Received 19 December 2011
Accepted 16 January 2012

A reviewer's guide to the grounded theory methodology in logistics and supply chain management research

Ila Manuj and Terrance L. Pohlen
Department of Marketing and Logistics, University of North Texas, Denton, Texas, USA

Abstract

Purpose – The purpose of this paper is to analyze previous grounded theory articles and, based on this analysis, to provide a framework to assist reviewers in evaluating grounded theory research and increasing the rigor and credibility of this methodology in logistics and supply chain journals.

Design/methodology/approach – An analysis of existing articles appearing in the leading logistics and supply chain journals combined with an extensive review of the grounded theory method literature were used to develop a comprehensive framework for evaluating grounded theory research.

Findings – The paper finds that no standard criteria for publication of grounded theory research exists in logistics and supply chain journals. Grounded theory is routinely confused with other qualitative methodologies. Overall, this situation leads to publications that do not adequately address or report on the process for developing a grounded theory.

Research limitations/implications – Reviewers can use this paper to establish the quality of grounded theory research. Reviewers who are unfamiliar with or skeptical of the grounded theory method can use the framework to evaluate the rigor and credibility of a grounded theory study rather than rejecting such research. The checklist can be used to provide thorough and constructive reviews to authors.

Originality/value – The paper presents a framework that provides a ready reference for reviewers to assess whether the authors have taken appropriate action in selecting a grounded theory methodology, collecting and analyzing data, developing a theory grounded in the data, and for evaluating their research. Existing research is compared with the framework to identify potential shortcomings in the review process. The application of the framework to the review of future articles provides an opportunity to increase the credibility and rigor of grounded theory research in logistics and supply chain management journals.

Keywords Grounded theory research, Qualitative research, Guide for reviewers, Distribution management, Supply chain management, Research, Periodicals, Literature

Paper type General review

1. Introduction

Grounded theory (GT) has become increasingly accepted by logistics and supply chain management researchers as a valid research methodology (Davis-Sramek and Fugate, 2007). Although logistics and supply chain management have gained recognition as academic disciplines, the supporting theory has largely been borrowed from

International Journal of Physical Distribution & Logistics Management
Vol. 42 No. 8/9, 2012
pp. 784-803

0960-0035
DOI 10.1108/09600031211269758

Authors are listed in the alphabetical order of their last names. Both authors contributed equally to the paper.

other disciplines. Some of these theories, because of their lack of grounding in logistics and supply chain management, may not fit or sufficiently explain logistics and supply chain phenomena. The attraction of GT stems from the need for theory to develop creative perspectives, generate powerful insights into human interaction and business practices, and explain new and even well-researched complex social phenomena (Mello and Flint, 2009; Glaser and Strauss, 1967).

GT presents several challenges to reviewers tasked with assessing the rigor and credibility of research. Reviewers have to rely on the authors' accounts of the research process and explanations of how key evaluation criteria were satisfied. Reviewers must often sift through lengthy descriptions of the research process and determine the appropriate balance to be struck between a parsimonious description of the research and sufficient detail to demonstrate rigor and credibility. Furthermore, there is no comprehensive reference for the GT method. To obtain a thorough understanding of GT, reviewers must rely on several texts or articles focusing on specific aspects of the GT method. Many of these methodological developments have taken place outside the management literature (Jones and Noble, 2007).

Researchers frequently confuse GT with qualitative approaches such as case studies or content analysis and apply GT as a generic term to any qualitative approach that employs inductive analysis (Jones and Noble, 2007). Qualitative approaches frequently employ techniques such as coding and verification. This common terminology has produced situations of methodological slurring (Suddaby, 2006; Goulding, 2002) where researchers use interpretative methods to analyze realistic assumptions. Jones and Noble (2007) argue that GT is at risk of losing its integrity as many researchers have adopted only snippets of the methodology. As a result, GT research has often been perceived as lacking a consistent methodology, depth, and rigor (Mello and Flint, 2009).

The purpose of this paper is to develop a framework to assist reviewers in assessing the rigor and credibility of manuscripts employing a GT methodology. The supporting objectives include: analyzing articles appearing in the leading journals which employ GT methodology; identifying shortcomings in these articles regarding the application of the methodology; providing references to high quality GT studies to serve as a guide; distilling knowledge from multiple sources on GT research; and combining this knowledge with the best practices obtained from high quality articles. Four sections comprise the remainder of this paper. The next section identifies the need for a reviewer's framework. Section 3 provides a reviewer's perspective on GT development and the implications associated with aligning the method with a consistent paradigm of inquiry. The Section 4 summarizes the steps in GT analysis and identifies the major actions required to rigorously emerge a GT. The final section presents a framework of questions the reviewer should consider when evaluating GT research.

2. The need for a reviewer's framework

To understand the rigor of GT studies in the logistics and supply chain management, a review was conducted of articles published during the past 12 years appearing in the journals ranked among the top five in the discipline (Carter *et al.*, 2009; Gibson and Hanna, 2003; Menachof *et al.*, 2009): *Journal of Business Logistics*, *Transportation Journal*, *Journal of Supply Chain Management*, *International Journal of Physical Distribution and Logistics Management*, and *The International Journal of Logistics Management*. The review had two objectives. First, the articles were analyzed for

a rigorous application of the GT method to identify methodological issues, advances, and techniques for inclusion in the reviewer's framework. Second, articles with rigorous execution were identified as examples to guide reviewers.

The review produced an initial sample of 33 articles with "grounded theory" appearing in the title or abstract. Articles combining GT with other qualitative methods were included in the analysis but excluded from being identified as examples. 22 articles employed only GT. Of these, five were selected as examples based on a justification and comprehensive description of the methodology and rigorous analysis. Table I provides a summary of the five articles. The following paragraphs summarize the key issues identified during the review.

Introduction and literature review

The review suggested "methodological slurring" (Jones and Noble, 2007) in the majority of articles. Several mentioned GT in the introduction, but the subsequent discussion and methodology did not use GT. In other instances, studies combined GT with other qualitative approaches. These studies typically did not focus on theory development; instead, they presented findings drawn from rich descriptions and comparisons of the phenomenon in a limited number of instances.

Methodology

The review of the methodology produced two categories of articles. The first included research that mentioned GT but applied only bits and pieces of the method. For example, studies combining GT with case studies adopted techniques such as coding and constant comparative analysis. However, the studies began with a pre-selected number of cases. The coding tended to be descriptive rather than explanatory. Theoretical sampling and saturation were not used, and how a theory emerged from the development of categories or theoretical memos was unclear. In the second category, a GT approach was employed, but these articles did not clearly establish the context under investigation. In situations where the context was specified, a connection generally did not exist between the characteristics of the context and its relevance to the phenomenon. Specifying the context is important because the resulting theory can only explain phenomena within the context of the study.

Three major issues were identified with data collection. First, most studies employed in-depth interviews as the main source of data and stated that theoretical sampling was employed. The choice and number of participants was usually explained, but information regarding theoretical sampling of subsequent participants and the rationale for reaching theoretical saturation were weak or missing. For example, interviews stopped when no new information was obtained. This cursory treatment of theoretical sampling and theoretical saturation raises questions regarding methodological rigor. Second, most studies claimed to use multiple sources of data but rarely provided insights regarding how this data was analyzed or shaped the findings. Third, the majority of studies used an interview protocol but failed to provide or describe it which raises concerns regarding the objectivity of the questions, whether the questions biased the findings, and how the interview protocol evolved to support theoretical sampling.

Most articles did not adequately explain the analysis process stating merely that guidelines for open and selective coding were followed and constant comparative analysis was used. Justifications for which GT approach (Glaser, Strauss and Corbin,

Table I.
Review of selected GT articles

Name and year	Introduction and literature review	Methodology	Findings, discussion, and contributions
Mello *et al.* (2008)	Purpose is to develop a data-driven descriptive model of logistics outsourcing strategy development process Inconsistency between prescriptive and descriptive models. Outsourcing strategy development process not addressed in literature	Use both literature and field research to develop GT GT used for investigating processes where limited theory exists. No specific context. Companies from different industries. Purposive sampling based on extent of outsourcing. Theoretical sampling used. No protocol provided. Interview guide with specific questions, follow-up questions, and outline of research used. Unstructured interviews Coding followed established GT guidelines. Open coding to construct categories, properties and dimensions. Selective coding of core categories. Discovery of underlying uniformities Working hypotheses. Looking for negative or contradictory evidence	Criteria of fit, relevance, workability, modifiability, scope and parsimony. Descriptive model of logistics outsourcing process and additional factors developed
Mollenkopf *et al.* (2007)	Purpose is theory development within supply chain strategy. Limited attention to theory-based research in managing returns. Functional integration is ignored. Research seeks to understand linkages between marketing and logistics at strategic and operational levels as they deal with returns management	Qualitative method adopted as phenomenon not previously explored from cross-functional perspective and is poorly understood; no control over the phenomenon, and because other researchers have used the methodology. Firms in Italy chosen because of European economic integration and a pan-European regulatory environment. Research setting enables study of firm-level and external factors. Modified theoretical sampling. Industry sectors and supply chain echelons included to maximize diversity. Participants were influential decision makers. Invitations extended until diversity achieved. Bracketing-type technique for reducing bias. Open-ended questions. Interview guide used but not provided. Interviews digitally recorded. Information redundancy achieved via modified theoretical sampling, hence inferred that theoretical saturation reached. No description of process of data analysis	Two sets of criteria: (1) credibility, dependability, transferability, confirmability, and integrity (2) fit, understanding, generality, and control. A causal model is proposed. Four research propositions are offered
Manuj and Sahin (2011)	Purpose is to develop a model of supply chain decision-making complexity. Relationship between supply chain complexity and decision-making is poorly understood. A theoretical model needed to advance research. Identify specific research questions. Use literature to identify gaps	GT chosen because objective is to develop a comprehensive model grounded in actual business practice. No specific context mentioned. Initial sample based on participants' ability to provide relevant data, varied job titles and profiles, articulation skills, and willingness. Further participants selected based on emerging questions. Semi-structured interview process. Broad questions followed by specific and directed questions. Interview protocol provided. Interviews continued till theoretical saturation defined as "no new information or insights from additional interviews". Simultaneous collection, coding, and analysis of interview data. Secondary data sources also used	Trustworthiness evaluated based on credibility, dependability, transferability, confirmability, integrity, fit, understanding, generality, and control. A theoretical model developed and propositions offered

(*continued*)

Table I.

Name and year	Introduction and literature review	Methodology	Findings, discussion, and contributions
Randall *et al.* (2010)	Purpose is to present a theoretical framework for performance-based logistics from perspective of service-dominant logic. SDL suitability to advancing theory in supply chain management and PBL justified. The research problem stated as lack of a theoretical structure of PBL	GT appropriate since objective is to emerge a theory. PBL and SDL lack explanatory and predictive framework, studying how complex systems adapt to their environment. Defense context selected. A variety of participants from companies with PBL, no PBL, and converting to PBL were included. Multiple sources including case studies, archival review, practitioner conference interaction, and interviews used. Data collection, coding, and analysis described in detail. Constant comparative method, memoing, and theoretical sampling described. Theoretical saturation described as no new insights, themes, and issues. Additional interviews demonstrated consistent group of constructs and relationships between them. Detailed tables suggestive of extensive memoing	Evaluation based on fit. A detailed theoretical structure for PBL presented. SCM community introduced to SDL concept as a potential foundation for investigating phenomena
Flint *et al.* (2005)	Purpose is to construct a theory of logistics innovation from ground up and compare to extant literature. Concept of innovation is defined and explored in a logistics context. Specific research question proposed that deals with innovation in a logistics context	GT used due to social nature of innovation, problematic nature, likely importance of processes, and phenomena not deeply explored. No specific context selected. Variety of participants from logistics service provider and other industries selected. Relied on discovery-oriented, open-ended questions for interviews and site visits. Sampling explained in detail starting with convenience sample with theoretical sampling later. Sample of 33 participants identified as enough for theory building. Interview guide provided and modified based on emergent theory. Guidelines for GT briefly described. Used GT and hermeneutic processes for data analysis. Constant comparative method. Use of software. Theoretical memoing	No specific evaluation criteria presented. Discussion identifies specific contributions and ties them to existing literature. Managerial implications discussed. Specific future research directions identified

Charmaz, and others) was adopted or how the approach influenced the coding and subsequent analysis were starkly missing; studies explaining these elements were rare exceptions. Moreover, the information regarding the use of memos, development of working hypotheses or interim theories, and use of data other than interviews was limited.

Findings, discussion, and contributions

The review indicated a need to improve explanations regarding how the emergent theory satisfied the criteria for a GT and advanced the understanding of the phenomenon under investigation. The studies typically recognized that the proposed theory needed to conform to several criteria such as fit, relevance, and modifiability (Glaser, 1978). However, the majority of studies simply provided a list and description of key findings, which failed to provide any theoretical insight or establish connections between the new findings and existing theory.

Future research was frequently not linked to the context of the research and failed to explain how expanding the context could lead to formal, or eventually grand, theories with greater explanatory power. The researchers rarely addressed whether the paradigm of inquiry and GT approach adopted in the research support follow-on verificational studies and quantitative testing of the hypotheses and theory.

The results of this review produced five major implications for theory development which are summarized below.

Implications for developing and advancing GT in logistics and supply chain management:

(1) Continued methodological slurring due to confusion between GT and qualitative methods.

(2) Perceived lack of credibility and rigor of GT within the discipline.

(3) Grounded theories not accepted by researchers in other disciplines.

(4) Challenging review process resulting from a lack of methodological standards.

(5) Difficulty in obtaining acceptance in top journals.

3. A reviewer's perspective on theory development and adoption of a consistent paradigm of inquiry

An initial task for reviewers is to determine whether: GT is appropriate, the authors have adopted an approach consistent with their paradigm of inquiry, and the selected approach has been consistently applied throughout the research.

Appropriateness of GT

GT can be defined as "a systematic approach to qualitative research that facilitates theoretical abstraction from field data through a focus on comparative analysis" (Mello and Flint, 2009, p. 109). The objective is to develop "an abstract theoretical understanding of the studied experience" (Charmaz, 2006, p. 4). The role of the researcher is not to provide a perfect description of an area but to develop a theory that accounts for much of the relevant behavior. The resulting theory may "be presented as a well-codified set of propositions or in a running theoretical discussion, using conceptual categories and their properties" (Glaser and Strauss, 1967, p. 31).

Reviewers should anticipate the emergence of a substantive or formal theory. GT typically produces substantive, or mid-range theory. Substantive theories initially may

be developed through comparative analysis within the same substantive or empirical area to fill a gap in our theoretical understanding of a phenomenon. The research could develop a formal theory where comparative analysis occurs among several different substantive areas. Reviewers should note that the focus is not on achieving generalizability but on increasing the ability to explain the phenomena under study within and across substantive areas.

Systematic approach and consistency with paradigm of inquiry

GT employs a systematic approach for analyzing qualitative and quantitative data. A manuscript should describe a series of steps leading to the development of a theory. The GT approach initially developed by Glaser and Strauss (1967) responded to criticism that qualitative research was not methodologically sound to produce credible results. Reviewers can preempt such criticism by ensuring the manuscript contains an explanation of the approach and how theoretical insights emerged from the data. Reviewers must encourage, guide, and require authors to go beyond providing a "neat and tidy" descriptive account of the phenomena to conceptualizing a theory that captures the theoretical "whole" of a basic social process including stages, properties, conditions, and consequences (Glaser, 1978, p. 11). Reviewers should be wary of a "sleight of hand" that produces little insight into the analytical approach, does not add significantly to our understanding of the phenomena under investigation, relies on description, and invites the reviewer to take it on trust that theory somehow emerged from the data without providing an adequate explanation of development of theoretical insights (Barbour, 2001).

GT methodology continues to evolve since the initial publication of *The Discovery of Grounded Theory*. One philosophical paradigm is emergent (following Glaser), and the other is more formulaic and directive (following Strauss). Other variations such as the constructivist approach by Charmaz have also emerged.

A reviewer must assess whether the use of GT is congruent with the paradigm of inquiry adopted in the manuscript (Annells, 1996) because divergent philosophical approaches to GT have implications for the conduct of the research in terms of process as well as outcomes of research (Mello and Flint, 2009; Annells, 1996; Walker and Myrick, 2006; Goulding, 2002; Kelle, 2005; Duchscher and Morgan, 2004). Guba and Lincoln (1994) identified four paradigms of inquiry: positivisim, postpositivism, critical theory, and constructivisim. These divergent paradigms guide the interaction of the researcher with the data, introduction of outside concepts and conditions, and the use of subsequent or sequential verification studies.

The "classic" school of GT methodology advocated by Glaser has frequently been identified as postpositivist due to its roots in social interactionism and an objectivist orientation (Annells, 1996; Goulding, 2002). Reviewers should find the researchers remain separate from the method, do not introduce outside concepts and constructs, and strictly adhere to the premise that theory should be allowed to emerge from the data. This paradigm accepts subsequent verificational research which is consistent with Glaser (1992) – GT "should be seen in sequential relation" to verificational research with the ultimate aim of both being the "building up of scientific facts". The key implications for reviewers when evaluating manuscripts adopting classic GT include ensuring that the researchers remained objective and outside the method, allowed the theory to emerge from the data, did not interact with the data by introducing outside concepts or constructs, and accept sequential verificational research.

The approach to GT advocated by Strauss and Corbin (1990, 1998) falls in the constructivist paradigm of inquiry but contains remnants of postpositivism (Annells, 1996). The reviewer should find the researcher interprets and interacts with the data to create knowledge. This approach imposes a coding paradigm to explain the construction of a framework of grounded categories (Kelle, 2005). A conditional matrix is employed which is likely to introduce issues such as class, gender, race, and power into the analysis (Annells, 1996). Verification is not assumed to be possible only through follow-on quantitative analysis but instead can occur through constant comparative data analysis throughout the course of the research (Strauss and Corbin, 1994). The key implications for reviewers when evaluating manuscripts adopting this approach are: the researchers became participants in the research by interacting and interpreting the data; the research introduces outside conditions, concepts, or constructs into the analysis to assist in constructing a theory; and verification may occur through constant comparative analysis.

In summary, the reviewer needs to determine whether the GT method employed is consistently applied, and is consistent with the paradigm of inquiry and the philosophical approach adopted in the research. Reviewers will find few differences in the methodology; a few differences were discussed in the previous paragraph (Annells, 1996; Goulding, 2002; Charmaz, 2004 for detailed explanations of these differences). However, the paradigm of learning adopted in the paper will affect the analysis and interpretation of the data, whether the research and resulting GT take an objectivist, positivist stance or a constructivist, interactionist stance, and whether subsequent verificational research is recommended for future research.

4. Major steps in a GT development

This section describes the major steps within a GT analysis and actions for each step that should be considered when evaluating a manuscript. The eight steps (Figure 1) are presented in a linear fashion; however, GT proceeds in a non-linear, iterative manner.

Step 1: explaining a phenomenon

Reviewers should consider why existing theory does not adequately explain a phenomenon. The authors may have reviewed the literature to reveal a lack of relevant theory, the inability of existing theory to fully explain the phenomenon, or the need to develop formal theory from previous substantive theories. Some manuscripts do not contain a literature review. A common misconception is researchers should enter the field devoid of any knowledge (Suddaby, 2006) to minimize preconceptions and bias. The leading grounded researchers have largely debunked this notion (Flint *et al.*, 2005; Charmaz, 2004; Goulding, 2002) and suggest that authors understand the substantive area and can explain and interpret it in their work (Glaser, 1978).

Theoretical sensitivity refers to the ability of a researcher to work with the data in both theoretical and sensitive ways (Glaser, 1978). Reviewers need to assess the ability of the authors to "theoretically and conceptually think about the data from a distance, while simultaneously maintaining an in-close level of sensitivity and understanding about the process and their involvement in that process" (Walker and Myrick, 2006, p. 552).

Reviewers can take several cues to assess the theoretical sensitivity of the authors. The phrasing of the problem as a general, broad statement that leaves open possible theoretical outcomes suggests the authors understand the substantive area but have left possible alternative explanations open for discovery. The research questions

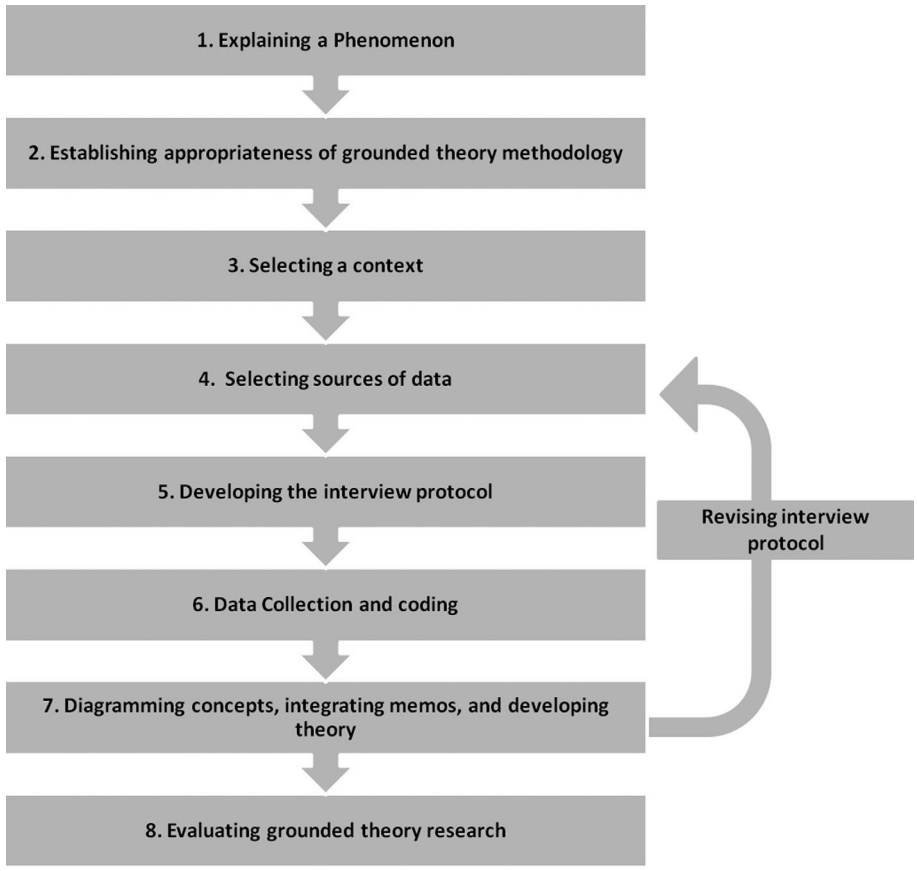

Figure 1.
Major steps in a grounded theory development

should identify the need for theory, further explanation, or interpretation in a substantive area. The authors must avoid verifying existing theory or hypotheses during the research.

Step 2: establishing the appropriateness of GT methodology

The goal of GT is to conduct research, grounded in empirical data, which increases our understanding of a phenomenon by developing theory with explanatory and predictive power (Glaser and Strauss, 1967, p. 24). Therefore, the justification should indicate an attempt to conceptualize how or why the phenomenon occurs. The reviewer must ensure that the justification goes beyond simply stating that GT was employed due to the exploratory nature of the research, that qualitative research was necessary due to a gap in extant research, or that authors adopted GT to describe a phenomena in detail. Fischer and Otnes (2006) identified four areas that support the use and goals of GT: investigating questions about the nature of a new construct; raising questions about the adequacy of a previously well-established construct; investigating previously unrecognized facilitators or implications of a construct; and addressing questions about the adequacy of prior conceptualizations of facilitators or implications of a construct.

Step 3: selecting a context

The reviewers need to consider the context employed within which the phenomenon is investigated. "A context is the set of conditions that give rise to the problems or circumstances to which individuals respond by means of action/interaction/emotions" (Corbin and Strauss, 2008, p. 229). Some practical considerations may guide the selection of context such as access to: willing companies, companies that can provide rich data, and knowledgeable people that are experiencing or have experienced the phenomenon. While convenience factors are acceptable, the reviewers must primarily look for the rationale for fit between the phenomenon, research questions, and the context.

Step 4: selecting sources of data

Reviewers should detect clear evidence of a multi-step, iterative process for data collection and sample selection. The categories and relationships that will emerge as the research progresses cannot be known in advance. Therefore, the criteria for collecting data during initial sampling depend upon the research questions posed before the investigation commences. This approach is frequently identified to as "purposive sampling" and refers to the calculated decision to sample a specific set of events, settings, or people based on criteria that are decided in advance. Reviewers should check the adequacy of the initial sample by asking questions such as whether the initial participants: fit the context, had visibility over part or the entire phenomenon, were knowledgeable, willing to participate, and experienced with, and engaged in the phenomenon being studied. To judge the reliability of data, reviewers may pose questions regarding how access was gained to obtain the appropriate data for the research.

Reviewers must be convinced theoretical sampling was employed for subsequent data collection. Theoretical sampling is the:

> [. . .] process of data collection where the analyst collects codes, and analyzes the data and decides what data to collect next and where to find them based on the emergent theory (Mello and Flint, 2009, p. 112).

Evidence that coding and constant comparative analysis guided the authors in determining the participants or data sources that were subsequently selected must be sought. Manuscripts often report the number of interviews or data sources used in the research which may indicate the authors did not go beyond an initial sample. A description of how sources were initially and subsequently selected during the research process is as important as reporting the sample size and composition.

As codes and categories are developed, the research should reflect their theoretical development with properties and connections to other categories until each category is saturated. Theoretical sampling continues until each category is saturated, elaborated, and integrated into an emerging theory (Glaser, 1992). When a category is saturated, constant comparison will not reveal new properties or dimensions for the category or relationships between categories. Additional information may be sought to fill in gaps in the emerging theory or to further understand differences detected between incidents. Research may employ theoretical sampling numerous times (see Step 7) during the investigation as one category is saturated and the research progresses to saturate other categories or exploring the relationships between them.

Reviewers should be wary of the four common mistakes related to use of theoretical sampling: sampling to address initial research questions, sampling to reflect population

distributions, sampling to find negative cases, and sampling until no new data emerge (Charmaz, 2009, p. 100). Initial sampling represents a point of departure; whereas, theoretical sampling is used to saturate categories or concepts for theoretical refinement and elaboration. Sampling to reflect population distributions implies verification or hypothesis testing. Theoretical sampling, on the other hand, focuses on increasing explanatory power and our understanding of a phenomenon. Sampling to obtain negative cases may be inconsistent with GT. When negative cases emerge in the data, theoretical sampling may be applied to understand the negative cases with the resulting theory being modified to incorporate these cases. The reviewer should challenge instances where negative cases are sought out or artificially imported into the analysis. Finally, many authors confuse theoretical sampling with data gathering (Charmaz, 2004, p. 102) and do not satisfactorily achieve theoretical saturation.

Step 5: developing the interview protocol

The reviewer should check whether the researcher developed an interview protocol or observation guide. The initial set of questions may be based on insights derived from literature, experience, or preliminary field work and should typically be broad and general in scope and reflect authors' theoretical sensitivity. Since the objective is to emerge a theory from data, reviewers need to look for evidence that these initial questions were reworded, refined, or even discarded over the course of the research (see revision loop in Figure 1).

The interview protocol should be open-ended, unstructured, and allow the authors to probe new ideas as they emerge. The interview questions should avoid the forcing of preconceived theories or concepts. Reviewers should find the questions were initially broad but became more focused as the research progressed and theoretical sampling was employed. Reviewers should look for the logic of modifying the interview protocol and find evidence that the interview protocol was modified in order to refine constructs, develop theory, or to increase the depth of understanding of the phenomenon.

Step 6: data collection and coding

Data collection and coding represent the first stage of the analytic process within GT. The manuscript should explain how data will be captured. If the interview protocol indicates that the permission of interview participants to record the interview was obtained or other sources of data were captured for future use, then it is easier to establish that data can be accurately captured and revisited, and if needed, multiple researchers can review the same data. If the data is not captured in its entirety, then reviewers need confirm how the accuracy of data capture was maintained.

Comparative analysis, or constant comparison, represents a core tenet of GT. Comparisons, both within and between incidents and with new and previous data, should have been used to identify similarities, differences, trends and patterns in the data as well as to make modifications of the emerging theory. These comparisons support the generation of abstract categories and properties which can be used to develop a theory explaining the behavior under observation (Glaser and Strauss, 1967). Although presented in a sequential manner, reviewers should be able to determine whether the research actually iterated numerous times across multiple steps (Figure 2).

Reviewers need to consider several points when evaluating the coding, memoing, and theoretical sampling phase of the methodology. First, they must determine

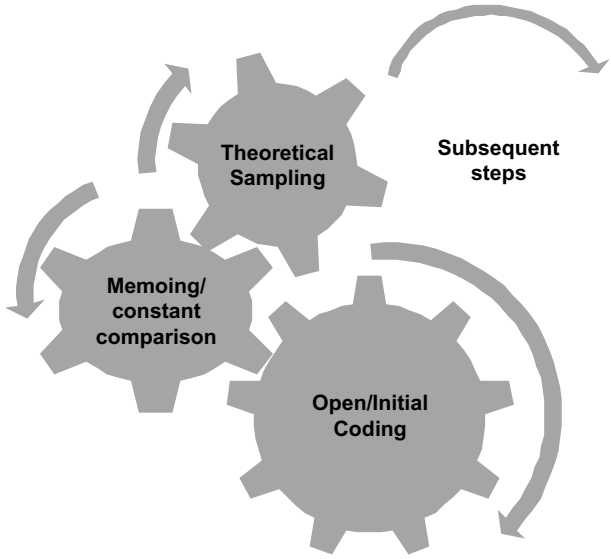

Figure 2.
Iterative nature of the grounded theory methodology

whether coding occurred at a descriptive or conceptual level. Descriptive coding simply describes a phenomenon and will provide little to no explanatory power. Conceptual coding and the resulting categories will reflect greater theoretical insight such as why or how the phenomenon is occurring.

Second, the use of constant comparative analysis should appear in the description of how new codes were created, the formation of selective or theoretical coding, analyses captured in theoretical memos, the discovery of categories and properties, the resulting relationships between categories, and theory development. Reviewers will need to determine how the authors employed theoretical memos to raise their coding to a conceptual level and to emerge categories, to identify tentative relationships between them and to identify the properties and dimensions of each category.

Third, constant comparison must be used to guide theoretical sampling. Reviewers will need to determine whether the authors used the data to guide theoretical sampling to obtain relevant data that would increase their insight and understanding. Finally, the constant comparison of data within the research should reflect a focused flexibility to keep transcending the analyses until the problem is theoretically saturated (Glaser, 1978). The reviewer must be convinced that constant comparison and the analysis of data occurred simultaneously and throughout the investigation process.

Step 7: diagramming concepts, integrating memos, and developing theory

The next phase of the GT methodology focuses on emerging theory from the codes, categories, and memos developed during the previous phase of research.

Theoretical saturation represents the stopping point for data collection and the "sample size" for the research. Reviewers should not accept statements that saturation was achieved after accomplishing an arbitrary number of interviews. For example, "Theoretical saturation was achieved after 17 interviews." The research should report how saturation was reached and not rely on arbitrary numbers, frequency counts, or citing repetition in the data.

Saturation occurs when constant comparison emerges no new properties or dimensions for each of the categories (Holton, 2010). "Saturation is not seeing the pattern over and over again. It is the conceptualization of these incidents which yields different properties of the pattern until no new properties emerge" (Glaser, 2001, p. 191). Unfortunately, no published guidelines or tests of adequacy exist for estimating the sample size required to reach saturation. Despite the lack of guidelines, the reviewers should ensure that the process used to reach saturation was rigorous, thorough and transparent. The authors "should specify what they did and how they did it, including how they handled data condensation and interpretation" (Bowen, 2008, p. 149).

Memo, or theoretical, sorting represents another process common to the different forms of GT. During this process, the reviewers should learn how the researchers created and defined links between categories (Charmaz, 2006). Sorting frequently produces more memos of a higher conceptual level that aid in producing a generalized, integrated model by which to write the theory. Reviewers may find that the relevant literature was integrated into the theory (Glaser, 1978). As part of their evaluation, reviewers should seek to understand how the authors used sorting to develop relationships between categories and used the memos to produce a theoretical framework and generate an integrated, dense and complex theory.

Reviewers will confront very different methods in integrating the categories into a final GT. Glaser (1978) suggests the use of theoretical coding to conceptualize how the categories relate to each other and may be integrated into a theory. To assist in this effort, he developed 18 theoretical coding families, but others may exist. These theoretical codes assist in developing a coherent, analytical story and help move it in a theoretical direction. The theoretical categories are eventually subsumed into a core category that:

> [. . .] pulls together all the strands in order to offer an explanation of the behavior under study. This is usually done when the theory is written and integrated with existing theories to show relevance and perspective (Goulding, 2002, p. 88).

Strauss and Corbin (1990, 1994) take a very different approach for integrating the final theory. They employ axial coding and a conditional matrix (Strauss and Corbin, 1990), or a conditional/consequential matrix (Corbin and Strauss, 2008), to enable the researcher to become theoretically sensitive in systematically relating conditions, action/interaction and consequences to a phenomenon. The matrix is operationalized by tracing conditional paths through it. Tracing paths of incidents from the level of action/interaction through the various conditional levels to consequences is used to determine how categories relate to one another.

Step 8: evaluating GT research

Finally, the reviewers should check whether the authors established the credibility and trustworthiness of the study. Several sets of criteria may be employed to evaluate the quality of a GT research. Charmaz (2006) suggests four criteria: credibility, originality, resonance, and usefulness. Glaser (1978) suggests fit, work, relevance, and modifiability. Flint *et al.* (2002) employ five criteria from interpretive research – credibility, transferability, dependability, confirmability, and integrity (based on Hirschman, 1986; Lincoln and Guba, 1985; Wallendorf and Belk, 1989) and four from GT – fit, understanding, generality, and control (based on Strauss and Corbin, 1990).

The reviewers should look for an explicit description of how the authors addressed at least one set of criteria. This set of criteria, in turn, must fit with the GT approach adopted by the authors.

The final product of a GT approach is a theory or theoretical framework. The reviewers need to determine whether the author has successfully explained how the theory is grounded in the data, that is, the methodology supports the development of theory. The reviewers must be convinced that the theory or theoretical framework logically flows from the research effort.

In terms of contribution to the body of knowledge, the reviewer should look for clear, meaningful, and significant theoretical contribution within the context in which the theory is proposed. Reviewers need to establish that the authors compared the theory to the extant literature to establish if the proposed theory is different or the proposed theory assists in our understanding of the literature. Any differences need to be described or explained. Reviewers should be convinced that the proposed theory increases our understanding of the phenomenon and provides hypotheses or propositions from the theory to guide future research. The reviewers must also look for managerial relevance and usefulness.

Table II summarizes the key considerations for each step. This list of questions can be used by the reviewer to ensure the methodology has been appropriately executed. Reviewers may also find the table useful in making a comprehensive evaluation of the research and the overall contribution.

5. Conclusions

GT can play an important role in the development of theory grounded in logistics and supply chain management, but the methodology continues to confront several hurdles which must be overcome. A review of 33 articles in the leading journals detected problems similar to those experienced in other disciplines: the lack of a clear distinction between GT and qualitative analyses, methodological slurring, unclear or missing descriptions of key methodological steps, a missing focus on theory development, and resulting perceptions of a lack of credibility and rigor. However, several articles demonstrated effective execution of the methodology and can guide reviewers in their assessment of manuscripts.

Reviewers can play a critical role in ensuring the rigor and credibility of published articles. However, reviewers confront several challenges, including making inferences in light of incomplete details on execution of research, deciding what to include while trying to meet the journal space constraints, and keeping pace with the evolution of the methodology. The framework put forth can serve as a guidepost to assist reviewers in two ways. First, the framework provides the two key characteristics of GT methodology which reviewers can use for assessing whether congruency exists between the research objectives, theory development approach, and the paradigm of inquiry adopted in the research. This discussion sensitizes the reviewers to the iterative, unstructured and fluid nature of the research methodology, such as evolving (rather than fixed) research questions, continually changing (rather than preset) interview protocol, and theoretical sampling (instead of a pre-determined sample). Second, the framework provides an eight step checklist for reviewers to use for assessing the execution and rigor of the GT methodology. Each step of the framework is elaborated in terms of what the reviewers should consider when evaluating the quality of a GT submission. Reviewers should avoid imposing or advocating

Table II.
A reviewer's guide to evaluating the quality of GT research

Step	Questions to ask
Step 1: describing the phenomenon	Does the author identify a process or phenomenon that will be examined and explain why theory is needed to increase our understanding? Does the author describe the lack of an extant theory capable of explaining the phenomenon? Does the author explain that the purpose is to generate a theory rather than achieve generalizability? Does the author address theoretical sensitivity? How does the author identify the theoretical issues and explain how the research will remain open to the emergence of new constructs, relationships and theory? Does the phrasing of the problem statement indicate that the researchers have remained open for discovery?
Step 2: establishing appropriateness of GT methodology	Is the problem statement phrased to indicate the need for a theory? Does the author provide justification for GT beyond the need to perform an exploratory study? Does the author explain which approach within the GT methodology is accepted? Are the key implications of adopting an approach within the GT methodology consistent with the approach in terms of author's interaction with the data, introduction of outside concepts and conditions, and the use of subsequent or sequential verification studies?
Step 3: selecting a context	Does the author explain why the context is appropriate for the purpose of the research? Is the context selected appropriate for building theory of the phenomenon or process being investigated?
Step 4: selecting sources of data	Does the author explain how the access will be gained to obtain the appropriate data for the research –, i.e. access to key personnel, data, observation points, etc. to conduct the research? Does the author explain how the initial participants were selected? Does the researcher indicate that the sample of firms/participants was identified in advance or was based on data collection? Does the author provide evidence that coding and constant comparative analysis guided the author in determining the follow-on sources or participants that were subsequently selected? Does the author indicate that theoretical sampling was continued until saturation was achieved? Does the author indicate that the purpose of theoretical sampling was to achieve more explanatory power and *not* to achieve generalizability of the results?
Step 5: developing the interview protocol	Are the interview questions phrased to avoid the forcing of preconceived theories or concepts? Do the interview questions allow the theory or framework to emerge from, or be grounded in, the data collected during the investigation? Is the interview protocol open-ended and structured in that it allows the author to probe new ideas as they emerge? Does the author indicate that interview protocol was modified to include more focused questions as interview progressed? Does the author provide evidence that interview protocol was modified to refine constructs, develop theory, or to increase the depth of understanding of the phenomenon?
Step 6: data collection and coding	How is the author capturing data? Are the interviews taped and recorded? Are the interviews transcribed? Is the author employing any tools to assist in the research such as MaxQDA? Does the author explain how the initial coding was performed? Does coding occur while the data collection is being conducted? How do the codes evolve over time? Does the author indicate how memos were used to capture ideas, potential categories, relationships, differences, or similarities in the data? How is constant comparison incorporated into the analysis of the data?

(*continued*)

Step	Questions to ask
	Do the authors provide evidence/examples that codes emerge from the data? Do authors provide evidence that codes are evolving as constant comparisons are made?
	What does the author report regarding the use of comparisons to refine categories and obtain a richer understanding of the phenomenon under investigation?
	How are memos used to identify themes that emerge from the data?
	How does the author report gaps in the investigation? How are the gaps targeted for follow-on investigation, or theoretical sampling?
Step 7: diagramming concepts, integrating memos, and developing theory	Theoretical saturation
	How does the author explain theoretical saturation? Does the author go beyond stating that "nothing new" was found and explain that the categories were determined to be saturated when no new theoretical insights or no additional information was obtained regarding the properties of the core categories?
	Does the author provide examples of how saturation was determined to have been achieved?
	Theoretical sorting
	Does the author explain how a theory or theoretical framework was accomplished through sorting, diagramming, or integrating theoretical memos developed during the research?
	How is the result of this process integrated into a theory or framework? Is the process of integrating the categories into a final GT consistent with the approach adopted (Glaserian, Straussian or any other)?
	Developing theory
	Does the author indicate whether theoretical sampling was used to enrich categories, to fill gaps in understanding, or to explore leads identified during the research?
	Does the author explain how focused coding occurred and how memos were used to develop the focused codes?
	How are conceptual categories identified and refined during the research?
	Is evidence from the research provided to support ideas or patterns?
Step 8: evaluating GT research	How does the author evaluate the research? Does the author use criteria such as Charmaz (credibility, originality, resonance, usefulness), Glaser (work, relevance, modifiability), Flint (credibility, transferability, dependability, confirmability, integrity), or Strauss and Corbin (fit, understanding, generality, control)?
	Does the author establish credibility with the reviewer?
	Does the author provide a theory or theoretical framework?
	Does the author explain the scope, based on the context, for the theory?
	Has the author successfully explained how the theory is grounded in the data – , i.e. how the investigation supports the development of theory?
	Does the theory or theoretical framework logically flow from the research effort?
	Was member checking, or some other technique, used to confirm that the author's analysis is consistent with the data (participants)?
	Does the researcher convincingly and clearly establish theoretical contribution?
	Does the author compare the theory to the extant literature? How does the theory assist in our understanding of the literature? How is this theory different? Why any differences?
	How does the theory increase our understanding of the phenomenon?
	Does the author propose any hypotheses or propositions from the theory to guide future research?
	Does the research establish managerial relevance and usefulness?

Table II.

an approach when reviewing a manuscript. Instead, they should ensure the authors have clearly identified the approach adopted for research, recognize the implications in their analysis and theory development, and remain consistent with the version of GT methodology employed.

Authors may err on the side of providing too many details than risk rejection if the manuscript is perceived as not being rigorous. Or, authors may not include adequate details because of space constraints. Reviewers should be prepared for either approach, to establish a dialog with authors, and to make recommendations on striking a balance between parsimony and credibility in the final version of the manuscript approved for publication.

Limitations

The framework is not a substitute for in-depth knowledge of GT methodology. Readers may refer to the references section for well-respected and widely accepted sources of knowledge on the methodology. The contribution of this paper is that for researchers educated in the methodology, it provides a comprehensive set of considerations for evaluating a manuscript. For reviewers with limited knowledge, the framework provides a fast and easy reference to start the process of learning about the methodology.

Contributions

More effective reviews of GT research are expected to help the body of knowledge in three ways. First, reviewers aware of critical methodological issues will act as gatekeepers ensuring that only credible and high quality research is published. This awareness will also prevent qualitative research being evaluated against inadequate, or even worse, incorrect criteria. The discussion on consistency among the paradigm of inquiry, philosophical approach, and method of GT provides a foundation to evaluate GT research. These considerations may lead to higher rigor with which the GT methodology is evaluated and increase the credibility of GT as a valid methodology. Second, reviewers who are skeptical of the rigor of GT methodology can learn about the methodology from the discussion and use the checklist to establish the rigor of a GT study rather than rejecting or discarding research. Third, the discussion and examples provided in this paper and the table could be used to provide thorough and constructive reviews that will enable researchers to improve the quality of their current and subsequent research efforts.

Summary

The execution of GT research in logistics and supply chain management has yet to receive wide-spread acceptance. The method continues to encounter skepticism due to a perceived lack of rigor and credible findings. However, well-executed GT research can make significant contributions to the discipline through theory development and increasing our understanding and ability to explain key phenomenon. Through a combination of an extensive review of the method and leading GT theory research published in the leading journals, a framework was developed to assist reviewers. By increasing their expertise and focusing their attention on key steps and considerations within the GT method, the framework can assist reviewers in their critique and of GT research. A more robust and thorough review of GT manuscripts will not only increase rigor and credibility but advance the development of theory grounded in logistics and supply chain data and practice.

References

Annells, M. (1996), "Grounded theory method: philosophical perspectives, paradigm of inquiry, and postmodernism", *Qualitative Health Research*, Vol. 6 No. 3, pp. 379-93.

Barbour, R.S. (2001), "Checklists for improving rigour in qualitative research: a case of the tail wagging the dog?", *British Medical Journal*, Vol. 322 No. 7294, pp. 1115-17.

Bowen, G.A. (2008), "Naturalistic inquiry and the saturation concept: a research note", *Qualitative Research*, Vol. 8 No. 1, pp. 137-52.

Carter, C.R., Easton, P., Vellenga, D. and Allen, B.J. (2009), "Affiliation of authors in transportation and logistics academic journals: a reevaluation", *Transportation Journal*, Vol. 48 No. 1, p. 42.

Charmaz, K. (2004), "Premises, principles, and practices in qualitative research: revisiting the foundation", *Qualitative Health Research*, Vol. 14 No. 7, pp. 976-93.

Charmaz, K. (2006), *Constructing Grounded Theory*, Sage, Thousand Oaks, CA.

Charmaz, K. (2009), "Shifting the grounds: constructivist grounded theory methods", in Morse, J.M., Stern, P.N., Corbin, J.M., Bowers, B. and Clarke, A.E. (Eds), *Developing Grounded Theory: The Second Generation*, University of Arizona Press, Walnut Creek, CA.

Corbin, J. and Strauss, A. (2008), *Basics of Qualitative Research*, 3rd ed., Sage, Thousand Oaks, CA.

Davis-Sramek, B. and Fugate, B.S. (2007), "State of logistics: a visionary perspective", *Journal of Business Logistics*, Vol. 28 No. 2, pp. 1-34.

Duchscher, J.E.B. and Morgan, D. (2004), "Grounded theory: reflections on the emergence vs. forcing debate", *Journal of Advanced Nursing*, Vol. 48 No. 6, pp. 605-12.

Fischer, E. and Otnes, C.C. (2006), "Breaking new ground: developing grounded theories in marketing and consumer behavior", in Belk, R.W. (Ed.), *Handbook of Qualitative Research Methods in Marketing*, Edward Elgar, Northampton, MA, pp. 19-30.

Flint, D.J., Woodruff, R.B. and Gardial, S.F. (2002), "Exploring the phenomenon of customers' desired value change in a business-to-business context", *Journal of Marketing*, Vol. 66 No. 4, pp. 102-17.

Flint, D.J., Larson, E., Gammelgaard, B. and Mentzer, J.T. (2005), "Logistics innovation: a customer value-oriented social process", *Journal of Business Logistics*, Vol. 26 No. 1, pp. 113-48.

Gibson, B.J. and Hanna, J.B. (2003), "Periodical usefulness: the US logistics educator perspective", *Journal of Business Logistics*, Vol. 24 No. 1, pp. 221-40.

Glaser, B.G. (1978), *Theoretical Sensitivity*, The Sociology Press, Mill Valley, CA.

Glaser, B.G. (1992), *Basics of Grounded Theory Analysis*, The Sociology Press, Mill Valley, CA.

Glaser, B.G. (2001), *The Grounded Theory Perspective: Conceptualization Contrasted with Description*, The Sociology Press, Mill Valley, CA.

Glaser, B.G. and Strauss, A.L. (1967), *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Aldine Transaction, Piscataway, NJ.

Goulding, C. (2002), *Grounded Theory: A Practical Guide for Management, Business, and Market Researchers*, Sage, Thousand Oaks, CA.

Guba, E.G. and Lincoln, Y.S. (1994), "Competing paradigms in qualitative research", in Denzin, N.K. and Lincoln, Y.S. (Eds), *Handbook of Qualitative Research*, Sage, Thousand Oaks, CA, pp. 105-17.

Hirschman, E.C. (1986), "Humanistic inquiry in marketing research: philosophy, method, and criteria", *Journal of Marketing Research*, Vol. 23 No. 3, pp. 237-49.

Holton, J.A. (2010), "The coding process and its challenges", *The Grounded Theory Review*, Vol. 9 No. 1, pp. 21-40.

Jones, R. and Noble, G. (2007), "Grounded theory and management research: a lack of integrity?", *Qualitative Research in Organizations*, Vol. 2 No. 2, pp. 84-103.

Kelle, U. (2005), "Emergence vs. forcing of empirical data? A crucial problem of grounded theory reconsidered", *Forum Qualitative Sozialforschung* (*Forum: Qualitative Social Research*), Vol. 6 No. 2, Art. 27, available at: www.qualitative-research.net/index.php/fqs/article/view/467/1001 (accessed August 30, 2011).

Lincoln, Y.S. and Guba, E.G. (1985), *Naturalistic Inquiry*, Sage, Newbury Park, CA.

Manuj, I. and Sahin, F. (2011), "A model of supply chain and supply chain decision-making complexity", *International Journal of Physical Distribution & Logistics Management*, Vol. 41 No. 5, pp. 511-49.

Mello, J. and Flint, D.J. (2009), "A refined view of grounded theory and its application to logistics research", *Journal of Business Logistics*, Vol. 30 No. 1, pp. 107-25.

Mello, J.E., Stank, T.P. and Esper, T.L. (2008), "A model of logistics outsourcing strategy", *Transportation Journal*, Vol. 47 No. 4, pp. 5-25.

Menachof, D.A., Gibson, B.J., Hanna, J.B. and Whiteing, A.E. (2009), "An analysis of the value of supply chain management periodicals", *International Journal of Physical Distribution & Logistics Management*, Vol. 39 No. 2, pp. 145-66.

Mollenkopf, D., Russo, I. and Frankel, R. (2007), "The returns management process in supply chain strategy", *International Journal of Physical Distribution & Logistics Management*, Vol. 37 No. 7, pp. 568-92.

Randall, W.S., Pohlen, T.L. and Hanna, J.B. (2010), "Evolving a theory of performance-based logistics using insights from service dominant logic", *Journal of Business Logistics*, Vol. 31 No. 2, pp. 35-61.

Strauss, A. and Corbin, J. (1990), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Sage, Newbury Park, CA.

Strauss, A. and Corbin, J. (1994), "Grounded theory methodology: an overview", in Denzin, N.K. and Lincoln, Y.S. (Eds), *Handbook of Qualitative Research*, Sage, Thousand Oaks, CA, pp. 273-85.

Strauss, A. and Corbin, J. (1998), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, 2nd ed., Sage, Thousand Oaks, CA.

Suddaby, R. (2006), "What grounded theory is not", *Academy of Management Journal*, Vol. 49 No. 4, pp. 633-42.

Walker, D. and Myrick, F. (2006), "Grounded theory: an exploration of process and procedure", *Qualitative Health Research*, Vol. 16 No. 4, pp. 547-59.

Wallendorf, M. and Belk, R.W. (1989), "Assessing trustworthiness in naturalistic consumer research", in Hirschman, E. (Ed.), *Interpretive Consumer Research*, Association for Consumer Research, Provo, UT, pp. 69-84.

About the authors

Ila Manuj, PhD is an Assistant Professor of Logistics at the Department of Marketing and Logistics, University of North Texas. Her research interests are in the areas of risk and complexity management in global supply chains. She has published in *Journal of Business Logistics*, *International Journal of Physical Distribution & Logistics Management*, *International Journal of Logistics Management*, and *Transportation Journal*, co-authored several book

chapters, and presented at several national and international conferences. Ila Manuj is the corresponding author and can be contacted at: ila.manuj@unt.edu

Terrance L. Pohlen, PhD, is an Associate Professor and Director, Center for Logistics Education and Research, University of North Texas. His research focuses on the costing and financial management of logistics and supply chain management. Recently, he co-authored CSCMP's *The Handbook of Supply Chain Costing*. His research has been published in the *Journal of Business Logistics*, *International Journal of Logistics Management*, *International Journal of Physical Distribution & Logistics Management*, and the *Transportation Journal*. He chairs the professional certification in transportation and logistics and is an AST&L board member.