Journal of the Association for Information Systems (2021) **22**(5), 1224-1244
doi: 10.17705/1jais.00713

RESEARCH ARTICLE

ISSN 1536-9323

How Big Data Analytics Affects Supply Chain Decision-Making: An Empirical Analysis

Daniel Q. Chen[1], David S. Preston[2], Morgan Swink[3]

[1]M. J. Neeley School of Business, Texas Christian University, USA, d.chen@tcu.edu
[2]M. J. Neeley School of Business, Texas Christian University, USA, d.preston@tcu.edu
[3]M. J. Neeley School of Business, Texas Christian University, USA, m.swink@tcu.edu

Abstract

This study investigates *how* different types of "big data analytics" (BDA) usage influence organizational decision-making in the area of supply chain management (SCM). Drawing on decision-making theory and organizational information processing theory, we conceptualize two patterns of BDA usage for supply chain (SC) activities (BDA use for SC optimization and BDA use for SC learning) and report two complementary channels via which the two BDA usage patterns impact a supply chain organization's BDA-enabled decision-making capability. An analysis of questionnaire data from supply chain managers representing 157 companies based in North American suggests that BDA use for SC optimization is directly associated with better decision-making capability. In contrast, the influence of BDA use for SC learning does not impact decision-making directly but indirectly, as its effect is fully mediated by organizational integration. We discuss the implications of these findings for future academic research and for managers in practice who seek to maximize business values from BDA implementations.

Keywords: Big Data Analytics, Supply Chain, Information Processing, Decision-Making, Internal Integration

Jack Jiang was the accepting senior editor. This research article was submitted on February 21, 2018 and underwent three revisions.

1 Introduction

Many organizations today are confronted with the rapidly growing volume, variety, velocity, and limited veracity of data in both structured and unstructured forms (Abbasi et al., 2016; Agarwal & Dhar, 2014; Grover et al., 2018; Tallon, 2013). The increasing availability of "big data" presents both opportunities and challenges. It has fueled advances in an array of powerful analytical techniques and information technologies (i.e., big data analytics) (Chen et al., 2012), which purportedly enable companies to capitalize upon new insights into demand, operational processes, asset productivity, and environmental conditions. Growing anecdotal evidence suggests that firms can exploit big data analytics (BDA) to create advantages over competitors (Davenport, 2006). However, many companies appear to be slow in exploiting big data, because they do not know its true value or because they lack the organizational capabilities needed to unlock its potential (Tallon, 2013).

Though the practitioner literature offers compelling testimonials about the success of BDA first-movers (Davenport, 2006), empirical validation linking BDA usage to improved organizational decision-making remains limited. Furthermore, more needs to be known about the means by which BDA impacts business outcomes (Abbasi et al., 2016). We note two salient research gaps. First, extant IT management literature surprisingly lacks theory-driven, contextually based studies of how BDA usage improves an organization's decision-making capability. There have been calls for research of BDA's business value and role in supporting more effective decisions (Abbasi et al., 2016; Kitchens et al., 2018). Second, existing empirical studies of BDA impacts on organizational outcomes tend to treat BDA monolithically (Chen et al., 2015; Müller et al., 2018), though different structures and applications of BDA exist (Choi et al., 2018). An examination of different *patterns* of organizational BDA usage and their impacts on organizational decision-making capability may be able to highlight ways that managers might better exploit BDA for value creation.

The current study aims to bridge these research gaps by developing theoretical and empirical evidence in the area of supply chain management (SCM). SCM offers a rich opportunity for the examination of the BDA phenomenon. Organizational issues regarding BDA are particularly relevant because of the cross-functional and boundary-spanning nature of SCM activities, which are generally characterized as complex and data intensive. Researchers have recently recognized the potential of BDA to revolutionize supply chain dynamics, including both how supply chains are designed and managed, thus prompting calls for theory-based empirical studies to investigate the outcomes of BDA applications to SCM decision-making (Sanders & Ganeshan, 2015; Waller & Fawcett, 2013).

Inspired by the classic decision-making framework of Gorry and Scott Morton (1971), we first conceptualize two patterns of BDA usage in the context of SCM corresponding to the two types of decisions (structured and unstructured) that firms conduct on a regular basis. Typically, structured decisions for SCM can be described as *optimization*, which seeks to maximize the efficient use of SCM resources (e.g., network design, production scheduling, and inventory management, etc.) (Choi et al., 2018). On the other hand, unstructured, *learning*-oriented decisions in SCM seek to uncover novel insights regarding transactions, frequencies, exceptions, or other patterns describing supply chain processes (e.g., sourcing, purchasing, customer services, etc.) (Choi et al., 2018). Accordingly, we distinguish between (1) BDA usage for SC optimization, and (2) BDA usage for SC learning.

Our research objectives are to examine *whether* and *how* these two types of BDA usage improve a supply chain organization's *BDA-enabled decision-making capability*, defined as the extent to which managers are able to execute supply chain decision-making quickly and effectively with the help of advanced BDA technologies (Grover et al., 2018; Joshi et al., 2010; Roberts and Grover, 2012). We focus on BDA-enabled rather than general decision-making capability for several reasons. First, an organization's overall decision-making capability could be constrained by its level of IT resources. Firm-level decision-making capability describes an organization's capacity to select a course of action in preference to a number of alternatives to help the firm maximizing the benefits brought by opportunities and minimizing the risks of threats (Thomas et al., 1993). Based upon the idea of bounded rationality (Simon 1955), most businesses cannot obtain all the information required for utility-maximizing calculations, either because it is too costly or because it is cognitively unfeasible to do so. As such, it is difficult to directly assess such organizational capability without considering the information accessible to the organization. Based on this perspective, the IT management literature has provided empirical evidence suggesting that IT can be leveraged to enable firms to sense and respond to opportunities and threats (Chakravarty et al., 2013; Roberts & Grover, 2012). Second, IT-enabled decision-making capability itself is a broad concept and the IS literature has defined many different types of IT in terms of how they support specific business tasks (Goodhue & Thompson, 1995). Recent literature has called for more work with high levels of granularity to understand the relationship between specific IT components and related business tasks (Park et al., 2017). BDA technologies (e.g., dashboard applications, sensors, etc.) not only allow firms to store and manage codified knowledge and business rules, but also provide individuals with direct access to enterprise-wide databases (e.g., data warehouse) and features like what-if analyses, data explorations, and visualizations. These BDA tools provide firms the capacity to automatically monitor and keep watch for important business events and have been recognized as essential to enable timely decision-making. As such, the current study focuses on examining *how* supply chain organizations' usage of BDA tools may improve their BDA-enabled decision-making capability.

Drawing upon organizational information processing theory (Galbraith 1973), we propose two channels leading to BDA-enabled decision-making capability: a direct path, which we consider to be more relevant for BDA use for SC optimization, and a mediated path, in which internal cross-functional process integration serves as a generative means for BDA impact. We expect the mediation effect to be more relevant for BDA use for SC learning, as learning processes are often embedded within a firm's supply chain organization (Schoenherr & Swink, 2012; Williams et al., 2013).

To uncover empirical evidence regarding our hypotheses, we designed a questionnaire used to collect data from supply chain managers representing 157 North America-based companies. The remainder of this paper describes the research model and theoretical support for the testable hypotheses, the methodology including the data collection process, the measures, and the survey sample, the results of our empirical analysis, and the discussion of the research findings, including the contributions, the limitations of the research design, and implications for future research and practice.

2 Theoretical Background and Research Hypotheses

Some technology-related components of BDA have been in place within organizations for decades, leading some executives to question whether BDA is simply "old wine in a new bottle" (Agarwal & Dhar, 2014). In order to more effectively articulate the value propositions of BDA, we provide a brief summary of the distinguishing characteristics of BDA, and then describe the theoretical streams shaping the research model and hypotheses (Figure 1).

2.1 The Opportunities Provided by Big Data Analytics

Big data is an umbrella term that describes data sets that are large and complex. Over the last decade, major contributors to the extremely rapid growth of vast quantities of both structured and unstructured data include online trade and social transactions, as well as data produced by an assortment of sensors and devices. Such data can be difficult to capture, process, and manage in a timely fashion using conventional data management and intelligence systems and tools (Chen et al., 2012). Conventional systems generally have the capability to capture structured data into database management software for processing. However, the requirements for processing large amounts of unstructured data often exceeds the capacity of conventional database systems because the data set is too big, moves too fast, or doesn't fit the structure of existing database architectures (Abbasi et al., 2016; Grover et al., 2018). In order to gain value from such data, an organization must seek alternative ways to process them.

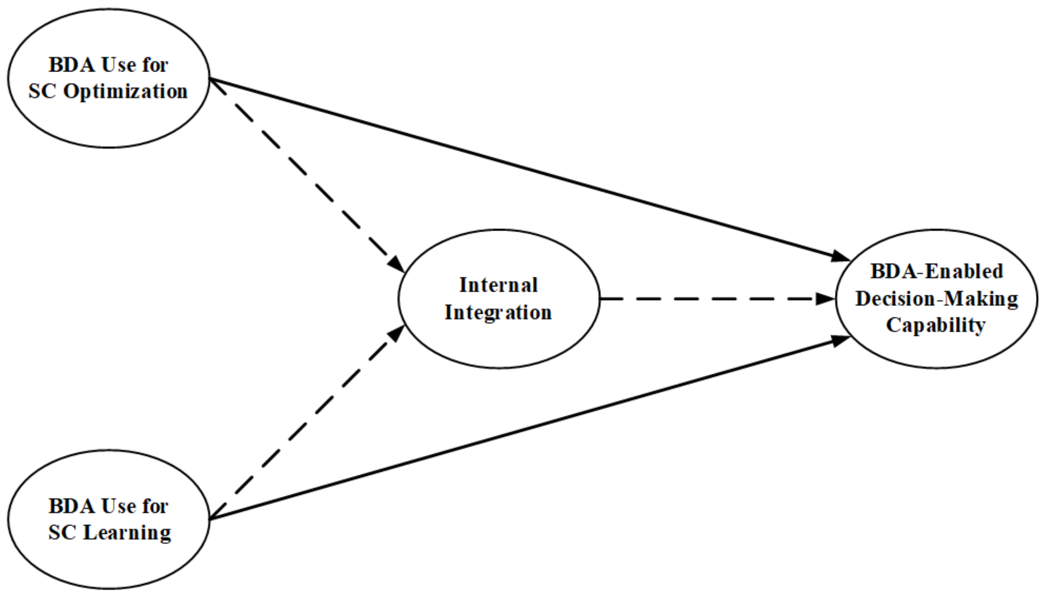
Advancements in innovative data processing and analytical technologies are often labeled as big data analytics. BDA technologies are generally categorized as data management (e.g., massively parallel-processing databases), open-source programming (e.g., Hadoop, MapReduce), statistical analysis (e.g., sentiment analysis, time-series analysis), solution generation and optimization tools (e.g., simulation, scenario evaluation, and optimization algorithms), visualization tools that help structure and connect data to uncover hidden patterns, anomalies, correlations, and other actionable insights, and in-memory computing (IMC) (e.g., SAP's HANA) (Chen et al., 2012; Grover et al., 2018). BDA technologies have considerably expanded opportunities for inquiry and understanding of human discourse as the dimensionality of data sets available has grown.[1] More importantly, the nature of these inquiries has also changed, fueled by computing machines that combine massive computing power with greater "intelligence" through better-designed algorithms and new information technologies that allow decision-makers (both humans and systems) to be more inherently informed through observations and interactions. Such engagements with data prompt new questions and feed new algorithms, which can be continuously refined with little or no human intervention (Agarwal & Dhar, 2014). Meanwhile, computer systems have become active question-generating machines rather than merely question-processing tools. In other words, data are not collected only to test a priori hypotheses or to evaluate solutions to targeted problems; data are also used to develop observations and patterns not envisioned at the time of data collection. Hence, BDA tools play more significant roles in knowledge discovery and decision-making (Agarwal & Dhar, 2014).

Although there is the potential for improved decision-making capability, empirical evidence supporting this assertion is still quite limited (Grover et al., 2018). The extant SCM literature emphasizes the role of information seeking and sharing to cope with uncertainty in increasingly complex business environments (Srinivasan & Swink, 2015). Researchers have also begun to examine the drivers of BDA adoption in a supply chain context (Chen et al., 2015). Our study examines the impact of the usage of BDA as an enabler of the supply chain organization's decision-making capability. We view these relationships through the lens of two classic organizational theories that address decision-making (Gorry and Scott Morton, 1971; Simon, 1960) and information processing (Galbraith, 1973), respectively.

2.2 The Gorry and Scott Morton Decision-Making Framework

Among existing decision-making theories, the Gorry and Scott Morton (1971) framework is perhaps one of the best known and most frequently cited in the IS field. Gorry and Scott Morton's framework is particularly well suited to guide our study because it "is one for managerial activities, not for information systems. It is a way of looking at decisions made in an organization" (Gorry and Scott Morton, 1971, (p. 56). The framework is, in essence, a mapping of Robert Anthony's hierarchy of management activities (i.e., operational control, management control, and strategic planning) (Anthony, 1965) across Herbert Simon's classification of decision-making (i.e., programmed and nonprogrammed decisions) (Simon, 1960). Gorry and Scott Morton (1971) replace Simon's "programmed" and "nonprogrammed" decisions terminology with the terms "structured" and "unstructured," because structured/unstructured decisions "imply less dependence on the computer and more dependence on the basic character of the problem-solving activity in question" (p. 56).

[1] For example, BDA allows online retailers to develop algorithms to predict what products individual customers would like to purchase next by tracking not only what they bought, but also their browsing history, their response to the promotions, reviews, recommendations, and web page design, as well as the similarities and differences across customer groups.



Note: Solid lines represent direct impacts of BDA *usage on decision-making capability* while dashed lines represent mediated effects via *internal integration*.

Summary of hypotheses: **H1:** BDA usage for SC optimization has a stronger positive impact on BDA-enabled decision-making capability than BDA usage for learning. **H2:** Internal integration has a stronger positive mediating effect to the relationship of BDA usage for SC learning and BDA-enabled decision-making capability than to the relationship of BDA usage for optimization and BDA-enabled decision-making capability.

Figure 1. Research Model

According to Simon (1960), there are three stages in human decision-making: (1) search for the problem, (2) design alternative solutions, and (3) select the best choice. In structured decision-making, all three stages can be specified through predefined algorithms, or decision rules (Gorry and Scott Morton, 1971). An example of structured decisions in the SCM domain is inventory optimization, where stocking decisions may be optimized through the application of simple economic order quantity (EOQ) formulas or highly complicated optimization programs, depending on the needs of the system. In contrast, unstructured decision-making typically does not follow an established procedure in that few or none of the stages may be structured. Rather than finding a solution to a known problem, an unstructured analysis might have opportunity identification as its goal. Decision makers must often apply judgment and domain knowledge to evaluations and the development of insights into problem definitions. Examples of unstructured SCM activities include searching for patterns or anomalies in commodity spending data, transportation data, in supplier or customer transactions, or relationship data. Insights developed in such analyses might suggest process improvements, new product ideas, or value-adding service offerings.

The Gorry and Scott Morton framework also recognizes the central role of information systems to facilitate decision-making within the organization.[2] In the next section, we propose a new typology based upon the Gorry and Scott Morton framework to conceptualize BDA usage in supply chain management.

2.3 BDA Usage for SC Optimization and SC Learning

Supply "chains" consist of complex networks of interdependent processes and organizations (Williams et al., 2013). Material and information flows are subject to numerous resource constraints that interact in unpredictable ways. As such, SCM decision-making requires the analytical processing of large amounts of data that describe sales, customers, processes, inventories, locations, asset conditions, schedules, and planned actions (Srinivasan & Swink, 2015). The combinatorial complexity of supply chain data can be massive, even for a moderately sized firm, given the interactions of time periods, locations, asset types, product variants, and organizational partners both internal and external to the firm. Moreover, data are often captured and organized differently by separate organizational functions and partners in a supply chain, necessitating large volumes of data cleaning, translation, and standardization to support analysis (Srinivasan & Swink, 2018).

[2] More specifically, during their era (i.e., early 1970s), Gorry and Scott Morton referred to the information systems supporting structure or unstructured decision-making as "structured decision systems" and "decision support systems," respectively.

The use of BDA tools likely supports SCM decision-making by automating the construction, processing, and manipulation of more comprehensive data sets. In addition, embedded intelligence in some BDA tools can offer new insights to managers. However, the overall assumption that BDA usage improves managerial decision-making has not been empirically supported. Importantly, prior studies treat organizational BDA usage as a monolithic construct (Chen et al., 2015; Müller et al., 2018), neglecting more nuanced views of the varying impacts that BDA may bring to businesses.

The framework of structured/unstructured managerial decisions (Gorry and Scott Mortonm, 1971) offers a useful theoretical lens[3] to examine how BDA can be used for facilitating different SCM activities (Choi et al., 2018). On the one hand, structured decisions in SCM generally involve a variety of *optimization* problems, such as inventory control, production scheduling, network design, etc. (Boone et al., 2018; Choi et al., 2018). As reflected by Gorry and Scott Morton (1971), a key characteristic of such optimization problems is a lack of ambiguity with regard to the goals sought. In addition, these problems are considered "organization independent" (i.e., the essence of the problems tends to remain the same across businesses). On the other hand, unstructured decisions in SCM represent the organizational need of *learning* new initiatives, such as exploring possible changes to transactions and relationships with suppliers and customers and locating opportunities for new products and processes (Cox, 2015; Kumar & Routroy, 2016; Wong & Wei, 2018). In contrast to optimization problems, there is generally a lack of routine procedures for dealing with learning situations. Often there exists equivocality in the problem definition because of the absence of formalization of the decision-making phases. Furthermore, contrary to optimization, the processes of learning and making unstructured decisions are organization dependent (Gorry and Scott Morton, 1971).

Although the two types of SCM activities can be correlated (i.e., a broad exploration of opportunities accomplished via *learning* may highlight patterns or anomalies that could provide insights for the development of *optimization* models), they command different facets of BDA tools and, more importantly, different analytical approaches to *use* in the context SCM. SC optimization activities generally use algorithms to evaluate alternative solutions to highly structured problems, using highly structured data. Almost all optimization efforts in the SC context create models (i.e., structure a solution space) using the same three types of elements: an objective function, variables, and constraints (Dantzig, 1963). Advanced mathematical programming algorithms evaluate the solution space to identify the solution(s) that maximize or minimize an objective function.[4] Any "learning" in this application of BDA is limited to developed insights regarding the sensitivity of the optimal solution(s) to parameter values, the binding or nonbinding nature of certain constraints, and so on. Given the discrete nature of SC planning elements (locations, physical units, chunks of capacity, defined time buckets, etc.), most optimization approaches follow this general process.

In contrast, the SC learning approaches that make use of BDA tools are broader in the types of questions asked and in the tools applied;[5] they may analyze structured or nonstructured (nonprogrammable) questions using structured or unstructured data (Choi et al., 2018). A supply chain manager attempting to learn from analyzing supplier relationship management data, for example, might seek to identify suppliers that are best/worst performing in multiple dimensions, determine how spending is allocated across suppliers of different types, or identify anomalies and variances that may offer opportunities for improvement. It is true that these types of analysis could also be used to estimate certain parameters (e.g., transportation cost per pound-mile) that could then be fed into the development of an optimization model, but again this is BDA of a distinctly different character.

Applying these concepts, we distinguish between two patterns of BDA usage within the SCM context. BDA usage for SC optimization refers to the extent to which an organization uses BDA to identify solutions for established objectives. BDA usage for SC learning describes the extent to which an organization uses BDA to identify new opportunities for supply chain management. This typology of optimization/learning is also consistent with another classic framework in the IS literature that evaluates the purpose of information to provide organizational benefits with the intent to automate vis-à-vis inform (Zuboff 1985). Zuboff was among the first scholars to acknowledge the duality of intelligent technologies: (1) to automate operations by increasing the continuity and control of work processes, and (2) to inform decision makers by creating information that may help innovate existing work processes. Whereas Zuboff (1985) cited the database as a primary example of intelligent technologies at the stage of technical development, we contend that this framework remains relevant to today's advanced BDA technologies. In particular, the purpose of BDA for optimization is to make existing supply chain processes more efficient,

[3] We are indebted to an anonymous reviewer for suggesting this theoretical view.

[4] For example, production optimization may seek the optimal schedule for production over a time horizon that minimizes a combination of production, inventory, and transaction costs.

[5] Typical techniques for SC learning include visual inspection, sorting and ranking, cluster analysis, discriminant analysis, regression, and other multivariate analysis tools.

while the purpose of BDA for learning is to facilitate the generation of new knowledge for innovation. Next, we turn to organizational information processing theory (OIPT) to explain the potential effects of these two patterns of BDA usage to respectively improve SCM decision-making.

2.4 Organizational Information Processing Theory

In accordance with OIPT (Galbraith 1973), an organization has two options with which to align its information processing capability with its information needs. The first option is to reduce the need for information processing through the creation of slack resources (e.g., excessive safety inventories) and self-contained tasks. However, as today's intensified competition has raised opportunity costs associated with these tactics, firms are alternatively likely to pursue the second option proposed by OIPT, which is to increase organizational information processing capacity. The availability of more high-quality data, coupled with advances in BDA technologies, has extended the limits of information processing capacity, thus making this a more applicable and relevant option for managers (e.g., especially supply chain managers).

In order to develop information processing capacity, Galbraith (1973) advises firms to both: (1) invest in technologies that enable greater collection and distribution of information (i.e., in Galbraith's term, vertical information systems[6]); and (2) create lateral relations within the organization (i.e., internal integration across functional units and business processes). Galbraith's theory has been widely used as a foundation for understanding the effects of IT for various organizational phenomena, including supply chain management (Peng et al., 2014; Srinivasan & Swink, 2015; Srinivasan & Swink, 2018). Based upon the two alternatives suggested by OIPT to improve organizational information processing capacity, we propose two different pathways that describe how BDA may affect supply chain decision-making.

2.5 Impacts of BDA Usage to BDA-Enabled Decision-Making Capability

2.5.1 The Direct Impacts

The first pathway is a direct one. As outlined above, OIPT suggests that BDA serves as a form of vertical information system that will directly improve organizational information processing capacity (Galbraith 1973). The popular press offers examples describing how companies have successfully implemented BDA tools to facilitate real-time decision-making for the management of supply chains. For example, P&G's analytics capability allows it to determine the best way to reroute trucks when necessary and still meet other delivery commitments to customers (Davenport, 2013). Similarly, Western Digital Corporation has integrated real-time dashboards into all managerial decision-making levels across its supply chain, stretching from the US to Asia (Galbraith, 2014). However, few empirical works have tested this direct impact of BDA on decision-making capability.

Although early management support systems (MSS) literature provides empirical evidence that extensive use of the conventional MSS can improve organizational decision-making capability (Leidner and Elam, 1995; Sabherwal & Sabherwal, 2005), studies of BDA usage are only just now emerging (Abbasi et al., 2016). As mentioned earlier, there are some unique attributes of BDA tools that distinguish them from traditional MSS. For example, a common characteristic of conventional MSS is their aim to improve individual-level decision-making (Clark et al., 2007). In contrast, the decision scope of BDA is often functional or enterprise-wide (Agarwal & Dhar, 2014; Galbraith, 2014; Waller and Fawcett, 2013). In addition, unlike most conventional MSS (e.g., decision support systems, executive information systems, knowledge management systems), BDA has the capacity to process not only structured but also unstructured data (e.g., data generated from the web and social media). Importantly, such unstructured data have recently become significant drivers for decision-making across supply chain activities (Waller & Fawcett, 2013).

As explained at the beginning of the paper, we focus on examining the effects of using BDA to improve the supply chain group's BDA-enabled decision-making capability. Today's supply chain environments manifest increasing levels of velocity and turbulence (Srinivasan & Swink, 2018). Prior research shows that organizational performance in high-velocity environments is directly tied to managerial decision-making (Bourgeois & Eisenhardt, 1988). Meanwhile, SCM decision-making demands data that are comprehensive, accurate, and timely (Rai et al., 2006). These circumstances create a tension between the need for highly rational and analytical decision processes and the need to make decisions quickly.

According to OIPT, extensive use of IT for data collection and distribution will greatly improve an organization's information processing capacity and therefore decision-making quality (Galbraith 1973). A supply chain organization that makes comprehensive use of BDA technologies will apply them in both optimization activities (e.g., inventory control, product delivery, network design, etc.) and learning activities (strategic analysis, customer service, etc.). Using such

[6] In particular, Galbraith describes the need for firms to condense the flow of information by developing special languages (e.g., accounting procedures) and computer systems, and he refers to these systems as vertical information systems.

BDA tools enables businesses to process large volumes and varieties of data in real time, thus expanding decision-making beyond human cognitive limitations, and, to the extent that BDA serves to integrate data from multiple sources, perhaps beyond the limitations imposed by functionally myopic analyses that impede an organization's capability to search for globally optimal solutions (Galbraith, 1973; Levinthal & March, 1993; March, 1991). As such, we expect that both patterns of BDA use in the SCM domain produce direct positive impacts to the supply chain organization's BDA-enabled decision-making capability.

2.5.2 The Mediating Role of Internal Integration

The second pathway we identify is indirect, in terms of decision-making capability. As discussed above, OIPT suggests that an alternative approach to improving information processing capacity is through the creation of lateral relationships (e.g., cross-functional teams, liaison and integrating roles, etc.). Along with functional myopia (Levinthal & March, 1993), agency problems (Eisenhardt, 1989) within organizations often lead functional units to exhibit conflicting self-interests. Lateral relationships encourage information sharing and joint decision processes that cut across lines of authority and move decision-making down in the organization to where the information exists, without the need to reorganize existing business functions (Galbraith, 1973). In doing so, the establishment of lateral relationships serves to override the limitations of decision-making imposed by myopia, agency problems, and political conflicts among and across functional units.

Internal functional integration, a noted form of lateral relationships, has received a substantial amount of attention in the SCM literature because of the increasing need of supply chain organizations to cope with environmental uncertainty. For example, scholars have examined a variety of topics regarding the implications of internal integration regarding supply chain value creation, including firm-level outcomes such as supplier and customer integration (Zhao et al., 2011), time-based and firm-level performance (Droge et al., 2004), and competitive advantage (Schoenherr & Swink, 2012; Williams et al., 2013). Consistent with the spirit of the extant literature, we define internal integration as the extent to which an organization fosters cross-functional collaboration and information sharing through interconnected and synchronized processes and routines (Schoenherr & Swink, 2012; Williams et al., 2013; Zhao et al., 2011). We specifically examine the mediating effect of internal integration on the relationship between organizational BDA usage and the supply chain group's decision-making capability.

There are fundamental reasons why the employment of BDA can contribute to internal integration. Research suggests that rational decision-makers prize greater amounts of information, which can also produce greater diversity in the viewpoints and alternatives to be evaluated (Eisenhardt & Zbaracki, 1992). One reason for such diversity of alternatives is based on the nature of organizations as coalitions of individuals with conflicting preferences (March, 1962). The other reason, as highlighted in extant IT literature, is the absence of information transparency or data integration (Goodhue et al., 1992). For businesses, the value of data is more thoroughly assessed when different sources of data are combined, integrated, and used. In addition, firms not only need high-quality data, but also advanced analytical tools and human talents to generate valuable insights and knowledge for timely decision-making.

BDA can serve these needs not only by accelerating the collection and analysis of information but by synergistically providing transparent data access across an organization (Kitchens et al., 2018). Furthermore, novel insights provided by BDA have the potential to align the goals of managers, who may initially have biased interests within the decision-making process. For example, dashboards can provide real-time access to the well-being of various activity systems within a firm (Davenport, 2013). Such abilities of BDA to generate valuable information and disseminate this information widely across a firm not only allows consistency in viewing the data but also facilitates more complete visibility of the firm's business processes and outcomes (Grover et al., 2018).

Thus, as an organization implements and uses BDA more extensively, a high level of internal integration will help its decision makers develop more comprehensive and holistic analyses and decisions. In this regard, internal integration can be viewed as an additional layer of information processing that uses BDA-generated outputs as its inputs. Accordingly, we posit that internal integration positively mediates the relationships between BDA use and supply chain decision-making capability.

2.5.3 Differential Impacts of BDA Use for SC Optimization vis-à-vis BDA Use for SC Learning

Although there is substantial support in the extant literature indicating that both patterns of BDA usage influence decision-making capability, such influence may be exercised through dissimilar pathways because of the differential nature of these two BDA usage patterns. As such, we posit that the ways that the two BDA uses impact BDA-enabled decision-making capability vary significantly. As described earlier, SC optimization activities are routinized, have specified objectives and formalized decision rules to follow, and are organization-independent (Gorry and Scott Morton, 1971). In contrast, approaches for enabling learning are less well-established. Furthermore, learning is organization dependent, as businesses have different histories, competitive factors, and challenges. Therefore, although using BDA can be well suited to account for the decision biases and errors caused by humans in both categories, decisions for SC learning activities are generally more difficult (than optimization activities) to automate. In other words, the employment of highly intelligent BDA tools for SC optimization requires less human intervention in both information selection and processing (Agarwal & Dhar, 2014). As such, BDA use for SC optimization may be expected to have a greater likelihood of directly influencing decision-making processes than BDA use for SC learning. Therefore, we hypothesize the following:

H1: BDA usage for SC optimization has a stronger direct positive influence on BDA-enabled decision-making capability than does BDA usage for SC learning.

In addition, the indirect path via internal integration to influence decision-making capability seems less relevant to BDA use for SC optimization than to BDA use for SC learning. Internal integration establishes rules, systems, procedures, and cross-functional relations that enable decision makers to achieve a shared understanding of opportunities and threats to the entire supply chain (Srinivasan & Swink, 2015). As already explained, SC optimization activities involve structured decisions that emphasize control and continuity regarding existing processes. In particular, most of these activities are performed in compliance with established decision rules that may be already be apparent across the organization. As such, the need for BDA use for SC optimization is less likely to require postprocessing via internal integration for its potential impact on organizational decision-making capability.

In contrast, internal integration seems more relevant for BDA use for SC learning. Organizational learning focuses heavily on developing insights and strategic plans for new opportunities, which is far less likely to have programmable paths to follow. The fact that BDA use for SC learning is organizationally embedded suggests that the success of such BDA usage depends on the inclusion of BDA in a firm's long-term business strategy and on the mechanisms in place to facilitate business alignment with this strategy (Grover et al., 2018). Such alignment involves streamlining processes, policies, procedures, organizational structure/governance, and corporate culture to leverage data for competitiveness. For this reason, internal integration is an essential mechanism that helps reconcile potentially myopic and inconsistent self-interested views by providing a relational platform for a wider range of supply chain decision stakeholders (e.g., purchasing, planning, production, logistics, etc.) to participate in joint evaluations and planning for the use of knowledge (Schoenherr & Swink, 2012; Swink et al., 2007).

Specifically, BDA use for SC learning allows for the creation of "communities of knowing" (Sambamurthy et al., 2003) among managers of various supply chain stages, characterized by frequent knowledge production and sharing, thereby commanding effective internal organizational integration. Meanwhile, internal integration facilitates functional goal-alignment, emphasizes operational interdependencies, and encourages the cooperative use of each functional area's resources and capabilities through information sharing and knowledge exchange (Schoenherr & Swink, 2012; Williams et al., 2013). Furthermore, internal integration is recognized as an important source of information processing capability for supply chain organizations in that it helps absorb (recognize, evaluate, assimilate, and apply) the insights generated by technologies (Swink & Schoenherr, 2015). Consequently, we expect that the mediating effect of internal integration plays a more significant role in translating BDA use for SC learning into better decision-making. Accordingly, we present the following hypothesis:

H2: Internal integration has a stronger positive mediating effect on the relationship between BDA usage for SC learning and BDA-enabled decision-making capability than on the relationship between BDA usage for optimization and BDA-enabled decision-making capability.

3 Research Methodology

To test the research hypotheses, we conducted an empirical study to collect data from supply chain executives via a questionnaire. The researchers designed the survey instrument in collaboration with the Computer Sciences Corporation (CSC, a major business technology consulting firm with a large supply chain IT practice) and the *Supply Chain Management Review* (SCMR, a trade journal). The survey instrument targeted supply chain executives as respondents who could appropriately represent the SCM organization as the unit of analysis. Invitations to the online survey were sent via email to the clients of CSC and to members of the Council of Supply Chain Management Professionals (CSCMP) who had responsibilities directly related to supply chain management. Three rounds of invitations were sent to prospective respondents over a six-week period. Targeted industries included manufacturing, wholesale, and retail sectors (see Table 1). The target population and resulting sample mainly included supply chain managers from businesses based in North America.

3.1 Measures and Questionnaire Administration

The research team worked with CSC to ensure face validity in the design of the measures, and also to provide relevance and readability for the respondents. The questionnaire contained both measures from the extant literature and new/adapted measures for key constructs in the research model: BDA use for SC optimization, BDA use for SC learning, internal integration, and BDA-enabled decision-making capability. In addition, we assessed control variables including industry, annual sales, number of employees, number of IT professionals, the respondent's managerial level, and strategic focus of the organization (i.e., operations focus, customer focus, new products/services focus) in the organization. All survey items were suggested, reviewed, and validated by the following groups: three academic researchers, two CSC partners, and five SC practitioners. Insights from these collaborators ensured that the survey used familiar terms and applications of BDA.

Specifically, following existing literature on supply chain BDA practices (Chen et al., 2015), measurement items were included in the scale to assess the BDA uses for SC optimization and SC learning.[7] We also adapted existing measurement items to assess both internal integration (Schoenherr & Swink, 2012; Williams et al., 2013) and BDA-enabled decision-making capability (Roberts & Grover, 2012; Williams et al., 2013). Specifically, with regard to BDA-enabled decision-making capability, these adapted measurement items entailed unique BDA techniques (e.g., sensors, data visualization apps, etc.) that help with organizational decision-making. Appendix A provides the sources and details describing the survey items and scales for all the constructs. The definitions and examples of big data and big data analytics were provided at the beginning of the survey to create a common understanding of these constructs among respondents.[8]

3.2 The Sample

After screening the responses and eliminating informants whose titles did not directly relate to a supply chain function, we additionally removed four outliers (Hoaglin & Iglewicz, 1987; Hoaglin et al., 1986), yielding a total of 157 usable responses representing multiple industries (see Table 1). The respondents held a range of senior and upper-level supply chain managerial positions, suggesting they possessed relevant knowledge for the study. Specifically, the titles included executive managers (C-level executive, president, senior vice president, and vice president), upper managers (senior director, director, head), and middle managers (senior manager, manager). The sample firms' sizes range from less than \$250 million to more than \$10 billion in revenue, and from less than 250 employees to greater than 30,000 employees. The total number of IT professionals in the organizations ranged from 1 to 20,000. Thus, our sample spanned a wide range of organizations, providing some degree of generalizability. A limitation of the method used to invite respondents to participate in this online survey is that it did not fully allow for the calculation of a conventional response rate.

We assessed nonresponse bias by first comparing performances (return-on-assets) of the sampled firms against their respective industry median values. None of the differences were statistically significant ($p > 0.05$). Second, we compared measurement item scores across early (i.e., first 25%) and late (i.e., last 25%) respondents, using late responses as representatives of nonresponse. Again, none of the differences were statistically significant, leading us to conclude that nonresponse bias is not an appreciable concern.

Because our sample consisted of respondents with different levels of job titles characterized as executive, upper-level managers, and middle managers, to further assess respondent validity, we also compared the performance (return-on-assets) of the sampled firms with their respective industry median values segmented into groups based on the respondent's organizational position. The findings indicate that none of the differences were statistically significant. In addition, we compared measurement item scores for each of the variables based on the respondent's position. Again, none of the differences were found to be statistically significant across the groups corresponding to job positions. We also included the respondent's job position as a control variable in our research model (which was not found to have a significant effect on either dependent variable).

4 Data Analysis and Results

We employed a two-step, partial least squares (PLS) regression analysis of the survey data, in order to examine the nomological validity of the research model. In the first step, we evaluated the measurement model for validity and reliability. In the second step, we evaluated the structural model to assess support for the hypotheses. Through this process, we evaluated the properties of all scales within the context of the structural model through an assessment of discriminant validity and reliability.

[7] Chen et al. (2015) specified three dimensions of BDA usage for SCM (reconfiguration, learning, and coordination) and followed a rigorous approach to develop valid measurement items. However, they treated BDA usage as a monolithic construct. In the current study, we adapted the items for SC reconfiguration and SC learning by Chen et al. (2015) to measure BDA use for SC optimization and SC learning, respectively.

[8] Please note, the time frame we asked the respondents to refer to when answering the questions in the survey was "the past three years."

Table 1. Summary Statistics of Variables

Variable (# items)	N	Mean	Std. Dev.	Min.	Max.
BDA-optimization (3)	157	2.54	1.14	1.00	5.00
BDA-learning (3)	157	2.46	1.12	1.00	5.00
Internal-integration (4)	157	3.46	0.80	1.50	5.00
BDA-enabled decision-making capability (4)	157	2.80	0.94	1.00	5.00
Number of IT professionals	157	505.67	2,052.16	1	20,000
Firm annual sales	157	< \$250 million = 36.3% \$251 million - \$500 million = 7.6% \$500 million - \$1 billion = 7.0% \$1 billion - \$10 billion = 34.4% > \$ 10 billion = 14.7%			
Number of employees	157	< 250 = 27.4% 250 - 1,000 = 14.0% 1,000 - 10,000 = 26.1% 10,000 - 30,000 = 14.7% > 30,000 = 17.8%			
Industry type	157	Automotive & Transportation = 3.8% Chemicals = 3.8% Consumer Goods = 10.8% Electrical Equipment = 1.9% Electronics = 1.3% Machinery & Industrial Equipment = 2.5% Mining & Metals = 1.3% Oil & Gas = 1.3% Pharmaceuticals = 1.9% Pulp & Paper = 0% Publishing & Printing = 2.5% Other Manufacturing Industry = 7.0% Financial = 0.006% Food = 8.3% Government = 0.6% Healthcare = 10.8% Retail = 9.6% Leisure & Recreation = 0% Professional = 7.6% Third Party Logistics Provider = 11.5% Wholesale/Distribution = 7.6% Other Service Industry = 5.7%			

4.1 Measurement Model

Table 1 presents the summary statistics of all the variables. Tables 2 and 3 provide the interconstruct properties of the scales and the results of the factor analysis, respectively. Consistent with the extant literature, all variables in the research model were modeled as reflective constructs. [9] Item loadings, reliabilities, and discriminant validity were observed to generally meet accepted guidelines. Table 2 provides interconstruct correlations, and the reliability (composite reliability and Cronbach's alpha) and square root of the AVE values, respectively. The composite reliabilities of most of the constructs are greater than or very close to 0.90, with the lowest value at 0.88. All Cronbach's alpha values are higher than the recommended value of 0.70 (Nunnally 1978). To assess convergent and discriminant validity (Chin 1998), we examined cross-loadings and average variance extracted. As the factor analysis results (Table 3) show, all items have a loading of greater than 0.70 (i.e., the "cutoff value") and also items were found to load more highly on their own construct than on other constructs. Furthermore, as shown by comparing the interconstruct correlations and AVE (shaded leading diagonal in Table 2), the square root of the AVE for each construct is

[9] Within the prior literature, "technology usage" has been modeled both formatively and reflectively. We follow Chen et al. (2015) to examine both BDA use for optimization and learning reflectively. Justification for such reflective modeling is validated via additional testing, which will be reported in the next section.

significantly higher than the correlation of that construct with all other constructs. That is, all constructs share more variance with their indicators than with other constructs (Fornell & Larcker, 1981; Wetzels et al., 2009). Thus, these results indicate satisfactory convergent and discriminant validity of the constructs. We do wish to note that there is a level of cross-loading between BDA use for SC optimization and BDA use for SC learning, which was expected to some degree, as these are two facets of BDA usage. Nevertheless, the findings indicate an acceptable discriminant validity. In addition, we evaluated the potential for multicollinearity of the independent variables (i.e., BDA use for optimization and BDA use for learning). The results indicate that the variance inflation factor (VIF) values range from 1.1 to 2.2, well below a threshold VIF of 5.0 for multicollinearity. As such, we conclude that there are no appreciable issues with multicollinearity based on these VIF values and also because correlations among variables are below 0.80 (Hair et al., 1998).

In addition, because the data come from a single respondent for each organization, it was important to assess potential common method variance (CMV) in the measurement model. Several factors suggest that CMV is not a serious source of bias in this study. The survey development process via industry experts provides validation of the measures prior to the distribution of the survey and the selection of knowledgeable supply chain executives as the key respondents mitigates the potential for CMV (Craighead et al., 2011). A series of steps were taken to prevent potential common method bias in the design of the instrument in accordance with established guidelines. The design of the survey includes features also aimed at mitigating CMV, including the separation of independent and dependent variables over the extent of the survey and, where possible, the use of dissimilar scale formats and anchors (Podsakoff et al., 2003). Furthermore, we conducted several statistical tests to assess the potential of CMV, including Harman's single-factor test, the correlational marker variable technique, and the examination of variance inflation factors. Collectively, the results from these tests suggest that CMV is not a major concern for this study.[10]

4.2 Structural Model

We employed PLS to test the structural model. The significance of the paths was determined using the *t*-statistic, calculated with the bootstrapping technique. As noted previously, we included control variables (annual sales, number of employees, number of IT professionals, industry, strategic focus, and the organizational position of the respondent) for both the mediating/dependent variable (i.e., internal integration) and dependent variable (i.e., BDA-enabled decision-making capability) in our research model. The results show that none of the control variables exerted a statistically significant influence on either of the dependent variables. Figure 2 shows the path coefficients for the structural model (including those of the control variables). We observe: (1) the direct path from BDA use for SC optimization to BDA-enabled decision-making capability is significant while the direct path from BDA use for SC learning is not significant, (2) the indirect path from BDA use for SC optimization to BDA-enabled decision-making capability via internal integration is not significant while the indirect path from BDA use for SC learning is significant.

To further validate the two reflectively modeled BDA use constructs, we reran the structural model with both BDA use for SC optimization and for SC learning modeled as formative items (rather than reflective) and found that the structural model remains essentially unchanged (i.e., all significant paths remain significant at the same level of statistical significance and all of the nonsignificant paths remain consistent). As such, interpretational confounding is not a concern in this case (Kim et al., 2010).

Based on the results of the structural model, we evaluate the support of the hypotheses in our research model. First, we observe that BDA use for SC learning has a significant influence on internal integration while BDA use for SC optimization does not. We note that the two BDA use patterns collectively explain 13.0% of the variance in internal integration (please note that the inclusion of the controls explains an additional 1.7% yielding a total of 14.7% of the variance in internal integration). Second, we find that BDA use for SC optimization has a direct significant influence on BDA-enabled decision-making capability while BDA use for SC learning does not. Furthermore, the findings indicate that internal integration has a significant influence on BDA-enabled decision-making capability. Internal integration in conjunction with two types of BDA usage collectively explains 33.6% of the variance in BDA-enabled decision-making capability (note: the inclusion of the controls explains an additional 2.3% yielding a total of 35.9% of the variance in decision-making capability).

We conducted mediation analyses to further assess the nomological network of the research model. As noted above, based on the results, internal integration does not have a mediating effect on the relationship between BDA use for SC optimization and BDA-enabled decision-making capability. As such, we focused on the examination of the potential mediating effect of internal integration on the relationship between BDA use for SC learning and BDA-enabled decision-making capability.

[10] The details and the results of the statistical analysis to assess CMV are provided in Appendix B.

Table 2. Reliability and Interconstruct Correlations

Construct (# items)	Reliability measures		Interconstruct correlations			
	Reliability	Cronbach's alpha	BDA-optimization	BDA-learning	Internal-integration	BDA-enabled DMC
BDA-optimization (3)	0.884	0.803	0.848			
BDA-learning (3)	0.912	0.857	0.737	0.881		
Internal-integration (4)	0.911	0.871	0.328	0.341	0.849	
BDA-enabled DMC (4)	0.896	0.846	0.540	0.477	0.353	0.827

Note: Reliability measures: Guidelines indicate that values for composite reliability and Cronbach's alpha should exceed or approach 0.900 and 0.700, respectively. Interconstruct correlations: The shaded numbers on the leading diagonal are the square root of the AVE. Guidelines indicate that the square root of AVE values should exceed 0.700 and also exceed the correlation of that construct with all other constructs (values on the vertical column).

Table 3. Results of Factor Analysis

Indicators	BDA-optimization	BDA-learning	Internal-integration	BDA-enabled DMC
BDA-Use-Optimiz1	0.847	0.625	0.212	0.503
BDA-Use-Optimiz2	0.816	0.655	0.271	0.436
BDA-Use-Optimiz3	0.878	0.597	0.351	0.435
BDA-Use-Learn1	0.722	0.906	0.343	0.487
BDA-Use-Learn2	0.617	0.917	0.295	0.403
BDA-Use-Learn3	0.597	0.818	0.251	0.356
Internal-Integ1	0.204	0.204	0.789	0.244
Internal-Integ2	0.276	0.309	0.886	0.273
Internal-Integ3	0.306	0.338	0.882	0.343
Internal-Integ4	0.310	0.287	0.834	0.324
BDA-EnabledDMC1	0.433	0.360	0.257	0.768
BDA-EnabledDMC2	0.383	0.346	0.242	0.829
BDA-EnabledDMC3	0.476	0.464	0.325	0.849
BDA-EnabledDMC4	0.484	0.396	0.333	0.858

Note: Guidelines indicate that for a factor analysis, all items should have a loading on their own construct factor of greater than 0.700 and also load more highly on their own construct than on other constructs (designated by each row in the table).

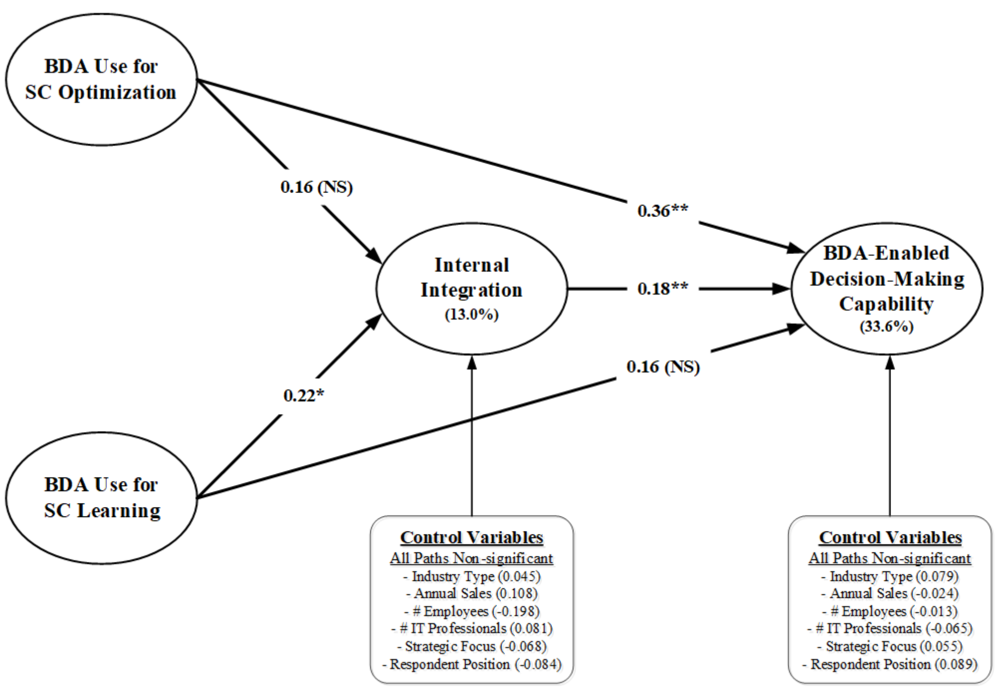


Figure 3. Data Analysis Results of the Structural Model

To conduct this additional assessment, we followed the mediation analysis procedures suggested by prior literature (Baron & Kenny, 1986; Sobel, 1982). Following the approach outlined by Baron and Kenny (1986), we removed the internal integration variable from the model (while retaining BDA use for SC optimization as an antecedent in the model) and observed that, in this case, BDA use for SC learning directly influences BDA-enabled decision-making capability ($t = 2.01$, $p < 0.05$). The Sobel (1982) tests provide additional evidence that internal integration mediates the effect of BDA use for SC learning ($t = 2.44$, $p < 0.01$). Therefore, the findings of this mediation analysis support that internal integration fully mediates the influence of BDA use for SC learning on BDA-enabled decision-making capability.

We further assessed the mediation of BDA use for SC learning by internal integration on BDA-enabled decision-making capability via the bias bootstrap method (Preacher & Hayes, 2008). This method examines indirect effects via a Monte Carlo simulation with significance levels determined by the bias-corrected bootstrap method (i.e., by employing a 99% confidence interval and 20,000 repetitions) (Preacher & Selig, 2012; Selig & Preacher, 2008; Swink & Schoenherr, 2015). This approach has several advantages: it provides a high level of statistical power, indirect effects can be measured directly rather than inferred, and it does not need to assume that the mediating effect is normally distributed (MacKinnon, 2008; Preacher, 2015; Vance et al., 2015). The results of the bootstrapping simulation found that the 99% confidence interval of the distribution of indirect effects did not include zero (1% lower bound = 0.004; 1% upper bound = 0.119). Thus, the finding provides additional support for mediation in this case (Preacher & Selig, 2012; Swink & Schoenherr, 2015). Collectively, the results of a series mediation analyses provide robust evidence that internal integration fully mediates the influence of BDA learning on BDA-enabled decision-making capability. As such, the findings provide empirical support for both hypotheses of the study.

5 Discussion

This study is among the first to develop and test an integrated theoretical model that explains how different patterns of organizational level BDA usage can influence internal integration and decision-making capability within the supply chain management domain. BDA offers clear functional advantages in exploiting the emerging availability of large and unstructured data sets, and the popular literature has been quick to present BDA applications as new opportunities for managers to create all types of performance gains. Getting beyond the hype, however, requires a more rigorous study of factors that explain how technologies such as BDA can be leveraged in specific contexts. We offer a more nuanced view to examine the "black box" through which BDA usage leads to greater decision-making capability within supply chain management, with expectations grounded in theory.

5.1 Theoretical Contributions

Our findings support the proposition that BDA usage is instrumental in improving an organization's decision-making capability. In other words, it is only through effective use that organizations may reap benefits potentially enabled by BDA technologies. Nevertheless, the emerging literature on the business value of BDA generally treats BDA use as a whole (Chen et al., 2015; Müller et al., 2018), even though decision processes may potentially vary across different organizational activities. We enrich the literature by considering the distinction between two patterns of BDA usage (for optimization and for learning) to be an important contribution to this line of literature. By highlighting different patterns of BDA use for structured and unstructured decisions, we show that BDA usage can enhance organizational decision-making capability in multiple ways. Whereas the impacts of BDA usage on decision-making in structured SC optimization activities are quite straightforward and immediate, the impacts on decision-making in non- or less structured business activities, i.e., SC learning, may not be direct and easily recognized. These findings provide helpful explanation regarding the mixed findings of BDA value reported in early studies, particularly why some organizations were not able to timely capitalize on the benefits they had been expecting from their BDA investments (Grover et al., 2018; Tallon, 2013). Hence, depending upon the nature of the decisions (i.e., structured/nonstructured) for which BDA is employed, the additional effort of engaging functional representatives to collectively digest insights generated by BDA analysis could be essential.

In particular, SCM is largely a boundary spanning and integrative function that must address many cross-functional interdependencies (Ashenbaum & Terpend, 2010; Parker & Anderson, 2002; Zhang et al., 2011). Internal integration is an important mechanism through which organizational implementation of BDA generates improved decision-making capability. Consistent with the tenets of information processing theory in that BDA (as a novel approach to collect and distribute information) improves organizational decision-making capability by addressing its information processing needs, we offer empirical evidence to corroborate that internal integration (i.e., an organizational expression of lateral relations) is a key factor mediating the influence of BDA use for SC learning on BDA-enabled decision-making capability. We view this result as evidence of the benefit of creating organizational mechanisms that serve to assimilate and integrate information developed by various functionally applied information technologies. Internal integration can also be viewed as an aspect of group cohesion, or the ability for decision makers across functions to work together in a complementary fashion. These findings extend the emerging literature on the interactions of vertical and lateral information processing systems (Galbraith, 1973; Kim & Narasimhan, 2002; Srinivasan & Swink, 2015). Future research might explore how team dynamics and lateral organizational relationships enhance knowledge creation and BDA-enabled decision-making phenomena.

5.2 Managerial Implications

The results of this study have important implications for managerial practice. We noted earlier that there has been intensive promotion within the practitioner press with regard to the potential impacts of BDA. Industry discussion has often emphasized that BDA has the potential to influence organizational outcomes. However, despite the purported promise of BDA, the means through which BDA can have such impact have generally remained opaque. Our findings offer an understanding of a mechanism and conditions through which BDA can influence decision-making capability within the supply chain context. For programmable routine activities (e.g., inventory monitoring, production scheduling, etc.), appropriate BDA usage may offer instant decision-making benefits. However, for nonroutine strategic initiatives, managers should recognize that organizational factors are essential precursors to successful exploitation of BDA. Specifically, information technologies that serve to analyze, combine, and disseminate insights from BDA appear to be essential for enhancing decision-making capability. Meanwhile, managers should design processes, teams, and other integrative mechanisms to process BDA inputs from various functional areas, thus creating more holistic and globally optimal decisions.

5.3 Limitations and Future Research

We note some potential limitations of the current study. First, because of the difficulty of collecting subjective data from supply chain executives, the sample size (n = 157) of the study, although acceptable for executive-level studies, is relatively small. Also, we wish to note that the measurement items were developed after review of the extant literature and conversations with industry practitioners and, as such, appropriately represent a current and relevant phenomenon of interest. However, we recognize that both BDA and the SC landscape will evolve over time and encourage future research to attempt to match the relevant phenomenon. As such, we call for future larger-scale studies (perhaps using multiple sources of data) to replicate and extend our research model. In particular, future research will need to ensure that it is conducted in accordance with industry conditions.

Second, the theories we employ use causal terms to describe the relationships included in the research model. However, our cross-sectional research design does not provide firm conclusions of causality. While our results are certainly consistent with the hypothesized relationships, longitudinal research would provide additional support for causal effects. Such a longitudinal study might provide additional insight into the temporal nature of how different patterns of BDA use (i.e., for optimization and for learning) lead to greater supply chain decision-making capability. We consider such research both important and timely. Our data indicate a mean near the center of the range of the scale (1 to 5) and also an appreciable degree of variance in levels for both BDA use for SC optimization (mean = 2.54; std. dev. = 1.14) and for SC learning (mean = 2.46; std. dev. = 1.12) in our sample, all of which suggest a broad range of coverage for the firms but with a greater proportion of firms at a potentially earlier stage of implementation of BDA usage. Longitudinal studies would allow greater insight into this phenomenon as BDA becomes more mature. In particular, we recognize that, in some firms, technologies supporting BDA use for SC optimization and SC learning may be at different stages of development, respectively. As learning-oriented technologies become more mature, for example, their outputs may be more seamlessly integrated into operational plans and responses, thus requiring less interpretation and application by cross-functional teams. Given the considerable evolution of BDA technologies, it would be interesting for researchers to examine whether or not our findings continue to hold in the future.

Third, while the contextual focus of this study is a strength, it is also a limitation to some degree. We focus on BDA practices within the organizational context of supply chain management for firms in a multitude of industries. We leave it to future research to explore similar effects in other functional areas and in other industrial settings. Such future studies could potentially expand the generalizability of the current findings and also enrich its basis in other contexts.

Despite the limitations outlined, this study extends the existing body of knowledge with regard to how organizational intelligence is key to yielding greater decision-making capability within the organization’s supply chain function. Specifically, the research model, which is cohesively grounded in theoretical lenses that include theories of OIPT (Galbraith, 1973) and decision-making (Gorry & Scott Morton, 1971), is readily applicable and generalizable to future studies on the value creation process of knowledge creation and decision-making systems. The current study develops a bridge between BDA usage and BDA-enabled decision-making capability and differential effects via internal integration; however, additional development of the nomological network is warranted in future research to provide a richer understanding of this phenomenon. Future studies might investigate the roles of specific SCM systems in feeding BDA processes. Such systems include relationship management tools, transaction management and execution systems, and focused planning tools. In addition, researchers might consider additional contextual settings that may govern the roles and values of different types of BDA. Similarly, the integration of OIPT with decision-making theory provides theoretical positioning supporting the current nomological network; however, future research should look to integrate complementary theoretical perspectives, where applicable.

6 Conclusion

Our study contributes to a nomological network that examines the influence of BDA usage (i.e., in terms of BDA use for SC optimization and BDA use for SC learning) on BDA-enabled decision-making capability as facilitated through internal integration. Our findings provide evidence that BDA use for SC optimization has a direct influence on BDA-enabled decision-making capability, and internal integration does not play a facilitating role in this process. Furthermore, the findings support that BDA use for SC learning does not have a direct influence on BDA-enabled decision-making capability, and BDA use for SC learning does indirectly affect BDA-enabled decision-making capability fully through internal integration. Collectively, the findings provide a theory-based understanding of how organizational BDA usage influences decision-making within supply chain management, which offers key implications for organizational managers.

References

Abbasi, A., Sarker, S., & Chiang, R. H. L. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2), i-xxxii.

Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research*, *25*(3) 443-448.

Anthony, R. N. (1965). *Planning and control systems: A framework for analysis*. Harvard University Press.

Ashenbaum, B., & Terpend, R. (2010). The purchasing-logistics interface: A scope of responsibility taxonomy. *Journal of Business Logistics*, *31*, 177-194.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical concerns. *Journal of Personality and Social Psychology*, *51*(6), 1173-1182.

Boone, C. A., Hazen, B. T., Skipper, J. B., & Overstreet, R. E. (2018). A framework for investigating optimization of service parts performance with big data. *Annals of Operations Research*, *270*(1-2), 65-74.

Bourgeois, L. J., & Eisenhardt, K. M. (1988). Strategic decision processes in high velocity environment: four cases in the microcomputer industry. *Management Science*, *34*(7), 816-835.

Chakravarty, A., Grewal, R., & Sambamurthy, V. (2013). Information technology competencies, organizational agility, and firm performance: enabling and facilitating roles. *Information Systems Research*, *24*(4), 976-997.

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, *32*(4), 4-39.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, *36*(4), 1165-1188.

Chin, W. W. (1998). Issues and opinion on structural equation modeling. *MIS Quarterly*, *22*(1), vii-xvi.

Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, *27*(10), 1868-1883.

Clark, T. D., Jones, M. C., & Armstrong, C. P. (2007). The dynamics structure of management support systems: Theory development, research focus, and direction. *MIS Quarterly*, *31*(3), 579-615.

Cox, A. (2015). Sourcing portfolio analysis and power positioning: towards a “paradigm shift” in category management and strategic sourcing. *Supply Chain Management*, *20*(6), 717-736.

Craighead, C. W., Ketchen, D. J., Dunn, K. S., & Hult, G. T. M. (2011). Addressing common method variance: Guidelines for survey research on information technology, operations, and supply chain management. *IEEE Transactions on Engineering Management*, *58*(3), 578-588.

Dantzig, G. B. (1963). *Linear programming and extensions* (1st ed.). Princeton University Press.

Davenport, T. (2006). Competing on analytics. *Harvard Business Review*, *84*(1), 98-107.

Davenport, T. (2013). Analytics 3.0. *Harvard Business Review*, *91*(12), 65-72.

Droge, C., Jayaram, J., & Vickery, S. K. (2004). The effects of internal versus external integration practices on time-based performance and overall firm performance. *Journal of Operations Management*, *22*, 557-573.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, *14*, 57-74.

Eisenhardt, K. M., & Zbaracki, M. J. (1992). Strategic decision making. *Strategic Management Journal*, *13*, 17-37.

Fornell, M., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50.

Galbraith, J. R. (1973). *Designing complex organizations*. Addison-Wesley.

Galbraith, J. R. (2014). Organization design challenges resulting from big data. *Journal of Organization Design*, *30*(1), 2-13.

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, *19*(2), 213-236.

Goodhue, D. L., Wybo, M. D., & Kirsch, L. J. (1992). The impact of data integration on the costs and benefits of information systems. *MIS Quarterly*, *16*(3), 293-311.

Gorry, G. A., & Scott Morton, M. S. (1971). A framework for management information systems. *Sloan Management Review*, *13*(1), 55-70.

Grover, V., Chiang, R. H. L., Liang, T.-P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, *35*(2), 388-423.

Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis with readings* (5th ed.). Prentice Hall.

Hoaglin, D. C., & Iglewicz, B. (1987). Fine Tuning some resistant rules for outlier labeling. *Journal of American Statistical Association*, *82*, 1147-1149.

Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of some resistant rules for outlier labeling. *Journal of American Statistical Association*, *81*, 991-999.

Joshi, K. D., Chi, L., Datta, A., & Nan, S. (2010). Changing the competitive landscape: Continuous innovation through it-enabled knowledge capabilities. *Information Systems Research*, *21*(3), 472-495.

Kim, S. W., & Narasimhan, R. (2002). Information system utilization in supply chain integration efforts. *International Journal of Production Research*, *40*(18), 4585-4609.

Kim, G., Shin, B., & Grover, V. (2010). Investigating two contradictory views of formative measurement in information systems research, *MIS Quarterly*, *34*(2), 345-365.

Kitchens, B., Dobolyi, D., Li, J., & Abbasi, A. (2018). Advanced customer analytics: strategic value through integration of relationship-oriented big data. *Journal of Management Information Systems*, 35(2), 540-574.

Kumar, C. V. S., & Routroy, S. (2016). Case application of a methodology for determining a manufacturer's preferred customer status with suppliers. *Engineering Management Journal*, *28*(1), 25-38.

Leidner, D. E., & Elam, J. L. (1995). The Impact of Excutive Information systems on organizational design, intelligence, and decision making. *Organization Science*, *6*(6), 645-664.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal*, *14*, 95-112.

March, J. G. (1962). The business firm as a political coalition. *The Journal of Politics*, 24, p 662-678.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71-87.

Müller, O., Fay, M., & Brocke, J. V. (2018). The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics. *Journal of Management Information Systems*, *35*(2), 488-509.

Nunnally, J. C. 1978. *Psychometric theory* (2nd ed.). McGraw-Hill.

Park, Y., El Sawy, O. A., & Fiss, P. C. (2017). The role of business intelligence and communication technologies in organizational agility: A configurational approach. *Journal of Association for Information Systems*, *18*(9), 648-686.

Parker, G. G., & Anderson, E. G. (2002). From buyer to integrator: The transformation of the supply-chain manager in the vertically disintegrated firm. *Production and Operations Management*, *11*, 75-91.

Peng, D. X., Heim, G. R., & Mallick, D. N. (2014). Collaborative product development: The effect of project complexity on the use of information technology tools and new product development practices. *Production and Operations Management*, *23*(8), 1421-1438.

Podsakoff, P. M., Scott, B. M., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879-891.

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures*, *6*(1), 77-98.

Rai, A., Patnayakuni, R., & Patnayakuni, N. (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS Quarterly*, *29*(4), 625-652.

Roberts, N., & Grover, V. (2012). Leveraging information technology infrastructure to facilitate a firm's customer agility and competitive activity: An empirical investigation. *Journal of Management Information Systems*, *28*(4), 231-269.

Sabherwal, R., & Sabherwal, S. (2005). Knowledge management using information technology: Determinants of short-term impact on firm value. *Decision Sciences*, *36*(4), 531-567.

Sanders, N. R., & Ganeshan, R. (2015). Call for papers: Special issue of production and operations management on big data in supply chain management. *Production & Operations Management*, *24*(4), 681-682.

Schoenherr, T., & Swink, M. (2012). Revisiting the arcs of integration: Cross-validations and extensions. *Journal of Operations Management*, *30*(1), 99-115.

Selig, J. P., & Preacher, K. J. (2008). *Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects*.www.quantpsy.org/medmc/ medmc111.htm

Simon, H. A. (1960). *The new science of management decisions*. Harper.

Sobel, M. E. (1982). *Asymptotic confidence intervals for indirect effects in structural equation models*. In Leinhardt S. (Ed.), Social Methodology, 290-312. Jossey-Bass.

Srinivasan, R., & Swink, M. (2015). Leveraging supply chain integration through planning comprehensiveness: An organizational information processing theory perspective. *Decision Sciences*, *46*(5), 823-861.

Srinivasan, R., & Swink, M. (2018). An investigation of visibility and flexibility as complements to supply chain analytics: An organizational information processing theory perspective. *Production and Operations Management*, *27*(10), 1849-1867.

Swink, M., Narasimhan, R., & Wang, C. (2007). Managing beyond the factory walls: Effects of four types of strategic integration on manufacturing plant performance. *Journal of Operations Management*, *25*(1), 148-164.

Swink, M., & Schoenherr, T. (2015). The effects of cross-functional integration on profitability, process efficiency, and asset productivity. *Journal of Business Logistics*, *36*(1), 69-87.

Tallon, P. (2013). Corporate governance of big data: perspectives on value, risk, and cost. *IEEE Computer*, *46*(6), 32-38.

Thomas, J. B., Clark, S. M., & Gioia, D. A. (1993). Strategic sensemaking and organizational performance: linkages among scanning, interpretation, action, and outcomes, *Academy of Management Journal*, *36*(2), 239-270.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84.

Wetzels, M., Odekerken-Schroder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly*, *33*(1), 177-195.

Williams, B., Roh, J., Tokar, T., & Swink, M. (2013). Leveraging supply chain visibility for operational responsiveness: The moderating role of internal integration. *Journal of Operations Management*, *31*(7), 543-554.

Wong, E., & Wei, Y. (2018). Customer online shopping experience data analytics: Integrated customer segmentation and customised services prediction model. *International Journal of Retail & Distribution Management*, *46*(4), 406-420.

Zhang, C., Viswanathan, S., & Henke, J. W. (2011). The boundary spanning capabilities of purchasing agents in buyer-supplier trust development. *Journal of Operations Management*, *29*, 318-328.

Zhao, X., Huo, B., Selen, W., & Yeung, J. H. Y. (2011). The impact of internal integration and relationship commitment on external integration. *Journal of Operations Management*, *29*(1), 17-32.

Zuboff, S. (1985). Automate/Informate: The two faces of intelligent technology. *Organizational Dynamics*, *14*(2), 5-18.

Appendix A: Construct Scales and Items

Definitions provided at the beginning of the survey:

"Big Data" refers to large structured and unstructured data sets that require new forms of processing capability to enable better decision making. Examples include sales data, process operating data and other information captured by sensors, web server logs, Internet clickstream data, social media activity reports, mobile-phone call records, etc**.**

"Big Data Analytics" is the process of examining Big Data using advanced technologies. These include data management (e.g., massively parallel-processing databases), open-source programming (e.g., Hadoop, MapReduce), statistical analysis (e.g., sentiment analysis, time-series analysis), visualization tools that help structure and connect data to uncover hidden patterns, anomalies, unknown correlations, and other actionable insights, and in-memory computing (e.g., SAP's HANA).

Big Data Analytics Use for SC Optimization [source: (Chen et al., 2015; Teece, 2007), Expert Interviews] To what extent has your organization implemented Big Data Analytics in each area? (1=little or no usage…3=moderate usage…5=heavy usage) 1) Network design/optimization 2) Production run optimization 3) Inventory optimization
Big Data Analytics Use for SC Learning [source: (Chen et al., 2015; Teece, 2007), Expert Interviews] To what extent has your organization implemented Big Data Analytics in each area? (1=little or no usage…3=moderate usage…5=heavy usage) 1) Sourcing analysis 2) Purchasing spend analytics 3) CRM/customer/patient analysis
Internal Integration [source: (Schoenherr & Swink, 2012; Williams et al., 2013)] Please indicate your level of agreement with the following statements: (1=strongly disagree…3=neutral…5=strongly agree) 1) Functional teams have a common prioritization of customers in case of supply shortages and how allocations will be made 2) Operational and tactical information is regularly exchanged between functional teams 3) Planning decisions are based on plans agreed upon by all functional teams 4) All functional teams use common plans and procedures
BDA-Enabled Decision-Making Capability [source: : (Galbraith 2014; Roberts & Grover, 2012; Williams et al., 2013)] Please indicate your level of agreement with the following statements: (1=strongly disagree…3=neutral…5=strongly agree) 1) We easily combine and integrate information from many data sources for use in our decision making 2) We often deploy dashboard applications / information to our managers' communication devices (e.g., smart phones, computers) 3) Our systems automatically make operational changes, based on performance criteria/business rules, in response to signals from sensors 4) Our systems give us the ability to decompose information to help root cause analysis and continuous improvement

Appendix B: Statistical Tests and Results for Common Method Bias

We conducted three statistical tests to assess the potential of CMV, including Harman’s single-factor test, correlational marker variable technique, and the examination of variance inflation factors. In accordance with established guidelines pertaining to Harman’s single-factor test, CMV does not appear to be problematic since: (1) several factors were identified, (2) the first factor did not account for the majority of the variance, and (3) there is no general factor in the unrotated factor structure (Podsakoff & Organ 1986; Podsakoff and et al., 2003).

The use of a marker variable technique is an essential statistical assessment of CMV, specifically within the fields of information systems and supply chain management research (Craighead et al., 2011). Through the application of this method, we introduced a measured marker variable (i.e., multi-item scale that assessed the general level of environmental volatility for the organization) that does not have a theoretical relationship to the other constructs in our research model. We assessed the correlation of the marker variable to that of all other latent variables in the model. The correlations between the marker variable and the other four latent variables range from 0.006 to 0.145, which are well below the threshold value of 0.300. Also, the maximum shared variance with the marker variable and any of the latent variables is 2.1% (with that of decision-making capability). In addition, we compared the proposed research model against a revised model with paths between the marker variable and each of the dependent variables in the research model. Correlations and path coefficients were not substantially different between the original and revised models, and the paths from the marker variable to internal integration and BDA-enabled decision-making capability (i.e., the endogenous variables within this path model) were observed to be nonsignificant. Collectively, the findings from applying a marker variable provide evidence that CMV is not an appreciable issue (Lindell & Whitney 2001; Malhotra et al., 2006).

We further tested for CMV issues via the evaluation of the VIF. As noted earlier in the text, we had examined VIF to determine that multicollinearity was not an issue between the two BDA use antecedents in the model. The assessment of VIF can also be employed to test for CMV by running structural models with each rotating latent variable as the sole dependent variable (e.g., four separate models in this case) and examining the resultant factor level VIF. The VIF results ranged from 1.1 to 2.4 (which are well below the threshold value of 3.3), thus providing support that there were not issues of CMV in our research model (Kock, 2015). In summary, the results from these series of statistical tests provide support that mitigates potential concerns associated with CMV.

About the Authors

Daniel Chen's research addresses the topics of digital strategy and IS leadership, business analytics and implications, technology and innovation management, and IT and supply chain management. His research has appeared or is forthcoming in several leading IT and operations management journals such as *Decision Sciences Journal, Decision Support Systems, IEEE Transactions on Engineering Management, Information Systems Research, Journal of Management Information Systems, Journal of Operations Management, Journal of Strategic Information Systems, MIS Quarterly,* and *MIS Quarterly Executive,* among others. Dr. Chen has served on the editorial boards of several premier IS and operations management journals. Currently, he is a senior editor for *Journal of the Association for Information Systems*, and an associate editor for *Journal of Operations Management* and *Decision Sciences Journal.*

David Preston is a professor of information systems at Texas Christian University, Fort Worth, Texas. He received his PhD in business administration and MBA from the University of Georgia, and a BS and MS in engineering from the University of Florida. His research interests include the role/impact of executive IS/business management with regard to organizational IS, IS strategic alignment, and the intersection of analytics and supply chain management. His work has been published or is forthcoming in a variety of journals, including *MIS Quarterly*, *Information Systems Research*, *Journal of Management Information Systems*, *Journal of Operations Management*, *Decision Sciences*, *IEEE Transactions on Engineering Management*, *Journal of Strategic Information Systems*, and *MIS Quarterly Executive*. He served two terms as an associate editor for *MIS Quarterly* and currently serves as a senior editor for the *Journal of the Association of Information Systems*. He also serves on the editorial board of *IEEE TEM* and the *Journal of Strategic Information Systems.*

Morgan Swink teaches and leads research in areas of supply chain management, innovation management, project management, and operations strategy. His current research projects address digital transformation, innovation management, servitization, visibility, collaborative integration, and financial impacts of supply management policies. Dr. Swink heads the Center for Supply Chain Innovation, a collaborative venture that engages business partners, faculty, and students. He is the former co-editor-in-chief for the *Journal of Operations Management,* a top academic journal, and he continues to serve in consulting and associate editor roles at several top journals. Dr. Swink has co-authored two supply chain operations textbooks, one managerial book on supply chain excellence, and more than 75 articles in a variety of academic and managerial journals.

Copyright © 2021 by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints, or via email from publications@aisnet.org.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.