apply this same approach to
> analyzing Q4, although I will find different results I imagine. And
> comparing the results of Q3 and Q4 could also be interesting. For
> example, did I find the same predominant methodology used in both
> questions? Did I find the same frequency of using multiple decision-
> making methods between questions?

Additionally, I wrote memos that I attached to the interviews in NVivo12 including my initial impression of each interview. An excerpt from a memo attached to the P3 transcript was the following:

P3 did an excellent job of explaining both the process for creating the KPI

and the process for the annual target setting process. There was language

that I could align with all three decision-making schools of thought.

Anchoring/adjust, expertise, utility theory, and satisficing came to mind

most often as I reviewed and then coded this interview.

Furthermore, I shared in the analysis section how I did the initial coding and then how I refined and modified the coding as I progressed through the data.

Transferability

Houghton et al. (2013) stated that transferability is determined by the reader reviewing the thick descriptions of the specific context, the research protocol, and any raw data provided by the researcher of the original study. Elo et al. (2014) stated that it is important for the researcher to include clear descriptions of the participants' characteristics. Sousa (2014) said internal consistency is achieved by demonstrating a consistent process of participant selection, interview questions, data collection, and data analysis. Sousa added that external consistency includes the linkage among the application of concepts, theme generation, and writing the results.

Participant selection added to the transferability of my study since I collected the experiences of supply chain leaders from several companies within a variety of industries. I followed a consistent and specific set of characteristics to select the

participants. I also shared participants' characteristics in the demographics section of this chapter. I provided the detailed research protocol I followed as described in Chapter 3 and in Appendix D. I used a standardized research protocol for the interview and member-checking processes. And I applied a consistent approach to analyzing each of the interview questions.

To aid in transferability of the study, I also provided the coding methodology which was based on the conceptual framework of decision-making theory. I provided descriptions of how I moved from raw data to themes and then results thereby making salient possible alternate interpretations of the results. And I included direct quotes from the participants to add to the richness of the interpretations of the phenomena.

Dependability

Houghton et al. (2013) said dependability comes from gathering data in a reliable and repeatable manner. I rigorously followed the data gathering process described in Chapter 3 and Appendix D. Houghton et al. also recommended using an audit trail. I found this recommendation useful and created an Excel file to track every step of the data gathering and analysis phases (Appendix E). The authors also recommended using a reflexivity journal, an important tool to ensure the study results are dependable and confirmable. I found that the most efficient method for journaling was to create a series of memos in NVivo12 that described my initial impressions of each interview, the steps in the research process, as well as my thoughts, and the decisions I made during the analysis.

Confirmability

According to Houghton et al. (2013), confirmability comes from demonstrating the development of the research findings from the data. Berger (2015) explained that reflexivity is the recognition of the impact of the researcher on results of a study. Berger highlighted factors that might influence the researcher's positioning such as age, gender, and professional experience. I did find that my professional experience influenced the line of questioning. I had to mentally remind myself that the responses needed to reflect the participants' experiences with the phenomenon.

Berger (2015) also shared three areas that required the researcher to reflect upon while conducting the study. The first was the ability of the researcher to access and engage the participants. The second was the impact of the researcher-participant relationship. And the third areas was the worldview of the researcher and the impact that worldview could have on how questions were posed and results were interpreted.

I found that getting time on busy senior level supply chain professionals' calendars was difficult. But once we began the interview I found the participants were interested in the questions and became well-engaged in the dialogue. Since I had only met one of the 15 participants previously and only in a professional setting, I felt that there was little if any impact from a researcher-participant relationship on the study results. As I stated earlier I tried to keep top-of-mind that I would keep as much as possible to the interview protocol. If the conversation veered away it was only because that was where the participant took the dialogue, and not the other way around. During

the entire study a goal I kept at the forefront was that the results must reflect the views of the participants with as little impact from the researcher as possible.

Study Results

Q1 and Q2: Name and Define the Most Important Supply Chain KPIs

For Q1, the participants were asked to select an important KPI for their organization and in Q2 to define that KPI. From these responses, I created a table (Table 1) to characterize the results in general themes.

Table 1

Most Important Type of KPI by Participant

KPI			
Service	Cost	Quality	Corporate social responsibility
P1			
P2		P2	
	P3		
P4			
P5			
	P6		
	P7		
		P8	P8
	P9		
P10			
P11			
P12			
P13			
P14			
		P15	
9	4	3	1

Note. Participants P2 and P8 referenced KPIs that included two attributes.

Nine of the 15 responses included service as a KPI, three responses included cost metrics, three included quality, and one included a CSR metric. The first finding was that

60% of the participants in the data set had service as one of their most important KPIs. P4 elaborated on this stating, "We want to get a pulse of how the customers feel as a result of that service." Cost and quality were tied at 20% each with CSR at 6%. Two participants, P2 and P8, cited KPIs that were a combination of two themes, as shown in Table 1.

Q3: Decision-Making Processes Used to Choose Selected KPIs

I used the results of Interview Questions 1 and 2 to frame the discussion for Interview Questions 3 (Q3) and 4 (Q4). Q3 and Q4 were the interview questions I used to get at the heart of the research question, which was, What are the decision-making strategies used by supply chain leaders in choosing the KPIs and goals for their organizations? From the Excel file that showed, by participant, the number of coding references for each decision-making theory (Appendix F), I created a binary table showing whether or not a participant used language that could be associated with a particular type of decision-making theory (Table 2).

Table 2 shows that five participants (P1, P2, P4, P5, and P7) used all decision-making theories. Seven participants (P3, P6, P8, P9, P11, P12, and P14) used three of the decision-making theories. Two participants (P10, P15) used two decision-making theories, and P13 used one decision-making theory. Additionally, 100% of the participants used naturalistic decision-making theory, 86.7% used rational decision-making theory, 80% used fast and frugal heuristics, and 40% used heuristics and biases decision-making theory. When I combined fast and frugal heuristics with heuristics and biases under a generic heuristics decision-making category, I found that 86.7% of the

participants used heuristics as a decision-making theory in selecting their KPIs; this was the same percentage that used rational decision-making theory.

Table 2

Number of Decision-Making Theories Used by Each Participant in Selecting KPIs

Participant ID	Fast and frugal heuristics	Heuristics and biases	Naturalistic decision-making	Rational decision-making	Total decision-making theories
P1	1	1	1	1	4
P2	1	1	1	1	4
P3	1	0	1	1	3
P4	1	1	1	1	4
P5	1	1	1	1	4
P6	1	0	1	1	3
P7	1	1	1	1	4
P8	0	1	1	1	3
P9	1	0	1	1	3
P10	1	0	1	0	2
P11	1	0	1	1	3
P12	1	0	1	1	3
P13	0	0	1	0	1
P14	1	0	1	1	3
P15	0	0	1	1	2
Total # Participants	12	6	15	13	
% of Decision-making theories used	80.0%	40.0%	100.0%	86.7%	

Note. A "1" designates the participant did use that particular decision-making theory, a "0" designates the participant did not.

From Table 2 it can be seen that P13 only applied naturalisti*c* decision-making theory. I went back and reviewed the sanitized transcript for P13 with the thought that perhaps I had missed something. I had coded P13's response: "So a couple of things. One

is getting closer to the customer and understanding what they need and looking at how

our performance helps improve their experience with our products and service," to the

subnode *situation comprehension*. I coded P13's response: "We have been on a journey

of transformation . . . We know that will be a competitive advantage for us going

forward," to the subnode s*ituation recognition*. Both of these subnodes fall under the

parent node *naturalistic decision-making*. Overall P13's response to Q3 was very concise

and to the point regarding the company focus on "Our continuous desire to improve and

better serve our customers."

Additionally, I noticed from Table 2 that P13 and P15 were the only participants

who did not use some form of heuristics-based decision-making theory. As stated above,

P13 only applied language indicative of naturalisti*c* decision-making whereas P15 used

both naturalistic decision-making and rational decision-making language.

Although Table 2 shows how commonly each decision-making theory was used

across all participants, the raw data in Appendix G was used to generate Table 3 that

shows the relative intensity by capturing the frequency with which each decision-making

theory was used. In other words, the raw data in Appendix G captures how many times

each participant used language consistent with a decision-making theory.

As shown in Table 3, rational decision-making language was used 42.2% of the

time by participants when describing the process used to select KPIs. Naturalistic

decision-making language was used 40.6% of the time. And language from rational or

naturalistic decision-making was used much more frequently than the combined

heuristics usage at 17.2% (heuristics and biases at 10.9% and fast and frugal heuristics at 6.3%).

Table 3

Frequency of Use of Each Decision-Making Theory Across All Participants in Selecting KPIs

	Rational decision-making	Naturalistic decision-making	Heuristics and biases	Fast and frugal heuristics
Frequency	108	104	28	16
% Frequency	42.2%	40.6%	10.9%	6.3%

A finding of the study was that although heuristics decision-making theory and rational decision-making theory was used by the same number of participants at 86.7% each (Table 2), heuristics decision-making language was used much less often than either rational or naturalistic decision-making language was used (Table 3) in KPI selection.

Q4: Decision-Making Process Used to Determine Goal for Selected KPIs

From the Excel file in Appendix G, I created a binary table that shows whether or not a participant used language that could be associated with a particular type of decision-making theory (Table 4). Table 4 shows that three participants (P9, P11, and P13) used all four decision-making theories. Ten participants (P1-7, P10, P12, and P14) used three decision-making theories. P15 used two decision-making theories and P8 used one decision-making theory. Additionally, 42.2% of the participants used some form of heuristics theory (13.3% used fast and frugal heuristics and 28.9% used heuristics and

biases) whereas naturalistic decision-making and rational decision-making theory were

used 29.9% and 28.9% of the time respectively.

Table 4

Number of Decision-Making Theories Used by Each Participant in Selecting Goals

Participant	Fast and frugal heuristics	Heuristics and biases	Naturalistic decision-making	Rational decision-making	Total decision-making theories
P1	0	1	1	1	3
P2	0	1	1	1	3
P3	0	1	1	1	3
P4	0	1	1	1	3
P5	0	1	1	1	3
P6	0	1	1	1	3
P7	0	1	1	1	3
P8	0	0	0	1	1
P9	1	1	1	1	4
P10	1	1	1	0	3
P11	1	1	1	1	4
P12	0	1	1	1	3
P13	1	1	1	1	4
P14	1	0	1	1	3
P15	1	1	0	0	2
Total # Participants	6	13	13	13	
% of Decision-making theories used	13.3%	28.9%	29.9%	28.9%	

Note. A "1" designates the participant did use that particular decision-making theory, a "0" designates the participant did not.

Participants used heuristic*s* decision-making theory at the same level as they used

rational decision-making theory (86.7%) and somewhat less frequently than they used

naturalist decision-making theory (100%) in the process of selecting KPIs; but used some

form of heuristics decision-making theory much more frequently (42.2%) than they used either naturalistic decision-making theory (29.9%) or rational decision-making theory (28.9%) in the process of determining goals. A third finding in the study was that there was a difference in the most frequently used decision-making theory in the process of selecting KPIs (naturalistic decision-making theory) versus the process used in determining goals (heuristics decision-making theory).

There were also differences in the number of decision-making theories each participant used in the process of deciding on KPIs versus the number of theories used in the process of choosing goals (Table 4). Five participants used all four decision-making theories and seven participants used three decision-making theories in the process of KPI selection; whereas three participants used all four decision-making theories and 10 participants used three decision-making theories in the process of goal setting. Conversely only one participant used a single decision-making theory in selecting the KPIs (P13) or in determining goals (P8). Another finding was that a majority of the respondents (80.0%) used three or more decision-making theories in the process of either selecting KPIs or in deciding on goals for those KPIs (86.7%).

Table 5

Frequency of Use of Each Decision-Making Theory Across All Participants in Selecting

Goals

	Rational decision-making	Naturalistic decision-making	Heuristics and biases	Fast and frugal heuristics
Frequency	8	22	43	27
% Frequency	27.0%	43.0%	22.0%	8.0%

In Table 5, naturalistic decision-making language was the most frequently used decision-making theory when selecting goals at 43.0% of the time although rational decision-making language was used 27% of the time and heuristics decision-making language was used 30.0% of the time (heuristics and biases at 22.0% and fast and frugal heuristics at 8.0%). An additional study finding was that although heuristics decision-making theory was the most frequently used decision-making theory, naturalistic decision-making language was the most commonly used language used by participants when deciding on goals.

Additionally, naturalistic decision-making language was used fairly consistently between selection of KPIs (40.6%) and determination of goals (43.0%). However rational decision-making language was used much more often in KPI selection (42.2%) than in goal determination (27%). And heuristics decision-making language was used much less often in KPI selection (17.2%) than in goal determination (30%)

Q5: How KPI is Used Within the Organization

The flow chart developed from the responses to Q5 (Figure 1) shows that KPIs and the results of those KPIs are broadly shared throughout the organization. P4 stated that KPIs are published and presented at the North American level as well as in each business unit. P4 added that team members drill down into the reasons when a KPI result is below target noting the deviation is "an indicator of something that's happening out there." Furthermore, results are shared externally with customers: "We do go all the way down to comparing what the customer is telling us versus what our system is telling us," (P1); and with vendors: "Certainly, carrier performance is a key driver and so there are carrier performance reviews" (P2).

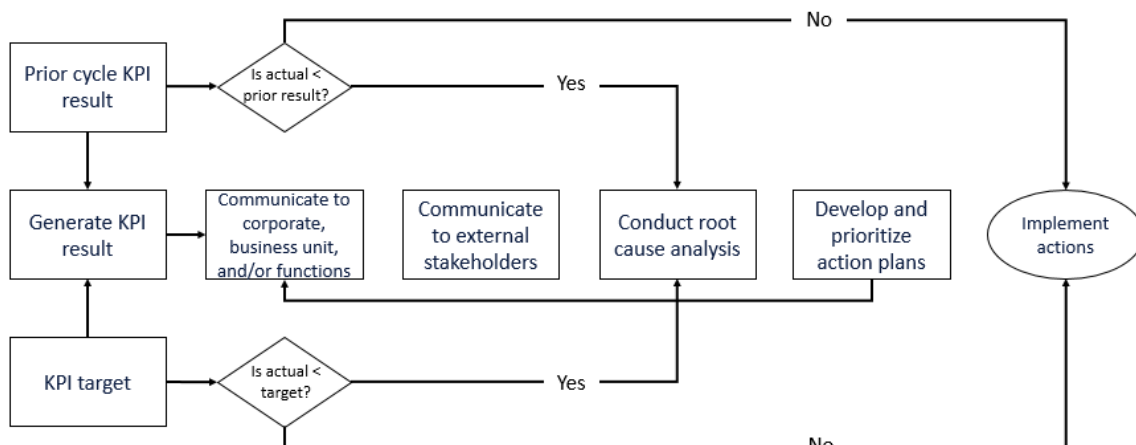*Figure 1*. Process flow describing how participants use their KPIs.

Figure 1 shows that the actual company results are compared to prior period company results and to company targets. P11 elaborated that they not only look at the actual measure but also at any gaps from company target performance. P10 stated that they look at how the specific team results changed versus the prior period, comparing this

to the best team result in the organization. P10 added that vendors and customers provide information that allows for productive dialogue to "breach the gap between reality and perception."

The process flow in Figure 1 moves on to root cause analysis, development of action plans, and implementation of these plans as described by P10:

We define . . . certain activities that we need to put in place with the teams

in order to improve the KPI . . . . We will continue . . . or we change

direction, [and] improve some of the activities that we put in place.

The finding from Q5 and the resultant flowchart is that KPIs are used both as communication vehicles and to drive improved outcomes through a process of feedback, problem analysis, action planning, and implementation.

Q6: Changes in Design or Use of KPI

After the final coding of the responses to Q6 in which participants were asked, if they could, what would they like to change regarding their KPIs. The result was four major themes across the 15 participants. *Better KPIs* was cited by 12 participants and was the most commonly referenced area for improvement. Improvement suggestions also included making data easier and faster to gather along with having more leading versus lagging indicators. P4 described that there is ongoing work to have more predictive indicators, so that leaders can "minimize the risk by taking proactive actions."

The second most frequently mentioned improvement area, with five references, was *experimenting with new metrics*. P3 offered that there is a large initiative in their organization to bring in more of the voice of the customer. This was followed by the

suggested improvement of having an *external view* with four responses. P12 said, "I want to know how the customers are measuring our performance," although P15 was looking for internal alignment on the external view of the company performance. And participants felt KPIs could be better at helping to improve people's behavior as evidenced by P11 who stated their greatest obstacle is when people are "confronted with this topline measure that is outside [their] control; [they] are reluctant to embrace it and action it and . . . to have it in personal goals and objectives."

Q7: Anything I Did Not Ask but Participant Wanted Me to Know

Q7 was intended to act as the catchall question where participants could give one last piece of advice around KPIs. The responses were quite varied with six themes. Ten participants responded in what could be termed as *KPI structure.* Some participants talked about the importance of a balanced set of metrics including P12 who shared, "That is the balanced score card of how we look at things. . . . There are two or three metrics now we look at to give our supply chain teams a good overview of how we assess their performance." Others mentioned that KPIs should be outcomes focused as well as the importance of developing the right KPI for the right level in the organization.

Additional topics that were brought up included talking about other important KPIs and the role organizational structure can play. One person was satisfied with the interview stating, "No, nothing comes to mind" (P3).

Summary

I conducted a descriptive phenomenological study employing a conceptual framework of decision-making theory to answer the central research question for my

study which was, What are the decision-making strategies used by supply chain leaders in choosing the key performance measures and goals for their organizations? I found that a majority of respondents used multiple decision-making strategies in selecting KPIs (80.0% of participants) and in choosing goals (86.7% of participants). I found that all 15 participants used naturalistic decision-making strategies when choosing their measures. And participants used some form of heuristics decision-making theory much more frequently (42.2%) than they used either naturalistic decision-making theory (29.9%) or rational decision-making theory (28.9%) in the process of determining their goals.

Additionally, participants described a consistent general process for KPI usage as a communication mechanism and as a vehicle to drive strategic outcomes via a process of feedback, problem analysis, action planning, and implementation. I created a process model from these descriptions (Figure 1). When asked how their KPIs could be improved, participants wanted the data to be more readily available. Participants wanted more KPIs that were externally focused. And participants wanted leading indicators to enable team members to get in front of situations and drive proactive solutions. The most commonly cited KPIs were service-related following by cost-related and quality-related (at the same frequency) and then sustainability-related KPIs.

Mayoh and Onwuegbuzie (2015) stated that in a descriptive phenomenology the researcher is looking for commonality across the participants. The participants in my study used language that could be mapped to one or more decision-making theories in either choosing KPIs or deciding on targets. The participants described very similar processes in how KPIs were used within their respective organizations. There was

consistency among participants in what they would like to see improved in the design or use of their KPIs. And what clearly came through the interviews was a laser focus on the customer and the business.

In Chapter 4, I shared the data collection process, the data analysis process, how I made sure the study was trustworthy, and the results of the data analysis. In Chapter 5, I interpret these results and highlight limitations of the study along with recommendations for future study. Finally I discuss the implications of the study and then bring the chapter to a close.

Chapter 5: Discussion, Conclusions, and Recommendations

The purpose of this descriptive phenomenological study was to improve the understanding of the decision-making strategies used by supply chain leaders in their selection of key performance measures and associated goals. I interviewed 15 participants from companies recognized as having exemplary supply chains. Then I applied a conceptual framework of decision-making theory to analyze the responses.

There were several findings that could prove beneficial to the broader supply chain community with regard to KPI and goal selection. The majority of participants used multiple decision-making strategies when selecting KPIs (80.0%) or setting goals (86.7%). Naturalistic decision-making strategies, characterized as making decisions in-situ, were used by 100% of the participants in selecting their KPIs. Heuristics decision-making and rational decision-making strategies were used equally often at 86.7% of the time in KPI selection. Heuristics decision-making strategies were used most frequently when choosing goals (42.2%) followed by naturalistic decision-making (29.9%) and rational decision-making (28.9%) strategies being used at nearly the same frequency in goal selection. KPIs that focused on the customer and the business were key.

Interpretation of Findings

Finding 1: Most Elements of the Balanced Scorecard Selected as KPIs

The majority of respondents in my study (60%) selected a service metric as one of their most important KPIs followed by cost and quality at 20% each and then CSR at 6%. Hoque (2014) shared that the four aspects of the BSC include a financial view, a customer view, an internal process view, and a learning and development view. The

results from my research are consistent with three of the four elements of the BSC with the customer view being measured as a service KPI, the finance view being measured via cost KPIs, and the internal process view being captured via quality KPIs. Only the aspect of learning and development was missing in the participant responses. Although the focus on service, cost, and quality was expected, receiving only one response on CSR was not consistent with the view of Bhattacharya et al. (2013) that a sustainability perspective be added as an additional perspective to the traditional BSC. Woods and Van der Meulen (2016) shared that Gartner had changed the criteria for the Top 25 Supply Chains recognition to include CSR. Given that all of the participating companies in my study were recognized as a Gartner Top 25, only a 6% response rate for CSR as a KPI was surprisingly low.

Finding 2: Majority of Participants Used Multiple Decision-Making Strategies

Another key finding of the study was that the majority of participants used multiple decision-making strategies when selecting KPIs (80.0%) or setting goals (86.7%). This finding is consistent with the perspective of Gigerenzer (2015) who stated that there are benefits and drawbacks to each decision-making approach. Franklin (2013) stated that although rational decision-making has a formal and repeatable process structure that includes generating an exhaustive list of solutions, valuing each option, and then choosing the option with the greatest value, the rigidity imposed by the process is problematic. Tversky and Kahneman (1974) found that people regularly made decisions that violated the rules of rational decision-making because humans had difficulty determining probabilities and calculating values that were required of rational decision-

making models. Tversky and Kahneman found that individuals frequently used heuristics to simplify decision-making. Hafenbrädl et al. (2016) labeled many such rules as fast and frugal heuristics because research showed that these decision-making strategies approached or even outperformed solutions using rational rules of decision-making. Klein (2015) introduced a naturalistic decision-making model that balanced relying on intuition for speed with an intentional analytical strategy when the situation required. Gigerenzer stated that there has been a shift in decision-making research, from assuming logic or statistics is normative in all decision-making situations, to conducting research through the lens of ecological rationality, in which the environment determines the preferred decision-making approach--heuristics or otherwise.

I determined which decision-making strategy each participant used in either selecting KPIs or determining goals by breaking down the participant's language into specific meaning units that were coded to one or more decision-making theories. Shan and Young (2016) found that both naturalistic decision-making and fast and frugal heuristics shared some of the same belief structures including the proficiency of the decision-maker, existence of a defined process, and decision rules based on decision-action matching. Consistent with Shan and Young's findings, I found that the language in a single meaning unit often resulted in me coding that meaning unit to more than one decision-making strategy. This coding led me to interpret that most participants employed more than one decision-making strategy when selecting KPIs or when choosing goals.

Finding 3: Naturalistic Decision-Making Used by All Participants in KPI Selection

A third key finding of the study was that naturalistic decision-making strategies were used by 100% of the participants in selecting their KPIs. Participants also used rational decision-making or heuristics decision-making strategies equally often at 86.7% each. Klein and Wright (2011) stated that naturalistic decision-making focused on creating useful models that could be applied to messy problems for which there is no single correct decision; factors that often exist in real work situations. Boyes and Potter (2015) stated that naturalistic decision-making is a combination of intuition and analysis, adding that greater experience leads to an expanded repertoire of prior decisions that could be drawn upon by the decision-maker to improve both the speed and the quality of expert-based decision-making in the field. Klein (2015) added that naturalistic decision-making research is about more than preventing errors in decision-making; it is about the improvement and contribution of experience to making good decisions in a variety of complex situations. Because the participants in this study were experienced supply chain leaders working in organizations recognized as having highly performing supply chains, it follows that all of these experts would use language consistent with the naturalistic decision-making process including (a) situation recognition, (b) situation comprehension, (c) mental simulation, and (d) sequential option evaluation.

Boyes and Potter (2015) stated that there is overlap in process between naturalistic decision-making and heuristics decision-making in matching the current situation to a prior similar situation. Okoli et al. (2013) stated that there is overlap in process between naturalistic decision-making and rational decision-making in option

evaluation. Based on prior research, it is logical that participants' language used in selecting KPIs and coded to naturalistic decision-making theory, could also be coded to rational decision-making theory or to heuristics decision-making theory thereby confirming the current literature of process overlap in decision-making theories.

Finding 4: Heuristics Decision-Making Used Most Frequently When Choosing Goals

I found that participants used heuristics decision-making strategies most frequently (42.2%) when choosing goals. Naturalistic decision-making and rational decision-making strategies were used much less often in goal selection at 29.9% and 28.9%, respectively. Tversky and Kahnemann (1974) found that people often use heuristics to simplify decision-making.

Braga et al. (2015) shared that there is often confusion between the widely researched representativeness heuristic and the availability heuristic. Braga et al. explained that when a decision is based upon how closely a situation is to a stereotypical representation of the target, an individual would use the representative heuristic. When the decision relies on how easily specific examples came to mind, an individual is more likely to use the availability heuristic (Braga et al., 2015). I found that participants used many examples of language consistent with the representativeness heuristic or the availability heuristic. Participants described, as part of the goal setting process, that they often participated in external benchmarking groups in which there was already an agreed to performance measurement framework. Participants used the representative heuristic when describing how their respective companies performed relative to the benchmarks.

The benchmark conversations centered around setting targets and determining the target level required to be considered good or more appropriately to this group*, excellent.* Additionally, participants stated that customers and regulatory bodies have made clear the target level that is considered acceptable performance using language consistent with the availability heuristic.

I also frequently used the anchor and adjust heuristics subnode while coding responses to the interview question on how supply chain leaders chose their goals. Tversky and Kahnemann (1974) described the anchoring and adjust heuristic process as one in which the individual has a starting value and then adjusts either up or down from that reference point by incorporating additional information. Cheek and Norem (2016) said that anchoring effects often occur in real world situations even among experts and particularly among individuals with an analytical thinking style. I found that several study participants described their goal setting process as starting with the current result (anchor) and then the participants described assessing the impact of various improvement initiatives to positively adjust the target along with any environmental conditions that would cause the participants to negatively adjust the target.

It became apparent to me upon reviewing participants' responses to the interview question on goal setting that an important step in the decision-making process was first to determine the anchor, either through beginning with an actual result, a benchmark, customer input, and/or other internal or external stakeholder feedback. And then naturalistic decision-making theory language or rational decision-making theory language

often became apparent as the participant described the process used to adjust up or down from the anchor.

Limitations of the Study

Myers (2013) stated limitations are the potential weaknesses of a study and are often out of the researcher's control. Regardless, the researcher must determine how to minimize the impact of the limitations on the study. The following is a description of limitations for this study along with mitigation strategies employed for each limitation.

Malterud (2012) cautioned that the researcher remain aware of the effects an investigator could have on the participants' responses. At the time of data collection, I was the chief supply chain officer in my company; therefore, participants might knowingly or unknowingly have modified their responses based on what they thought I wanted or expected to hear during the interview. To mitigate this risk, I shared my background and the reason for my interest in research on how supply chain leaders choose the KPIs and goals for their organizations. I openly expressed my admiration to these participants, as senior leaders, for their roles in helping their companies earn the reputation and public recognition of having outstanding supply chains. I also made it clear that as a researcher and a professional I was most interested in what I could learn from each of them.

Malterud (2012) highlighted that a limitation of a descriptive phenomenological study could occur during the process of de-constructing the data. One potential risk area for the study was that the process of assigning meaning units to one or more decision-making theories was somewhat of a subjective process. So rather than ask myself what

decision-making theory a particular meaning unit should be coded to, I compared the language in the meaning unit to the attributes (subnodes) of each decision-making theory in the codebook. When a match was found I coded it to that subnode. I continued this process of comparing the meaning units to the subnodes. The process meant I might have erred on the side of coding a meaning unit to more decision-making theories than if I took the alternate approach. As a second coding check, I reviewed all the meaning units coded to each subnode and checked again for fit. If I deemed the meaning unit did not belong to that subnode, I determined if that meaning unit was or should be coded to another subnode. If not, I removed the meaning unit from the subnode.

According to Sousa (2014), one limitation of qualitative research is constructing a framework of concepts that could be applied to a more general population. Sousa added that to overcome this limitation is to ensure transferability of a study through trustworthiness of the data collection, analysis, and reporting processes. I provided sufficient detail for each step in the process so that someone reading the study will have an adequate understanding of both the context and the steps taken to analyze, categorize, reduce, and reconstruct the common experiences of the participants in choosing the key performance measures and goals for their organizations. The participants in this study should be considered as extreme cases since they belong to organizations recognized as having top supply chain functions. Conducting the same study with a population coming from average or under-performing supply chain organizations could yield different findings.

Recommendations

Chan et al. (2014) asserted that there is a link between performance of the supply chain and the performance of the organization. And although there is a substantial amount of literature on the design, implementation, and use of PMSs, there is a gap in the literature on how leaders actually decide on the performance measures and goals for their respective organizations. I chose to apply a decision-making conceptual framework to improve the understanding and knowledge of the strategies used by senior supply chain professionals from companies having highly regarded supply chains and I shared those results and interpretations in this chapter. As a result of this study and other prior research I have several recommendations regarding further research.

The first recommendation is to conduct the same descriptive phenomenological study but rather than using participants from exemplar supply chain organizations, select participants from companies considered to be under-performing. Using the assumption provided by Franco-Santos et al. (2012), who stated that prior research indicates a positive relationship between the quality of performance measurement systems and company performance, a future researcher could use publicly available information to find companies performing below the norm against a widely used index such as the S&P 500. Results from the two studies could be compared and contrasted.

Giorgi (2012) stated that a core tenet of phenomenological research is capturing participants' experiences in their own words. The study I conducted was done from an ex-post facto perspective in which participants had to recall their experiences with the phenomenon. From the language in the interviews it seemed clear that in some cases the

experience was fairly current and in others the experience had happened further in the past. Replicating this study with participants from organizations who were actively going through a design or a refresh phase in developing their KPIs and targets could provide additional insights into the decision-making strategies used by supply chain leaders in selecting their KPIs or targets. The two studies could then be compared and contrasted. I expect that contemporaneous collection of data could potentially pose new obstacles I did not experience in my data collection phase; the most obvious obstacle being finding a sample population that is actively going through a design or refresh of their performance measurement systems.

A third area of potential research could include selecting a single decision-making theory and conducting an in-depth case study analysis with several participants from the same organization. Since the results in this study showed that all participants used naturalistic decision-making in selecting their respective KPIs, a future researcher could use one of the naturalistic decision-making models, such as Klein's (2015) recognition-primed decision-making model as the conceptual framework. The organization could be selected from the Gartner Supply Chain Top 25 (Aronow et al., 2016) list of high performing organizations. A more detailed study through the recognition-primed decision-making lens could provide additional insights into what aspects of naturalist decision-making were actually used by supply chain leaders in choosing their KPIs.

Since participants used a heuristics decision-making strategy in choosing their targets, a future researcher could conduct a study to explore which heuristics models are used most often in selecting KPI targets. The study could be a quantitative study

conducted via survey. Some heuristics models that seem relevant for such a study include the recognition heuristic (Artinger et al., 2015), the satisficing heuristic (Simon, 1995), and the anchor and adjustment heuristic (Tversky & Kahneman, 1974).

A side benefit of the study I conducted was the creation of a common process model (Figure 1) that describes how participants used the KPIs and goals in their organizations. Future quantitative research could be conducted to validate the effectiveness of the model in driving business outcomes. A future researcher could conduct an action research study to improve the performance of an organization by applying the proposed process model and coming up with a method of evaluating if the model provided any benefit to that organization.

Implications

Implications for Social Change

Cecere et al. (2016) found that companies with high performing supply chains provided higher levels of return for their shareholders. Ellinger et al. (2012) found a positive correlation between supply chain performance and customer service levels. Franco-Santos et al. (2012) stated that the design and use of performance measurement systems are key to organizational performance. The purpose of this qualitative phenomenological study was to improve the understanding of the decision-making strategies used by supply chain leaders, from highly performing companies, in their selection of key performance measures and goals for their organizations. Insights from this study could help improve supply chain performance leading to better customer service and increased shareholder benefit resulting in positive social change.

Most of the participants in the study chose a customer-facing KPI as their most important metric. Additionally, much of the language from participants regarding the goal-setting process reflected a strong focus on the customer, whether that focus came from within the organization or via feedback directly from customers. Participants also specifically stated that KPIs and results were used to drive improvement actions with the explicit goal of improving service to their customers.

A key recommendation from the study is to include a customer-oriented service measure as a KPI when designing performance measurement systems. Focusing on the customer can drive positive social change for both the customer organization as well as the servicing organization. High customer service means the customers, in turn, will have the product needed to meet their customers' needs and so on up the value chain, with each organization helping to ensure the upstream organizations' future viability. Likewise, this means individuals in the organizations will be able to prosper along with the companies and the communities within which the companies operate.

When asked about the decision-making strategies used to select KPIs, all respondents in the study used language consistent with naturalistic decision-making theory. Participants used language demonstrating that first they assessed the situation and then they relied on prior experience to develop a set of possible solutions, picking the solution with the closest fit to the current situation. This finding confirmed Boyes and Potter's (2015) research showing that a broader set of experiences lead to a larger portfolio of prior decisions that the decision-maker drew upon to improve the efficiency and efficacy of expert-based decision-making in the field. A logical recommendation

from the finding and from prior research is that managers should provide the opportunity

for other members of the organization to participate in the development of KPIs. Doing

so allows for the development of an expanded toolbox of experiences for each member

thereby building organizational muscle for continued KPI improvement, ultimately

leading to better organizational performance. And as stated above, improved performance

impacts the organization, the customers, and the communities within which they all

operate.

I also found that participants generally used multiple decision-making theories in

KPI selection. The use of multiple strategies is consistent with the research of Boyes and

Potter (2015) who found that a combination of intuition and analysis was used by

decision-makers to improve their respective decision-making processes in real life

situations. A third recommendation is that leaders should make their own decision-

making processes transparent so that they can intentionally incorporate more than one

decision-making strategy in KPI development. Including other stakeholders in the

process may also improve the result by getting multiple viewpoints and likely, multiple

approaches to decision-making. According to Franco-Santos et al. (2012), including

others in the decision-making process can have the additional benefit of increasing

owners' and designers' satisfaction with the performance measurement system

potentially increasing the performance of the organization.

Participants used heuristics decision-making nearly twice as often as they used

naturalistic decision-making or rational decision-making when establishing goals. More

specifically, participants used language consistent with anchor and adjustment,

representativeness, and satisficing heuristics. Leaders should make transparent the decision-making process they are using when working through goal-setting. Consistent with the findings of Franco-Santos et al. (2012), who found a positive relationship in performance when individuals were included in setting performance goals, making such processes salient will improve the acceptance and ownership of the goals by those who have participated in the process. Selecting the right performance measures can contribute to driving desired individual behavior (Marginson et al., 2014). And as stated above improvement in the goal setting process can positively impact the individual, the organization, and the stakeholders of the organization.

Implications for Theory

This study was conducted to address the research problem regarding a lack of knowledge and understanding with respect to the decision-making strategies used by supply chain leaders in selecting the key performance measures and goals for their respective organizations. I applied a conceptual framework of decision-making theory in analyzing the transcripts of 15 senior supply chain professionals from companies with highly regarded supply chains. Applying this unique lens to the study of how supply chain leaders actually chose their KPIs and goals allowed me to discover that these participants used more than one decision-making theory when selecting KPIs and choosing targets. I also found that all participants used naturalistic decision-making strategies when selecting KPIs. And I found that participants used heuristics decision-making strategies more often than other forms of decision-making theory when choosing goals. My study supported the recommendation of Gore et al. (2015) who stated that an

emerging area of decision-making research is the combination of naturalistic decision-making methods with rational as well as heuristics and biases research traditions.

Implications for Professional Practice

I constructed a general model (Figure 1) describing how participants used their KPIs, goals, and actual results as both a communication vehicle and a continuous improvement process within their respective organizations to drive performance outcomes. The model that I developed from the participants' responses was consistent with three of the four classifications Henri (2006) recommended for the use of KPIs: (a) monitoring, (b) focusing attention, and (c) strategic decision-making. Although I did not find explicit evidence of Henri's fourth classification, legitimization, that does not mean it was not present. Rather my interview questions were not designed to understand if the KPIs were or were not used to legitimize prior decisions made or actions taken.

The model I developed also supported Koufteros' et al. (2014) claim that managers should use their performance management systems for diagnostic and/or interactive purposes. The diagnostic piece as described by the participants included monitoring performance against targets as well as focusing the participants' attention on needed improvements. And the interactive aspect of the model included building organizational capabilities through dialogue that stimulated the development of new ideas and new actions.

I recommend that supply chain professionals apply the model I developed as an output of this study to determine if there are any gaps between the model and their own performance measurement systems. Koufteros et al. (2014) claimed that organizations

with well-developed PMSs should outperform those with less developed systems. Thus, conducting the analysis and closing gaps in an organization's PMS could help improve the effectiveness of the use of their KPIs and goals possibly leading to improved organizational performance. Additionally, I recommend regularly measuring actual performance against a target and then using any gaps identified to drive communication on results throughout the organization. I also recommend using those gaps to drive analysis and action for improved organizational outcomes.

I also found that participants were in strong agreement regarding the need to generate KPI results in a more efficient manner. This finding was consistent with the research of Chalmeta et al. (2012) and of Gutierrez et al. (2015) that barriers to success for a measurement system included difficulties in collecting complete, accurate, and relevant data from information systems. Participants in my study shared that KPIs must include leading indicators, in addition to the more traditional lagging indicators. Leading indicators will allow organizations to proactively implement corrective solutions. This perspective is consistent with the recommendation of Gibbons and Kaplan (2016) that scorecards must balance lagging financial indicators with leading indicators reflecting the customer, the internal process, and the learning and growth views of organizational performance. The advice of the participants, all of whom came from organizations with well-regarded supply chains, also supports the conclusions of Koufteros et al. (2014) that highly performing organizations deploy PMSs in a balanced manner to drive swift actions that are consistent with organizational goals and deliver strong performance over time.

Conclusions

I found that the supply chain leaders in my study relied on their broad and deep supply chain experience when deciding on the key performance measures they used to drive exemplary supply chain performance. It was evident to me that these leaders used various combinations of naturalistic, heuristic, and rational decision-making strategies based on the language they used when describing the processes they and their teams went through in selecting their KPIs. And in every case, the participants used language consistent with naturalistic decision-making in KPI selection.

I also found that the supply chain leaders in my study overwhelmingly used language consistent with heuristics decision-making theory when these leaders described how they established their goals. The participants used a variety of internal and external inputs when establishing an initial target for a KPI. And then they adjusted that target in a positive direction as a result of incorporating the expected benefits of the future action plans they had developed with their teams, their customers, their vendors, and other stakeholders. Additionally these supply chain leaders adjusted the target in a negative direction as a result of various environmental influences that could adversely impact the target. Targets were established using a well-developed process; not merely chosen based on someone's feelings on what a goal should be.

What also came through loud and clear was the process discipline in using the performance measures and goals to develop action plans either to improve performance against a previous result or to proactively get in front of a potential situation and keep that situation from adversely impacting results. It was apparent from the participant

interviews that KPIs and goal setting was not merely an exercise that was done to placate senior management or external parties. The participants were passionate and committed to delivering results. And they saw the selection of performance measures and goals as a key step in their overall supply chain strategy.

For supply chain leaders who might read my study, I want to leave a few key recommendations. Make the selection of your performance measures and goals an important first step in your overall supply chain strategic planning process. Work in a collaborative manner to decide on your KPIs and targets incorporating elements of naturalistic, rational and heuristics decision-making into the process. And use your KPIs, goals, and actual results as the basis for developing future state actions that will improve your service, reduce your cost, engage your teams, and help deliver on strong and sustained overall organizational performance and help you on your journey to performance excellence.

References

Abushaiba, I. A., & Zainuddin, Y. (2012). Performance measurement system design, competitive capability, and performance consequences: A conceptual like. *International Journal of Business and Social Science*, *3*(11), 184–193. Retrieved from http://ijbssnet.com/index.php/home

Agostino, D., & Arnaboldi, M. (2012). Design issues in Balanced Scorecards: The "what" and "how" of control. *European Management Journal*, *30*(4), 327–339. https://doi.org/10.1016/j.emj.2012.02.001

Albar, F. M., & Jetter, A. J. (2013). Uncovering project screening heuristics with cognitive task analysis: How do gatekeepers decide which technologies to promote? In *2013 Proceedings of PICMET '13: Technology Management in the IT-Driven Services (PICMET)* (pp. 459–467). Retrieved from http://ieeexplore.ieee.org

Allais, M., & Hagen, O. (1979). *Expected utility hypotheses and the Allais Paradox: Contemporary discussions of decisions under uncertainty with Allais' rejoinder*. Dordrecht, Germany: D. Reidel Publishing Company.

Allison, G. T. (1971). *Essence of decision; Explaining the Cuban missile crisis*. Boston, MA: Little, Brown.

Almy, B., & Krueger, J. I. (2013). Game interrupted: The rationality of considering the future. *Judgment & Decision-making*, *8*(5), 521–526. Retrieved from http://journal.sjdm.org

Anosike, P., Ehrich, L. C., & Ahmed, P. (2012). Phenomenology as a method for exploring management practice. *International Journal of Management Practice*, *5*(3), 205-224. https://doi.org/10.1504/IJMP.2012.048073

Aronow, S., Burkett, M., Nilles, K., & Romano, J. (2016). The Gartner Supply Chain Top 25 for 2016. Retrieved from http://www.gartner.com/document/3321523?ref=solrAll&refval=170440625&qid =b68dcb04f939844d8f2262e3a29491d1

Artinger, F., Petersen, M., Gigerenzer, G., & Weibler, J. (2015). Heuristics as adaptive decision strategies in management. *Journal of Organizational Behavior*, *36*(S1), S33–S52. https://doi.org/10.1002/job.1950

Bahník, Š., & Strack, F. (2016). Overlap of accessible information undermines the anchoring effect. *Judgment & Decision-making*, *11*(1), 92–98. https://doi.org/10.1016/j.copsyc.2016.06.005

Barberis, N. C. (2012). *Thirty years of prospect theory in economics: A review and assessment* (No. W18621). Retrieved from National Bureau of Economic Research website: www.nber.org/papers/w18621

Becker-Peth, M., & Thonemann, U. W. (2016). Reference points in revenue sharing contracts—How to design optimal supply chain contracts. *European Journal of Operational Research*, *249*(3), 1033–1049. https://doi.org/10.1016/j.ejor.2015.06.007

Berg, N. (2014). Success from satisficing and imitation: Entrepreneurs' location choice and implications of heuristics for local economic development. *Journal of

Business Research, *67*(8), 1700–1709.

https://doi.org/10.1016/j.jbusres.2014.02.016

Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in

qualitative research. *Qualitative Research*, *15*(2), 219–234.

https://doi.org/10.1177/1468794112468475

Bevan, M. T. (2014). A method of phenomenological interviewing. *Qualitative Health*

Research, *24*(1), 136–144. https://doi.org/10.1177/1049732313519710

Bhattacharya, A., Mohapatra, P., Kumar, V., Dey, P. K., Brady, M., Tiwari, M. K., &

Nudurupati, S. S. (2013). Green supply chain performance measurement using

fuzzy ANP-based balanced scorecard: A collaborative decision-making approach.

Production Planning & Control, *25*(8), 698–714.

https://doi.org/10.1080/09537287.2013.798088

Bititci, U. S., Garengo, P., Dörfler, V., & Nudurupati, S. (2012). Performance

measurement: Challenges for tomorrow. *International Journal of Management*

Reviews, *14*(3), 305–327. https://doi.org/10.1111/j.1468-2370.2011.00318.x

Blokdjik, G. (2008). *Change management 100 success secrets: The complete guide to*

process, tools, software and training in organizational change management.

Brisbane, Australia: Emereo Ltd.

Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing,

implementing and updating performance measurement systems. *International*

Journal of Operations & Production Management, *20*(7), 754–771.

https://doi.org/10.1108/01443570010330739

Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, *138*(3), 389–414. https://doi.org/10.1037/a0026450

Boyes, M., & Potter, T. (2015). The application of recognition-primed decision theory to decisions made in an outdoor education context. *Australian Journal of Outdoor Education*, *18*(1), 2–15. Retrieved from http://outdooreducationaustralia.org.au/joee/

Braga, J. N., Ferreira, M. B., & Sherman, S. J. (2015). The effects of construal level on heuristic reasoning: The case of representativeness and availability. *Decision*, *2*(3), 216–227. https://doi.org/10.1037/dec0000021

Brinkmann, S., & Kvale, S. (2015). *Interviews: Learning the craft of qualitative research interviewing* (3rd ed.). Los Angeles, CA: Sage.

Buchholz, W., & Schymura, M. (2012). Expected utility theory and the tyranny of catastrophic risks. *Ecological Economics*, *77*, 234–239. https://doi.org/10.1016/j.ecolecon.2012.03.007

Cecere, L., Hart, H., Denman, R., & King, H. (2016, July 26). 2016 supply chains to admire. Retrieved from http://supplychaininsights.com/2016-supply-chains-to-admire/

Chai, J., Liu, J. N. K., & Ngai, E. W. T. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications*, *40*(10), 3872–3885. https://doi.org/10.1016/j.eswa.2012.12.040

Chalmeta, R., Palomero, S., & Matilla, M. (2012). Methodology to develop a performance measurement system in small and medium-sized enterprises. *International Journal of Computer Integrated Manufacturing*, *25*(8), 716–740. https://doi.org/10.1080/0951192X.2012.665178

Chalyvidis, C. E., Ogden, J. A., & Johnson, A. W. (2013). Using supply chain interoperability as a measure of supply chain performance. *Supply Chain Forum: International Journal*, *14*(3), 52–73. Retrieved from http://www.supplychain-forum.com/index.cfm

Chan, F. T. S., Nayak, A., Raj, R., Chong, A., & Manoj, T. (2014). An innovative supply chain performance measurement system incorporating research and development (R&D) and marketing policy. *Computers & Industrial Engineering*, *69*, 64–70. https://doi.org/10.1016/j.cie.2013.12.015

Cheek, N. N., & Norem, J. K. (2016). Holistic thinkers anchor less: Exploring the roles of self-construal and thinking styles in anchoring susceptibility. *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2016.01.034

Cheng, M. M., & Humphreys, K. A. (2012). The differential improvement effects of the strategy map and scorecard perspectives on managers' strategic judgments. *Accounting Review*, *87*(3), 899–924. https://doi.org/10.2308/accr-10212

Choong, K. K. (2013). Understanding the features of performance measurement system: A literature review. *Measuring Business Excellence*, *17*(4), 102–121. https://doi.org/10.1108/MBE-05-2012-0031

Choong, K. K., & Burgess, T. (2014). The fundamentals of performance measurement systems: A systematic approach to theory and a research agenda. *International Journal of Productivity and Performance Management*, *63*(7), 879–922. https://doi.org/10.1108/IJPPM-01-2013-0015

Chorus, C. (2012). Random regret minimization: An overview of model properties and empirical evidence. *Transport Reviews*, *32*(1), 75–92. https://doi.org/10.1080/01441647.2011.609947

Coe, N., & Letza, S. (2014). Two decades of the balanced scorecard: A review of developments. *Poznan University of Economics Review*, *14*(1), 63–75. Retrieved from http://ebr.edu.pl/

Cohen, M. Z., Kahn, D. L., & Steeves, R. H. (2000). *Hermeneutic phenomenological research: A practical guide for nurse researchers*. Thousand Oaks, CA: Sage.

Cristancho, S. M., Vanstone, M., Lingard, L., LeBel, M.-E., & Ott, M. (2013). When surgeons face intraoperative challenges: a naturalistic model of surgical decision-making. *The American Journal of Surgery*, *205*(2), 156–162. https://doi.org/10.1016/j.amjsurg.2012.10.005

Cunliffe, A. L., & Karunanayake, G. (2013). Working within hyphen-spaces in ethnographic research: Implications for research identities and practice. *Organizational Research Methods*, *16*(3), 364–392. https://doi.org/10.1177/1094428113489353

Dahlgaard-Park, S. M., Chen, C.-K., Jang, J.-Y., & Dahlgaard, J. J. (2013). Diagnosing and prognosticating the quality movement: A review on the 25 years quality

literature (1987–2011). *Total Quality Management & Business Excellence*, *24*(1/2), 1–18. https://doi.org/10.1080/14783363.2012.756749

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, *22*(5), 1915–1953. Retrieved from https://rfs.oxfordjournals.org/

Deshpande, A. (2012). Supply chain management dimensions, supply chain performance and organizational performance: An integrated framework. *International Journal of Business and Management*, *7*(8), 2–19. Retrieved from http://www.theijbm.com/

Dikko, M. (2016). Establishing construct validity and reliability: Pilot testing of a qualitative interview for research in Takaful (Islamic Insurance). *The Qualitative Report*, *21*(3), 521–528. Retrieved from http://nsuworks.nova.edu/tqr/vol21/iss3/6

Dworkin, S. L. (2012). Sample size policy for qualitative studies using in-depth interviews. *Archives Of Sexual Behavior*, *41*(6), 1319–1320. https://doi.org/10.1007/s10508-012-0016-6

EFQM. (2015). EFQM model: Fundamental concepts. Retrieved from www.efqm.org

Ellinger, A., Shin, H., Northington, W. M., Adams, F. G., Hofman, D., & O'Marah, K. (2012). The influence of supply chain management competency on customer satisfaction and shareholder value. *Supply Chain Management*, *17*(3), 249–262. https://doi.org/http://dx.doi.org./10.1108/13598541211227090

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014).

    Qualitative content analysis. *SAGE Open*, *4*(1), 1–10.

    https://doi.org/10.1177/2158244014522633

Englander, M. (2012). The interview: Data collection in descriptive phenomenological

    human scientific research. *Journal of Phenomenological Psychology*, *43*(1), 13–

    35. https://doi.org/10.1163/156916212X632943

Estampe, D., Lamouri, S., Paris, J.-L., & Brahim-Djelloul, S. (2013). A framework for

    analying supply chain performance evaluation models. *International Journal of*

    Production Economics, *142*(2), 247–258.

    https://doi.org/10.1016/j.ijpe.2010.11.024

Fawcett, S. E., Fawcett, A. M., Watson, B. J., & Magnan, G. M. (2012). Peeking inside

    the black box: Toward an understanding of supply chain collaboration dynamics.

    Journal of Supply Chain Management, *48*(1), 44–72.

    https://doi.org/10.111/j.1745-493X.2011.03241.x

Fennell, J., & Baddeley, R. (2012). Uncertainty plus prior equals rational bias: An

    intuitive Bayesian probability weighting function. *Psychological Review*, *119*(4),

    878–887. https://doi.org/10.1037/a0029346

Franco-Santos, M., Lucianetti, L., & Bourne, M. (2012). Contemporary performance

    measurement systems: A review of their consequences and a framework for

    research. *Management Accounting Research*, *23*(2), 79–119.

    https://doi.org/10.1016/j.mar.2012.04.001

Franklin, C. L. (2013). Developing expertise in management decision-making. *Academy of Strategic Management Journal*, *12*(1), 21–37. Retrieved from

http://www.alliedacademies.org/academy-of-strategic-management-journal/

Fredrickson, J. E. (2013). Prospect theory and saving behaviors during the Great Recession: 2009 SCF Panel survey results. *Journal of Behavioral Studies in Business*, *6*, 1–17. Retrieved from http://www.aabri.com/jbsb.html

Gary, M. S., Wood, R. E., & Pillinger, T. (2012). Enhancing mental models, analogical transfer, and performance in strategic decision-making. *Strategic Management Journal*, *33*(11), 1229–1246. https://doi.org/10.1002/smj.1979

Gibbons, R., & Kaplan, R. S. (2015). Formal measures in informal management: Can a balanced scorecard change a culture? *American Economic Review*, *105*(5), 447–451. https://doi.org/10.1257/aer.p20151073

Gigerenzer, G. (2014). *Risk savvy: How to make good decisions*. New York, NY: Viking.

Gigerenzer, G. (2015). *Simply rational: Decision-making in the real world* (2nd ed.). New York, NY: Oxford University Press.

Gill, M. J. (2014). The possibilities of phenomenology for organizational research. *Organizational Research Methods*, *17*(2), 118–137. https://doi.org/10.1177/1094428113518348

Gilmore, D. (2012a, February 17). Metrics strategies and supply chain performance. *Supply Chain Digest*. Retrieved from

http://www.scdigest.com/assets/FirstThoughts/12-02-17.php?cid=5505

Gilmore, D. (2012b, March 9). Supply chain metric strategies and the 50 percent problem. *Supply Chain Digest*. Retrieved from http://www.scdigest.com/assets/FirstThoughts12-03-09.php?cid=5587

Giorgi, A. (2009). *The descriptive phenomenological method in psychology: A modified Husserlian approach*. Pittsburgh, PA: Duquesne University Press.

Giorgi, A. (2012). The descriptive phenomenological psychological method. *Journal of Phenomenological Psychology*, *43*(1), 3–12. https://doi.org/10.1163/156916212X632934

Glimcher, P. (2015). Understanding the hows and whys of decision-making: From expected utility to divisive normalization. *Cold Spring Harbor Symposia on Quantitative Biology*, 024778. https://doi.org/10.1101/sqb.2014.79.024778

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition*, *123*(1), 21–32. https://doi.org/10.1016/j.cognition.2011.12.002

Golicic, S. L., & Davis, D. F. (2012). Implementing mixed methods research in supply chain management. *International Journal of Physical Distribution & Logistics Management*, *42*(8/9), 726–741. https://doi.org/10.1108/09600031211269721

Gonzalez, C., Meyer, J., Klein, G., Yates, J. F., & Roth, A. E. (2013). Trends in decision-making research: How can they change cognitive engineering and decision-making in human factors? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 163–166. https://doi.org/10.1177/1541931213571037

Gopal, P. R. C., & Thakkar, J. (2012). A review on supply chain performance measures and metrics: 2000-2011. *International Journal of Productivity and Performance Management*, *61*(5), 518–547. https://doi.org/10.1108/17410401211232957

Gore, J., Flin, R., Stanton, N., & Wong, B. L. W. (2015). Applications for naturalistic decision-making. *Journal of Occupational and Organizational Psychology*, *88*(2), 223–230. https://doi.org/10.1111/joop.12121

Gutierrez, D. M., Scavarda, L. F., Fiorencio, L., & Martins, R. A. (2015). Evolution of the performance measurement system in the logistics department of a broadcasting company: An action research. *International Journal of Production Economics*, *160*, 1–12. https://doi.org/10.1016/j.ijpe.2014.08.012

Hafenbrädl, S., Waeger, D., Marewski, J. N., & Gigerenzer, G. (2016). Applied decision-making with fast-and-frugal heuristics. *Journal of Applied Research in Memory and Cognition*, *5*(2), 215–231. https://doi.org/10.1016/j.jarmac.2016.04.011

Hands, D. W. (2014). Normative ecological rationality: Normative rationality in the fast-and-frugal-heuristics research program. *Journal of Economic Methodology*, *21*(4), 396–410. https://doi.org/10.1080/1350178X.2014.965907

Heap, S. P. H. (2013). What is the meaning of behavioural economics? *Cambridge Journal of Economics*, *37*(5), 985–1000. https://doi.org/10.1093/cje/bes090

Henri, J. F. (2006). Organizational culture and performance measurement systems. *Accounting, Organizations and Society*, *31*(1), 77–103. https://doi.org/10.1016/j.aos.2004.10.003

Hensher, D. A., Greene, W. H., & Chorus, C. G. (2013). Random regret minimization or random utility maximization: An exploratory analysis in the context of automobile fuel choice. *Journal of Advanced Transportation*, *47*(7), 667–678. https://doi.org/10.1002/atr.188

Hoffmann, A. O., Henry, S. F., & Kalogeras, N. (2013). Aspirations as reference points: An experimental investigation of risk behavior over time. *Theory and Decision*, *75*(2), 193–210. https://doi.org/10.1007/s11238-012-9323-6

Hoque, Z. (2014). 20 years of studies on the balanced scorecard: Trends, accomplishments, gaps and opportunities for future research. *The British Accounting Review*, *46*(1), 33–59. https://doi.org/10/1016/j.bar.2013.10.003

Houghton, C., Casey, D., Shaw, D., & Murphy, K. (2013). Rigour in qualitative case-study research. *Nurse Researcher*, *20*(4), 12–17. Retrieved from http://journals.rcni.com/journal/nr

Huck, S., & Müller, W. (2012). Allais for all: Revisiting the paradox in a large representative sample. *Journal of Risk & Uncertainty*, *44*(3), 261–293. https://doi.org/10.1007/s11166-012-9142-8

Jacob, S., & Furgerson, S. (2012). Writing interview protocols and conducting interviews: Tips for students new to the field of qualitative research. *The Qualitative Report*, *17*(42), 1–10. Retrieved from http://nsuworks.nova.edu/tqr/vol17/iss42/3

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under

    risk. *Econometrica: Journal of the Economic Society*, *47*(2), 263–291. Retrieved

    from http://www.econometricsociety.org/publications/econometrica/

Kaplan, R. S. (1984). The evolution of management accounting. *The Accounting Review*,

    59(3), 390–418. Retrieved from http://aaajournals.org/

Kaplan, R. S. (2012). The balanced scorecard: Comments on balanced scorecard

    commentaries. *Journal of Accounting & Organizational Change*, *8*(4), 539–545.

    https://doi.org/10.1108/18325911211273527

Kaplan, R.S & Norton, D. P. (1996). *The Balanced Scorecard: Translating strategy into*

    action. Boston, MA.: Harvard Business School Press.

Kim, Y. (2011). The pilot study in qualitative inquiry: Identifying issues and learning

    lessons for culturally competent research. *Qualitative Social Work*, *10*(2), 190–

    206. https://doi.org/10.1177/1473325010362001

Klein, G. (2015). Reflections on applications of naturalistic decision-making. *Journal of*

    Occupational and Organizational Psychology, *88*(2), 382–386.

    https://doi.org/10.1111/joop.12122

Klein, G., & Wright, C. (2016). Macrocognition: From theory to toolbox. *Cognitive*

    Science, *7*, 54. https://doi.org/10.3389/fpsyg.2016.00054

Knight, J., Corbett, A., Smith, C., Watkins, B., Hardy, R., & Jones, G. (2012). "What

    made me stay?" A review of the reasons student nurses enrolled in a Bachelor of

    Nursing programme completed their studies: A descriptive phenomenological

study. *Nurse Education Today*, *32*(8), e62–e65.

https://doi.org/10.1016/j.nedt.2012.03.007

Kocaoğlu, B., Gülsün, B. & Tanyaş, M. (2013). A SCOR based approach for measuring a

benchmarkable supply chain performance. *Journal of Intelligent Manuf*acturing

24, 113–132. https://doi.org/10.1007/s10845-011-0547-z

Koufteros, X., Verghese, A., & Lucianetti, L. (2014). The effect of performance

measurement systems on firm performance: A cross-sectional and a longitudinal

study. *Journal of Operations Management*, *32*(6), 313–336.

https://doi.org/10.1016/j.jom.2014.06.003

Lai, P. (2017). Academic guides: Programs: PhD study. Retrieved May 29, 2017, from

http://academicguides.waldenu.edu/formandstyle/programs/phd

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice

under uncertainty. *The Economic Journal*, *92*(368), 805–824.

https://doi.org/10.2307/2232669

Lovallo, D., Clarke, C., & Camerer, C. (2012). Robust analogizing and the outside view:

Two empirical tests of case-based decision-making. *Strategic Management

Journal*, *33*(5), 496–512. https://doi.org/10.1002/smj.962

Macquet, A. C., & Skalej, V. (2015). Time management in elite sports: How do elite

athletes manage time under fatigue and stress conditions? *Journal of

Occupational and Organizational Psychology*, *88*(2), 341–363.

https://doi.org/10.1111/joop.12105

Malhotra, S., Zhu, P., & Reus, T. H. (2015). Anchoring on the acquisition premium decisions of others. *Strategic Management Journal*, *36*(12), 1866–1876. https://doi.org/10.1002/smj.2314

Malterud, K. (2012). Systematic text condensation: A strategy for qualitative analysis. *Scandinavian Journal of Public Health*, *40*(8), 795–805. https://doi.org/10.1177/1403494812465030

Manley, E. J., Orr, S. W., & Cheng, T. (2015). A heuristic model of bounded route choice in urban areas. *Transportation Research Part C: Emerging Technologies*, *56*, 195–209. https://doi.org/10.1016/j.trc.2015.03.020

Marginson, D., McAulay, L., Roush, M., & van Zijl, T. (2014). Examining a positive psychological role for performance measures. *Management Accounting Research*, *25*(1), 63–75. https://doi.org/10.1016/j.mar.2013.10.002

Maxwell, J. A. (2012). *A realist approach for qualitative research*. Thousand Oaks, CA: Sage.

Mayoh, J., & Onwuegbuzie, A. J. (2015). Toward a conceptualization of mixed methods phenomenological research. *Journal of Mixed Methods Research*, *9*(1), 91–107. https://doi.org/10.1177/1558689813505358

Melnyk, S. A., Bititci, U., Platts, K., Tobias, J., & Andersen, B. (2014). Is performance measurement and management fit for the future? *Management Accounting Research*, *25*(2), 173–186. https://doi.org/10.1016/j.mar.2013.07.007

Michalkiewicz, M., & Erdfelder, E. (2016). Individual differences in use of the recognition heuristic are stable across time, choice objects, domains, and

presentation formats. *Memory & Cognition*, *44*(3), 454–468.

https://doi.org/10.3758/s13421-015-0567-6

Microsoft. (2017). Visio Pro for Office 365. Retrieved from

https://products.office.com/en-us/Visio/microsoft-visio-pro-for-office-

365?wt.srch=1&wt.mc_id=AID522514_SEM_tFCxF4Ju

Mousavi, S., & Gigerenzer, G. (2014). Risk, uncertainty, and heuristics. *Journal of*

Business Research, *67*(8), 1671–1678.

https://doi.org/10.1016/j.jbusres.2014.02.013

Myers, M. D. (2013). *Qualitative research in business & management* (2nd ed.). Los

Angeles, CA: Sage.

NoNotes.com. (2017). from https://www.nonotes.com/index.php/

NVivo. (2017). NVivo for Mac [www.qsrinternational.com/nvivo-product/nvivo-mac].

Retrieved from http://www.qsrinternational.com/nvivo-product/nvivo-mac

Nørreklit, H., Nørreklit, L., Falconer, M., & Bjørnenak, T. (2012). The rise of the

balanced scorecard! Relevance regained? *Journal of Accounting &*

Organizational Change, *8*(4), 490–510.

https://doi.org/http://10.1108/18325911211273491

Ntabe, E. N., LeBel, L., Munson, A. D., & Santa-Eulalia, L. A. (2015). A systematic

literature review of the supply chain operations reference (SCOR) model

application with special attention to environmental issues. *International Journal*

of Production Economics, *169*, 310–332.

https://doi.org/10.1016/j.ijpe.2015.08.008

Okoli, J. O., Weller, G., Watt, J., & Wong, B. L. W. (2013). Decision-making strategies used by experts and the potential for training intuitive skills: A preliminary study. In *Proceedings of the 11th International Conference on Naturalistic Decision Marking (NDM 2013), Marseille, France, 21-24 May, 2013* (pp. 227–232). Paris, France: Arpege Science Publishing.

Okongwu, U., Brulhart, F., & Moncef, B. (2015). Causal linkages between supply chain management practices and performance: A balanced scorecard strategy map perspective. *Journal of Manufacturing Technology Management*, *26*(5), 678–702. https://doi.org/10/1108/JMTM-01-2013-0002

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Paulraj, A., Chen, I. J., & Lado, A. A. (2012). An empirical taxonomy of supply chain management practices. *Journal of Business Logistics*, *33*(3), 227–244. https://doi.org/10.1111/j.0000-0000.2012.01046.x

Peng, X., & Prybutok, V. (2015). Relative effectiveness of the Malcolm Baldrige National Quality Award categories. *International Journal of Production Research*, *53*(2), 629–647. https://doi.org/10.1080/00207543.2014.961207

Ramos, G. M., Daamen, W., & Hoogendoorn, S. (2014). A state-of-the-art review: Developments in utility theory, prospect theory and regret theory to investigate travellers' behaviour in situations involving travel time uncertainty. *Transport Reviews*, *34*(1), 46–67. https://doi.org/10.1080/01441647.2013.856356

Riegel, B., Dickson, V. V., & Topaz, M. (2013). Qualitative analysis of naturalistic decision-making in adults with chronic heart failure. *Nursing Research*, *62*(2), 91–98. https://doi.org/10.1097/NNR.0b013e318276250c

Robinson, O. C. (2014). Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative Research in Psychology*, *11*(1), 25–41. https://doi.org/10.1080/14780887.2013.801543

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Thousand Oaks, CA: Sage.

Sato, A. (2016). Quality control strategy in Japan after World War II: Role of the TQC: Advocated by an educator, W. Edwards Deming. *Japan Studies Association Journal*, *10*, 83–101. Retrieved from http://www.japanstudies.org/journal.html

Saunila, M., Pekkola, S., & Ukko, J. (2014). The relationship between innovation capability and performance: The moderating effect of measurement. *International Journal of Productivity and Performance Management*, *63*(2), 249–234. Retrieved from http://www.emeraldinsight.com./loi/ijppm

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

Searcy, C. (2012). Corporate sustainability performance measurement systems: A review and research agenda. *Journal of Business Ethics*, *107*(3), 239–253. https://doi.org/10.1007/s10551-011-1038-z

Seuring, S. (2013). A review of modeling approaches for sustainable supply chain management. *Decision Support Systems*, *54*(4), 1513–1520. https://doi.org/10.1016/j.dss.2012.05.053

Shaban, R. (2012). Theories of clinical judgment and decision-making: A review of the theoretical literature. *Australasian Journal of Paramedicine*, *3*(1). Retrieved from http://ro.ecu.edu.au/jephc/vol3/iss1/8

Shan, Y., & Yang, L. (2016). Fast and frugal heuristics and naturalistic decision-making: A review of their commonalities and differences. *Thinking & Reasoning*, 1–23. https://doi.org/10.1080/13546783.2016.1152999

Shosha, G. A. (2012). Employment of Colaizzi's strategy in descriptive phenomenology: A reflection of a researcher. *European Scientific Journal; Kocani*, *8*(27). Retrieved from http://eujournal.org/index.php/esj

Signori, P., Flint, D. J., & Golicic, S. (2015). Toward sustainable supply chain orientation (SSCO): Mapping managerial perspectives. *International Journal of Physical Distribution & Logistics Management*, *45*(6), 536–564. https://doi.org/10.1108/IJPDLM-07-2014-0160

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Smith, J. A., & Osborn, M. (2015). Interpretative phenomenological analysis as a useful methodology for research on the lived experience of pain. *British Journal of Pain*, *9*(1), 41–42. https://doi.org/10.1177/2049463714541642

Sousa, D. (2014). Validation in qualitative research: General aspects and specificities of the descriptive phenomenological method. *Qualitative Research in Psychology*, *11*(2), 211–227. https://doi.org/10.1080/14780887.2013.853855

Spaniel, W. (2014). *Game theory 101: The complete textbook*. Charleston, SC: Create

Space.

Swink, M. L., Johnson, R. R., & Quinn, F. J. (2012). The ninth annual global survey of

supply chain progress. *Supply Chain Management: An International Journal*,

16(2). Retrieved from

http://www.scmr.com/article/9th_annual_global_supply_chain_survey/

Szpiro, G. (2013). Economics: Value judgments. *Nature*, *500*(7464), 521.

https://doi.org/10.1038/500521a

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences.

Journal of Experimental Psychology: General, *142*(1), 151–162.

https://doi.org/10.1037/a0028232

Tappeiner, F., Howorth, C., Achleitner, A. K., & Schraml, S. (2012). Demand for private

equity minority investments: A study of large family firms. *Journal of Family

Business Strategy*, *3*(1), 38–51. https://doi.org/10.1016/j.jfbs.2012.01.001

Taticchi, P., Balachandran, K., & Tonelli, F. (2012). Performance measurement and

management systems: state of the art, guidelines for design and challenges.

Measuring Business Excellence, *16*(2), 41–54.

https://doi.org/10.1108/13683041211230311

Thoma, V., & Williams, A. (2013). The devil you know: The effect of brand recognition

and product ratings on consumer choice. *Judgment & Decision-making*, *8*(1), 34–

44. Retrieved from http://journal.sjdm.org/

Tilburg, W. A. P., & Igou, E. R. (2014). Moving onwards: An action continuation

strategy in finding the way. *Journal of Behavioral Decision-making*, *27*(5), 408–

418. https://doi.org/10.1002/bdm.1817

Timmermans, S., & Tavory, I. (2012). Theory construction in qualitative research: From

grounded theory to abductive analysis. *Sociological Theory*, *30*(3), 167–186.

https://doi.org/10.1177/0735275112457914

Trochim, W. M. K., Donnelly, J. P., & Arora, K. (2014). *Research methods: The*

essential knowledge base. Boston, MA: Cengage Learning.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases.

Science, *185*(4157), 1124–1131. Retrieved from http://www.jstor.org

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative

representation of uncertainty. *Journal of Risk & Uncertainty*, *5*(4), 297–323.

Retrieved from https://www.jstor.org/journal/riskuncertainty

Von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior,*.

Princeton, NJ: Princeton University Press.

Waal, A. A. de, & Kourtit, K. (2013). Performance measurement and management in

practice. *International Journal of Productivity and Performance Management*,

62(5), 446–473. Retrieved from http://www.emeraldinsight.com/loi/jppm

Wang, Q., Kim Koh, J. B., & Song, Q. (2015). Meaning making through personal

storytelling: Narrative research in the Asian American context. *Journal of*

Psychology, *6*(1), 88–96. https://doi.org/10/1037/a0037317

Woods, V. & Van der Meulen, R. (2016, May 19). Gartner Announces Rankings of the

    2016 Supply Chain Top 25. https://www.gartner.com/en/newsroom/press-

    releases/2016-05-19-gartner-announces-rankings-of-the-2016-supply-chain-top-

    25 Wu, H.Y. (2012). Constructing a strategy map for banking institutions with

    key performance indicators of the balanced scorecard. *Evaluation and Program*

    Planning, *35*(3), 303–320. https://doi.org/10.1016/j.evalprogplan.2011.11.009

Yin, R. (2014). *Case study research: Design and methods* (Fifth). Thousand Oaks, CA:

    Sage.

Zhou, Y., Fenton, N., & Neil, M. (2014). Bayesian network approach to multinomial

    parameter learning using data and expert judgments. *International Journal of*

    Approximate Reasoning, *55*(5), 1252–1268.

    https://doi.org/10.1016/j.ijar.2014.02.00

Appendix A: Letter of Cooperation

[Community Research Partner Name]
[Contact Information]

[Date]

Dear Elizabeth Nohe,

Based on my review of your research proposal, I give permission for you to conduct the study entitled *How Supply Chain Leaders Choose the Measures that Matter: A Descriptive Phenomenological Study* within the [Insert Name of Community Partner]. As part of the study, I authorize you to connect with supply chain leaders in our organization with the title of EVP, SVP, VP, Senior Director, Director, or Senior Manager from a list of names and contact information that I will provide to you. I also authorize you to conduct initial phone interviews to gather data as well as up to two short follow-up phone meetings to confirm and to clarify information from the original interview. You may also provide results of your study to the participants. Individuals' participation will be voluntary and at their own discretion.

We understand that our organization's responsibilities include:
- A commitment that each participant will be initially interviewed by phone, uninterrupted and in a private space for 45-60 minutes.
- That the participant will be available for up to two 15-minute follow-up conversations to review the original transcript and make any needed changes and/or follow-up interactions to clarify any comments made during the interview.

It is my understanding that you, as the researcher, will adhere to the National Institute of Health's standards as stated in the Protecting Human Research Participants as well the standards set forth by Walden University's Internal Review Board (IRB).

I also understand that neither the data collected nor any published documents as a result of the study will contain names, locations, or details that could identify the participants or our organization. I also understand that the data collected will remain entirely confidential and may not be provided to anyone outside of the student's supervising faculty/staff without permission from the Walden University IRB.

We reserve the right to withdraw from the study at any time if our circumstances change.

I confirm that I am authorized to approve research in this setting and that this plan complies with our organization's policies.

Sincerely,

[Authorization Official]
[Contact Information]

Walden University policy on electronic signatures: An electronic signature is just as valid as a written signature as long as both parties have agreed to conduct the transaction electronically. Electronic signatures are regulated by the *Uniform Electronic Transactions Act*. Electronic signatures are only valid when the signer is either (a) the sender of the e-mail, or (b) copied on the e-mail containing the signed document. Legally an "electronic signature" can be the person's typed name, their e-mail address, or any other identifying marker. Walden University staff will verify any electronic signatures that do not originate from a password-protected source (i.e., an e-mail address officially on file with Walden).

Appendix B: Letter of Invitation to Participants

From: [Point of Contact Information]
Date: [Date]

Dear [Participant Name],

Our company has been invited to participate in a doctoral research study titled *How Supply Chain Leaders Choose the Measures that Matter: A Descriptive Phenomenological Study*. You have been identified as a potential participant because of your qualifications. Your decision to participate in the study is entirely voluntary. Should you choose to participate you may also decide to withdraw from the study at any time.

Your commitment would consist of participating in an initial phone interview of 45-60 minutes and potentially up to two short follow-up phone meetings to confirm and to clarify information from the original interview. The entire interview and follow up process is to be completed over a 6-8 week period.

If you are interested in taking part in the study please:
- Read and sign the attached Letter of Consent.
- Keep a copy for your records.
- Send the Letter of Consent via e-mail to:

  Elizabeth Nohe, PhD Candidate – College of Management and Technology
  Walden University
  Elizabeth.Nohe@Waldenu.edu
  Mobile: [redacted]

As a participant in the study, your privacy is guaranteed under the standards set forth by Walden University's Internal Review Board (IRB) which adheres to the National Institute of Health's standards as stated in the Protecting Human Research Participants. Although there is no remuneration for your participation, the researcher would be pleased to send you the results of the study after completion.

Thank you in advance for considering participation in Elizabeth Nohe's study.

Regards,
[Point of Contact Name]

Attachment: Consent Form

Appendix C: Interview Protocol

How Supply Chain Leaders Choose the Measures that Matter

Participant ID
Date
Time
Location

Pre-interview Instructions:
Hi, my name is Elizabeth Nohe. Feel free to call me by my nickname, Betsey. I want to start by thanking you for agreeing to participate in my PhD research. In addition to pursuing my PhD, I am the VP of Supply Chain for Morton Salt. I live and work in downtown Chicago.

Before we start the interview there are a few housekeeping items we must cover.

This interview should take between 45-60 minutes. It's important that we not be interrupted during the interview.
Is this still a good time for you to be interviewed?
[Participant responds yes or no, researcher records response.]

Let me tell you a bit about the research. I am very passionate about the topic of performance measures and metrics. As I started researching the evolution of performance measurement in the business environment in general and more specifically in supply chain, I found there was a lot of research telling supply chain leaders what to measure; for example frameworks such as the SCOR model or the Gartner Hierarchy of Supply Chain Metrics.

But I couldn't find any research on how senior managers actually chose their measures and metrics. I decided that an interesting and new approach to the topic could be interviewing leaders in supply chain, such as you, and then analyzing the interviews through the lens of decision-making theory. And thus was born my study *How Supply Chain Leaders Choose the Measures that Matter: A Descriptive Phenomenological Study*. I hope that I can find some common decision-making approaches among those of you I'm interviewing that could provide new insights and help other Supply Chain leaders as they design or re-design their respective performance measurement systems. .

I previously e-mailed to you an Informed Consent form that you signed and returned to me. I want to confirm again that you are a willing participant in the study.
Are you in agreement with participating in the study?
[Participant responds yes or no, researcher records response.]

I also want to make sure that you understand that as a participant in the study confidentiality is guaranteed. All documents in the study will only refer to you by your participant ID number. I do have a document that maps your name to your participant ID and that document is password protected and stored on an external drive.
Are you in agreement that you believe your privacy will be assured? [Participant responds yes or no, researcher records response.]

In order to be sure I capture your responses completely and accurately I plan to record this interview. The interview will then be transcribed by an outside service called NoNotes. The transcription will be sent to me electronically. I will then make sure there is nothing in the transcript that could be construed as identifying information about you or your company. I will also delete the NoNotes audio file. Then I will send the sanitized transcript to you for your review.
Is this process acceptable to you?
[Participant responds yes or no, researcher records response.]
If the response is no then discuss a method to gather interview responses in a mutually agreeable manner.

As I mentioned earlier, the purpose of this interview is to explore how you, as a Supply Chain Leader, select the most important key performance measures and goals for your supply chain. I want to make sure you know that you may choose at any time to stop this interview or to withdraw from the study.
Would you please confirm your understanding?
[Participant responds yes or no, researcher records response.]

Do you have any questions for me?
[Participant responds yes or no, researcher records response.]

Ok, then let's begin!

Demographic Questions:

1. What is your title?

2. What gender do you identify with? [select female, male, or prefer not to answer]

3. How many years have you been a supply chain professional?

4. How many years have you been in your current position?

5. What is your role in selecting the key performance measures and goals for your organization [select Responsible, Accountable, Consulted, or Informed]?

Interview Questions:

1. What is the most important key performance measure or key performance indicator (KPI) for your supply chain organization?

2. For clarification, please define [insert name of KPI] for me.

3. Please tell me, with as much detail as possible, about the process you used to decide on [insert name of KPI] as a key performance measure for your organization. [Make sure to collect not only process but also context.]

   a. Follow on questions could include:
      i. Who is responsible and accountable for the KPI
      ii. Did others participate in the process? If so, please tell me how.
      iii. What other factors might have influenced the process of selecting this KPI?
         • Prior experience
         • Business context
         • External environment
         • Frequency of review and/or refresh
         • Formal or informal process

   b. Once the description seems to have been exhaustive respond, "Thank you, that description was wonderful and very insightful!"

4. Have you determined a goal or target for [insert name of KPI]?

   a. If **YES** then:
      i. Please describe, with as much detail as possible, how you came to decide on the goal of [insert goal].

   b. Follow on questions could include:
      i. Did others participate in the process? If so, please tell me how.
      ii. What other factors might have influenced the process of selecting this goal?
         • Prior experience
         • Business context
         • External environment
         • Benchmarking activity
         • Frequency
         • Formal or informal process

   c. If **NO** then:

       i. Please describe why you didn't establish a goal for [insert name of KPI].

5. Please tell me, with as much detail as possible, how you use [insert name of KPI] in your organization.
   a. Specific improvement strategies tied to KPI
   b. How broadly shared within organization
   c. Formal or informal performance review process

6. If you could change something in the design and/or use of [insert name of KPI] what would you change and why?
   a. May lead to follow on questions such as:
      i. What do you think would be the impact of implementing that change?
      ii. Are there any obstacles to implementing that change?

7. Was there anything I did not ask you but you wanted to be sure that I know?

Post-interview Process

I'd like to go over next steps with you now.

In a few days, I will e-mail you a copy of the sanitized transcript of this interview. Please read through it and let me know either that the interview and the transcript captured your experiences completely and accurately or let me know what changes I should make to the transcript.

Would it be possible for you to read and respond within 48 hours of receiving my e-mail with the transcript attached?
[If no, then come up with an agreed upon timeline.]

If possible I'd like to set up a 15-minute follow up appointment to review and make any changes to the transcript. This way we both have it on our calendars.
[Agree to a date/time. Immediately after send a follow up invitation.]

The goal of a descriptive phenomenological study is to find common elements among the participant group on their experience in selecting key performance measures and goals. Once the interviews are complete I will analyze the information from the interviews through the lens of decision-making theory. Then I will write up the findings and make recommendations that can be used both by other supply chain practitioners and other researchers interested in performance measurement. I would be happy to review the results with you too.
Would you like an executive summary of the study?
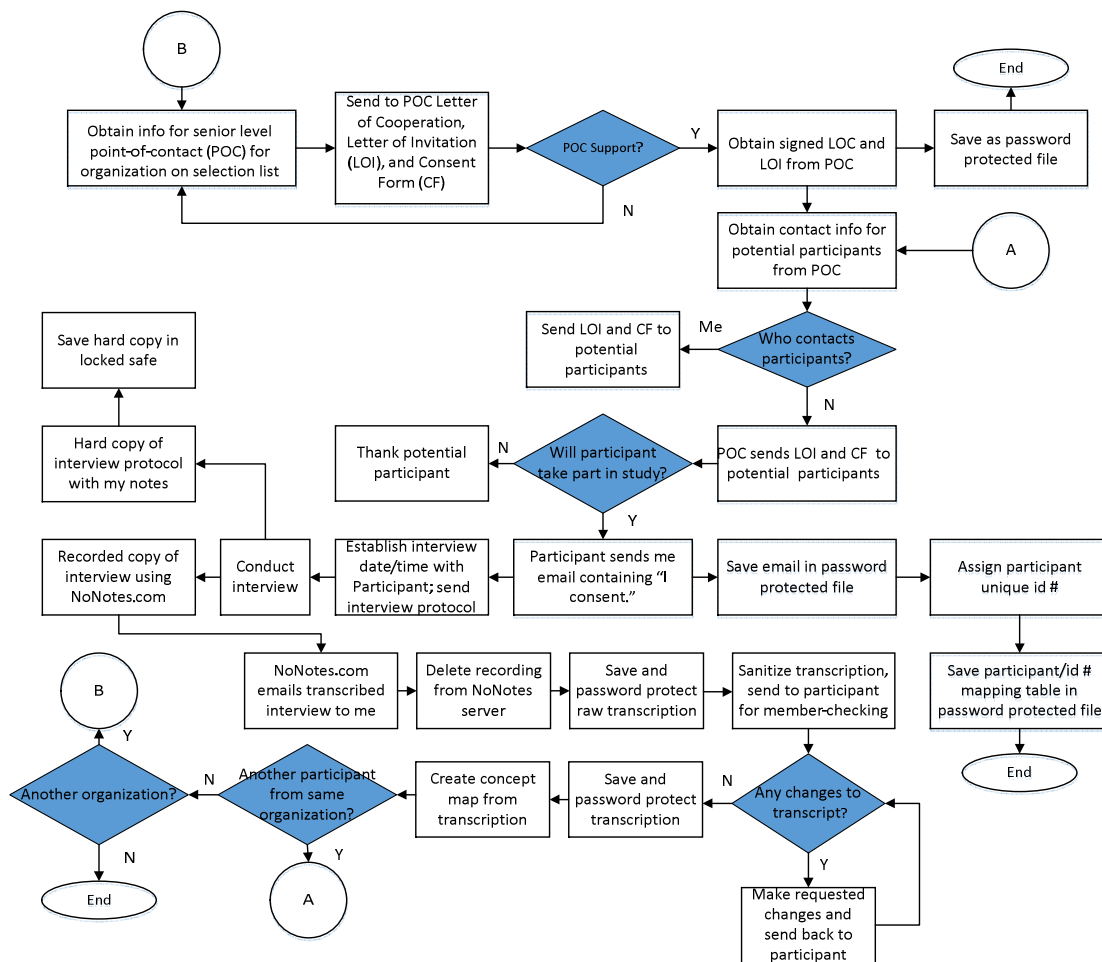[Participant responds yes or no.]

Before we close for today, let me give you my e-mail address (Elizabeth.Nohe@waldenu.edu) and my cell phone number (redacted) in case you want to contact me.

Thank you again for taking time out of your busy schedule for this interview. I really appreciate it. Have a great day!

Appendix D: Recruitment and Data Collection Process Diagram

Appendix E: Audit Trail Showing Status of Data Collection Process by Participant

Participant ID	Pilot	Interview	Recording	Transcript	Sanitized transcript sent	Member-checked transcript	Entered in Nvivo
P1	Y	Y	Y	Y	Y	Y	Y
P2	Y	Y	Y	Y	Y	Y	Y
P3	N	Y	Y	Y	Y	Y	Y
P4	N	Y	Y	Y	Y	Y	Y
P5	N	Y	Y	Y	Y	Y	Y
P6	N	Y	Y	Y	Y	Y	Y
P7	N	Y	Y	Y	Y	Y	Y
P8	N	Y	Y	Y	Y	Y	Y
P9	N	Y	Y	Y	Y	Y	Y
P10	N	Y	Y	Y	Y	Y	Y
P11	N	Y	Y	Y	Y	Y	Y
P12	N	Y	Y	Y	Y	Y	Y
P13	N	Y	Y	Y	Y	Y	Y
P14	N	Y	Y	Y	Y	Y	Y
P15	N	Y	Y	Y	Y	Y	Y

Note. P1 and P2 were the pilot participants and also included in main study.

Appendix F: Q3 Number of References to Each Decision-Making Theory by Participant

Participant ID	Fast and frugal heuristics	Heuristics and biases	Naturalistic decision-making	Rational decision-making
P1	1	4	6	3
P2	1	13	18	21
P3	2	0	5	6
P4	1	6	15	13
P5	2	2	8	10
P6	1	0	6	13
P7	1	1	8	16
P8	0	2	11	9
P9	2	0	4	4
P10	1	0	1	0
P11	2	0	4	1
P12	1	0	3	3
P13	0	0	2	0
P14	1	0	10	7
P15	0	0	3	2

Note. Raw data showing number of meaning units by decision-making theory by participant.

Appendix G: Q4 Number of References to Each Decision-Making Theory by Participant

Participant ID	Fast and frugal Heuristics	Heuristics and biases	Naturalistic decision-making	Rational decision-making
P1	0	4	5	5
P2	0	2	1	2
P3	0	1	3	2
P4	0	3	5	3
P5	0	1	1	1
P6	0	1	3	3
P7	0	1	5	3
P8	0	0	0	1
P9	1	1	2	1
P10	2	2	2	0
P11	1	2	4	1
P12	0	1	3	2
P13	2	2	4	2
P14	1	0	5	1
P15	1	1	0	0

Note. Raw data showing number of meaning units by decision-making theory by participant.