Analysis of Current Grounded Theory Practices

Corina Radulescu, University of Sydney
Iris Vessey, University of Queensland

BIS WP2011-01
Analysis of Current Grounded Theory Method Practices

Abstract
Use of the grounded theory research method (GTM) is increasing across many fields of inquiry. Understanding the GTM and how to apply it is therefore a key task for researchers examining the possibility of using this method in their research. Since its introduction by Glaser and Strauss in 1967, GTM has evolved into two major streams, and there has been a continual debate about the choice between them and their applicability. Examination of the extant literature reveals significant problems in applying GTM. In this paper, we take a first step in the quest for identifying current GTM practices and providing more effective procedures for conducting GTM research. To achieve our goal we start by analysing and identifying a number of difficulties encountered by researchers who have used GTM. We then examine and present ways in which they have resolved and addressed the problems.
Introduction
The grounded theory research approach (Glaser and Strauss 1967) is a method for addressing research questions for which little extant theory exists. It is a qualitative method that seeks to discover the essence of a phenomenon from rich empirical data (Schwandt 1997); that is, the theory emerges from, and is rooted in, the data. Grounded Theory Method (GTM) is most commonly used in association with case study research, which relies largely on data from interviews, observations, documents, audio and video materials, and other physical artifacts (Dey 1999; Strauss and Corbin 1998). The data used is therefore textual in nature. A major difference between GTM and other qualitative research methods is that GTM is based on continuous interplay between data collection and analysis (Urquhart et al. 2009).

Glaser and Strauss (1967) introduced the GTM as a method for conducting empirical research in sociology. The method (or methodology) has, however, been used in a wide variety of disciplines, among them, nursing, health, management, and marketing (see, for example, Becker 1993; Goulding 2004; Kendall 1999; Morse 2004; LaRossa 2005). GTM has also been used increasingly in information systems research, as evidenced by a number of publications in top journals (see, for example, Orlikowski 1993; Sarker et al. 2001; Volkoff et al. 2007).

Although, at first glance, GT methods may appear to be easy to operationalise, students or novices have experienced significant difficulties in applying them (see, for example, Kendall 1999; Partington 2000; Scott 2004). And indeed, the initial motivation for our paper stems from issues that challenged us in our attempt to use GTM, and was further motivated by the difficulty in finding solutions. While the literature on the theoretical and philosophical underpinnings of GTM is plentiful, there is little literature on how GT methods are actually applied (see, also, Balint (2003, p. 35) and Kendall (1999, p. 749)). It seems likely that difficulty in conducting GT research based on the procedures offered may explain why current studies using GTM contain few details of how the findings of “GT studies” are actually derived. Initial insight into some of the problems with GT research comes from studies of common pitfalls and misunderstandings (see, for example, Becker 1993; Suddaby 2006).

The objective of this paper is to take a preliminary step toward analysing and providing a rigorous, workable, and well-defined set of procedures for conducting GT research. We do so by examining a number of published articles (reflective or confessional in nature) that specifically focused on the issues researchers confronted in conducting a study using GTM and the ways in which they resolved them. The practical focus of our analysis differs from the majority of such analyses, which focus largely on the divide between the approaches of Glaser and Strauss. We believe that researchers, especially those engaging in GT research for the first time, will derive significant benefits from such a paper. First, it highlights important differences between the two streams of GT research that could help the researcher decide from the outset whether to use Glaser’s approach (1987, 1992) or that of Strauss and Corbin (1990, 1998). Second, it highlights a number of important problems that researchers have faced in conducting GT research. Third, it presents the ways in which they have actually addressed those problems and therefore may be particularly relevant for researchers conducting or seeking to conduct GT research.

The paper proceeds as follows. In the following section, we introduce the key features of GTM together with the two major streams, that of Glaser, on the one hand, and Strauss and Corbin, on the other. We then characterise a number of the problems that researchers have had with using GTM and present the way in which they resolved them. Finally, we present our conclusions.
Grounded Theory Research Streams

GTM has two central features, which are common to both streams: “constant comparison,” in which data are collected and analysed simultaneously, and “theoretical sampling,” in which decisions about the data to be collected next are determined by the theory that is being generated (Glaser 1978, 1992; Strauss and Corbin 1990, 1998). A key aspect of GTM is the development of categories (of concepts) and the relationships among them. Conceptualizations of the categories, their inter-relationships, and their properties are captured and described in detail in memos. The theory emerges and evolves as data collection and analysis proceed iteratively, until theoretical saturation of categories is achieved; that is, until nothing new emerges from the data. The final theory is developed by choosing a core category and relating the other key categories to the chosen core category.

Since the introduction of GTM in 1967, Glaser and Strauss have taken quite different paths (see, for example, Glaser 1978, 1992; Strauss and Corbin 1990, 1998). The two streams share the objective of grounding theory in the data. However, while Glaser remained true to the original tenets of allowing the theory to emerge unshackled from the data, Strauss, in association with Corbin, sought to proceduralise the process of conducting research using GTM. They argued that having a clear set of detailed procedures would be useful for learning qualitative analysis. Despite the fact that LaRossa (2005) presents a very careful analysis showing that some reconciliation of the two approaches is possible, the two proponents have authored numerous, sometimes acrimonious, books and journal articles. Below is a brief introduction to each stream.

Glaser’s method (1978, 1992) focuses on the data by “allowing the data to tell their own story.” He therefore prefers to use a true “discovery mode,” one that allows researchers to rely on their creativity when developing theory. Glaser used two coding stages, substantive and theoretical coding. Substantive coding consists of two sub-stages, open and selective coding, and is concerned with producing categories and their properties. Theoretical coding occurs at the most abstract conceptual level by relating the chosen core category to the other major categories. Glaser highlighted strongly the need for conceptualization and abstraction as the foundation for deriving a “conceptual theory.” As an aid to conceptualization, Glaser offers a set of 18 “theoretical coding families.” Although one or more coding families could be used as a general framework to help researchers develop and integrate their developing theory, it is not a core part of Glaser’s method.

Strauss and Corbin (1990, 1998), on the other hand, focused on developing a more formal set of procedures and techniques for data analysis that were to be followed in step-by-step fashion. They proposed a three-stage coding process: open, axial, and selective coding. Open coding is the “analytic process through which concepts are identified and their properties and dimensions are discovered in the data” (Strauss and Corbin 1998, p. 101). Although details offered in the literature differ, open coding typically involves increasing levels of conceptualization or abstraction, the subsequent formulation of categories from the highest-level concepts identified, the identification of the properties and dimensions of those categories (Strauss and Corbin 1998, pp. 114-119). Although it forms part of axial coding, Strauss and Corbin (p. 114) note that it is helpful to develop sub-categories based on when, where, what, how, why, etc. (See Strauss and Corbin (1998, p. 114-119) and LaRossa (2005) for in-depth analyses.) Note, also, that Strauss and Corbin’s definition of sub-category differs from database terminology.

Axial coding is the “process of relating categories to sub-categories” based on their “properties and dimensions” (Strauss and Corbin 1998, p. 123). In other words, the purpose of axial coding is to “begin the purpose of re-assembling the data that were fractured during open coding” in order to develop more precise and complete explanations about the phenomena (Strauss and Corbin 1998, p. 124). Strauss and Corbin recommend using what they call the coding paradigm or paradigm model to aid in axial coding.
The paradigm model consists of six pre-determined categories that guide data collection and analysis: conditions, phenomena, context, intervening conditions, actions/strategies, and consequences (often shortened to conditions, context, actions/interactions, and consequences), and derived by asking what, where, when, how, why, and with what consequences. According to Strauss and Corbin (1998), the paradigm model is designed to develop a “well integrated set of concepts that offer a thorough theoretical explanation” of the phenomenon under investigation, often in the form of a process model that reflects cause and effect.

Theoretical coding (Glaser) and selective coding (Strauss and Corbin) is the process of integrating and refining the theory. Specifically, the researcher identifies a core category, that category that accounts for most of the variation in the central phenomenon, and then identifies relationships among the core category and the remaining categories. Theory formulation and integration is supported by a process called “theoretical sorting,” i.e., the sorting of theoretical memos in Glaser’s method (1978) and the “conditional matrix” in Strauss and Corbin’s method. Strauss and Corbin (1990) introduced the conditional matrix, as an “analytic device to stimulate analysts’ thinking about relationships,” and a way of representing the evolving theory (see Figure 1). It is a spiral representing aspects of the world ranging from broad (i.e., global) on the outer circle to focused (i.e., individual) on the inner circle. A researcher can use the conditional matrix as an aid to “trace the paths of connectivity among conditions, actions/interactions, and consequences” (Strauss and Corbin 1998, p. 182).

Figure 1. The Conditional Matrix (from Strauss and Corbin 1998)

Glaser and Strauss differ, first and most notably, on Strauss and Corbin’s axial coding. Glaser (1992) believes that Strauss and Corbin’s method of forcing the data into the paradigm model constrains the emergence of true theory. Glaser believes, instead, that categories should emerge from the data and not be constrained by any kind of model or framework. Second, Strauss and Corbin also raised the ire of Glaser by extending their method to address the elaboration of existing theory. Third, Strauss and Corbin further recommend that researchers take a more active role during the inquiry than does Glaser, recommending the use of preconceived questions and categories, and even the development of hypotheses. Hence, in Glaser’s view, the data in Strauss and Corbin’s stream no longer “speak for themselves.”

Current Grounded Theory Method Practices
We analysed a number of GT articles to identify current GT practices. We found three main problems: the very foundations of the approach, the terminology and associated meanings, and coding procedures. Below we discuss each of them.
Terminology
LaRossa (2005) states that a major source of problems in GT research is the lack of precise definitions. He also suggests that the opaque and confusing guidelines for applying GTM may lead to different interpretations of the categorization process and therefore of how those methods should be applied. According to LaRossa, GTM could be put to better use if terms and procedures were well defined and if researchers knew exactly what steps to follow, a sentiment reflected in Sarker et al. (2001) and Walker and Myrick (2006). In the final outcome, LaRossa believes that the approaches of Glaser and Strauss and Corbin, with appropriate interpretation, may not differ markedly. He concludes that Strauss and Corbin’s axial coding is similar to Glaser’s theoretical coding.

Coding Procedures
Coding procedures appear to be at the root of most of the problems researchers identify. However, because coding is the fundamental analytical process used by GT researchers (Strauss and Corbin 1990, p. 12), it is extremely important that it be done well. Despite the extensive descriptions of coding procedures both in articles and books, numerous researchers from a variety of disciplines have challenged those procedures and provided their suggested solutions (see, for example, Allan 2003; Kendall 1999; LaRossa 2005; Sarker et al. 2001; Scott 2004). In fact, such solutions appear to be consistent with Glaser’s viewpoint (1999) that some researchers tend to use GTM selectively and adapt the method for their own purposes. The majority of papers addressing issues with GTM focus on data analysis. We identified certain problems and issues across all parts of the analysis/coding process.

Developing Concept Hierarchies in Open Coding
A number of authors report experiencing problems in identifying an appropriate set of concepts for further investigation, that is, in open or substantive coding. Recall that this initial coding calls for identifying concepts at increasingly higher levels of abstraction using the constant comparison method.

First, LaRossa (2005) provides details of the use of the indicator-concept approach for identifying concepts, which subsequently evolve into categories, or what he calls “variables.” Second, Allan (2003) focuses in considerable detail on the process of developing concepts and their evolution to categories. In his critique of GTM, the author notes the difficulties associated with the micro-analysis coding (close to the word level) recommended by Strauss and Corbin, a phenomenon that Glaser (1992) refers to as over-conceptualization. Allan states that dealing with a large number of codes leads to confusion and loss of focus with respect to the main ideas. He resolved this problem by reverting to Glaser’s approach (1992), which operationalised what he calls the “Key Point Coding process.” That is, Allan based his coding on identifying key points in the text. He then analysed and grouped the key points into themes, concepts, and then categories, all based on different levels of commonality. Georgieva and Allan (2008) present further details of his approach on a more comprehensive data set.

Focusing on Conceptual Description vis-à-vis a Conceptual Theory
As we have seen, since the appearance of Strauss and Corbin’s method in 1990, Glaser (1992) has constantly criticised it for being too rigid thereby constraining the emergence of theory. He also argues that their method leads to “conceptual description” rather than a “conceptual theory.” A number of authors have since joined the debate. Kendall (1999), for example, has attempted to shed light on the issue by illustrating how her data analysis led her to conceptual description rather than theoretical development.

Kendall’s study focused on the experiences of families who have a child with attention deficit hyperactivity disorder (ADHD). Her interest was in “the process of doing well when living with a chronic and stigmatizing condition” (p.750). Although she had used Glaser’s approach in prior research, Kendall was drawn to Strauss and Corbin’s method in the belief that it would be easier to apply than Glaser’s and that funding bodies would view more favorably a research proposal based on the more rigorous methodology. Rather than focusing on “doing well,” Kendall found herself drawn into the difficulties of living with an ADHD child; in effect, she focused on the “most cogent and pervasive data set to emerge:
disruption” (p. 750). The situation was exacerbated by the fact that participants had difficulty describing what doing well meant, characterizing it most frequently in terms of what it is not. As a result, Kendall found herself side-tracked into generating a large volume of “descriptive data,” which, while it described disruption and related categories well, did not answer her research question. She claims that using Strauss and Corbin’s coding paradigm resulted in her losing focus and limited her thinking to only the most obvious aspects of the phenomenon she was investigating.

Realising that it was necessary to focus on both conceptualization and theorizing in order to move away from conceptual description, Kendall later analysed her data using Glaser’s approach. Further, she analysed the data based on her research questions, rather than on Strauss and Corbin’s coding paradigm. Kendall’s re-examination supports Glaser’s view that one needs to reach a more abstract level of analysis to generate and integrate the categories, and thus to produce a theory, and that Strauss and Corbin’s method leads to conceptual description rather than to theory. Specifically, she found that axial coding is: “a distraction that interfered with the ability to engage in a higher level of abstract thinking required in GT analysis” (1999, p. 755).

Conducting Axial and Selective Coding without Using the Paradigm Model
As we have seen, Strauss and Corbin’s axial coding consists of developing relationships among the categories derived in the open coding stage. In seeking to use this approach to build theory in the area of virtual team development, Sarker et al. (2001) faced a number of problems (see p. 45). First, they encountered difficulties in distinguishing between properties and sub-categories. Second, the fact that they identified a large number of categories and sub-categories made it “virtually impossible” for them to validate proposed relationships based on combinations of identified categories and sub-categories. Third, they experienced difficulty in clearly classifying all categories and sub-categories into the paradigm model due to unclear definitions of causal, contextual, intervening conditions terms, etc. Fourth, they felt that their data did not provide them with direct evidence of causal relationships. Fifth, they also experienced difficulties in completing the selective coding stage; that is, identifying a core category and then developing the final grounded theory.

Specifically, with regard to axial (and selective) coding, the researchers state: “it became clear, given the large numbers of categories/sub-categories that were emerging, that it would be virtually impossible to hypothesize and deductively validate relationships (among sub-categories and categories during axial coding, and between the core category and other categories during selective coding) based on all combinations of properties” (p. 45). In effect, they could not effectively handle the paradigm model. (See the following section for a suggested way of doing so.) To overcome the issue of large numbers of codes, and to fulfill the goal of axial coding, the authors used their own “two-step process” (p. 45). First, they identified the major categories and related each to their sub-categories in a hierarchical manner. Like many other researchers, they emphasize the necessity at this stage of iterating between open and axial coding. Second, the authors developed integrative memos for each category in an attempt to integrate as many sub-categories as possible with each category.

The paradigm model is also a forerunner to identifying the core category. Sarker et al. (2001) report that they found the paradigm model too mechanistic and therefore too constraining to facilitate choosing the core category. Their recommended solution was to use meta-theories. In their study of virtual team development, they state that they used both structuration theory and symbolic interactionism, to arrive at their core category of “stages of team development.” They then developed their final theory by integrating five other major categories with the core category.

Using Additional Tools to Aid in Axial and Selective Coding
Scott (2004) identifies the ambiguity and vagueness of Strauss and Corbin’s procedures for conducting GT research and the difficulty in developing and defining relationships in axial coding as major issues in
using Strauss and Corbin’s GTM. Scott, together with McCaslin and Scott (2003) and Scott and Howell (2008), suggest that axial coding requires the introduction of further rigor and guidelines in order to be applicable. They propose that the need to manage a large number of initial codes, as well as the process of developing relationships among categories can be addressed by introducing two additional tools: 1) the Conditional Relationship Guide (CRG; Scott 2002); and 2) the Reflective Coding Matrix (RCM; McCaslin 1993).

Scott conducted a grounded theory study of persons over the age of 50 who sought to do something “exceptional,” and therefore totally different, with their lives. Scott (2002) developed the CRG to help handle the large number of codes she identified in her research (reducing 1908 total codes to 54 “elemental” or primary categories). The key to this representation is that the categories not classified as elemental are properties or dimensions of the elemental categories. The CRG representation is a matrix with each of the elemental categories on the Y-axis and the six-dimensions of what, when, where, why, how, and with what consequences on the X-axis. The author states that because this tool is at a higher level of abstraction than the coded data per se, it permits the researcher to view the data as a whole. Clearly, it is a case of putting like things together, or chunking the data, a familiar approach to handling cognitive complexity.

Scott (2004; Scott and Howell 2008) then used McCaslin’s RCM (1993) as a more rigorous method of building and understanding the relationships among the categories chosen for inclusion in the final theory. This tool aids in the transition to selective coding, which, as we have seen, is the process of identifying a core category and linking the other major categories to it. The RCM, as shown in Table 2 of Scott (2004) and in Tables 2a and 2b of Scott and Howell (2008), characterizes the chosen core category in terms of properties, processes, dimensions, contexts, and modes for understanding the consequences on the Y-axis and the major categories under consideration on the X-axis. The RCM captures, therefore, the higher level of abstraction necessary to accomplish selective coding and to interpret the emergent theory. Once a core category is determined, in Scott’s case, “commitment to extraordinary achievement,” all other categories become core category descriptors: the properties, processes, dimensions, contexts, and modes for understanding the consequences.

Scott and colleagues then suggest that researchers use their RCM to develop a conditional matrix, Strauss and Corbin’s (1998) highest level of analysis. The conditional matrix “serves as a model for representing the emergent theory” (Scott and Howell 2008). Scott’s (2004) conditional matrix presents the rationale for how a person moves through the process of what she calls “the theoretical position of congruous autonomy.” In this instance, the model is positioned on the lowest level of the conditional matrix, the level of the individual (Strauss and Corbin 1998). As is typical in certain disciplines, Scott presents the theory using the “story line” derived from the conditional matrix, rather than the models preferred by the IS community.

**DISCUSSION**

Our research in progress sought to identify the current GT practices, the problems that researchers have had in applying GTM, and the solutions they used to resolve them. With our focus on actual accounts of difficulties with the GT process rather than pure discussion papers, we sought articles across a number of disciplines, in particular, health-related disciplines, in addition to IS. Unfortunately, however, there are few reflective accounts of the use of GTM in the literature and the sample of studies we analysed is necessarily small.

It appears likely that many of the issues with GTM may have arisen as a result of the emergence of two streams that essentially seek only division rather than consensus. LaRossa’s careful analysis of both perspectives (2005) suggest that, with due consideration, proponents of the two streams may not be very far apart. Although Strauss and Corbin (1998, p. 13), view GTM “analysis as “both science and art,”
Walker and Myrick (2006) conclude that “perhaps it seems simply more a science with Strauss and more an art with Glaser” (p. 558). With the definitions and interpretations of the two streams at play, however, it is difficult to see where the reasoned arguments that could provide a unifying view of the area might originate.

Many of the issues associated with coding procedures relate to Strauss and Corbin’s approach. We speculate that a proceduralized approach such as that of Strauss and Corbin is attractive to many researchers because they may believe that they simply need to follow the prescribed steps and theory will result. Glaser’s approach, on the other hand, may not be subject to the same level of scrutiny because researchers know that developing theory is their responsibility, and that no method can replace interpretation, insight, and creativity.

A number of aspects of Strauss and Corbin’s axial coding have proved to be problematic. A key issue surrounds the role of the paradigm model in aiding the development of categories, sub-categories, and the relationships among them. We identified one study that successfully used two further tools to support the paradigm model (Scott 2004) and a second that chose not to use the paradigm model, preferring instead to use their own types of sub-categories and top-down linking of categories and sub-categories (Sarker et al. 2001). At this point, it is probably fair to say that there is no consensus on how to resolve axial coding issues.

Our study represents the first step in collating and contributing to the development of more effective procedures for conducting GTM research. Given the felt need among researchers, it is vitally important that researchers report their insights into how they actually conducted their GT studies. We are currently analysing a number of additional GT studies to support our claims and provide further directions for researchers seeking to conduct GTM. Only with this type of information can we succeed in determining how to provide more meaningful analysis and ultimately procedures for researchers wishing to conduct GT research.

CONCLUSION
Ours is the first paper of which we are aware to present the current GTM practices, problems and solutions of researchers engaging in the application of GTM. We have seen that there are significant problems not only with GTM procedures and their applicability, but also with basic definitions. Unfortunately there is no consensus on how to resolve those issues, and a number of researchers have developed their own workarounds and tools or techniques that facilitate the application of GTM.

In this paper, we have taken just the first step in the quest for a set of workable, integrated procedures for conducting GT research. We have also sought to raise awareness of the fact that the problem of applying GT methods is a multi-disciplinary one, hence answers may come from unexpected areas such as health-related communities, as well as from more traditional research communities.

Although, we acknowledge the implication of the researcher’s ontological and epistemological position on how GT research is conducted and data is analysed (Partington 2004; Urquhart et al. 2009), in this preliminary analysis we did not address the link between the philosophical underpinnings and the conduct of GTM. Our initial focus was solely on how GTM procedures are operationalised.

Because journal pages are consumed largely in explaining study outcomes, there is often insufficient space to provide the chain of evidence leading to a grounded theory, i.e., an accurate account of the methodology is rarely provided (Urquhart et al. 2009). Often researchers gloss over the issues they had with using the suggested techniques, making it difficult to identify how GT studies were actually conducted. We hope that our eventual aim of raising awareness of the current GTM practices and producing a rigorous, well-defined set of procedures for conducting GT research will be realised as more researchers open the black box on how they actually pursued the process of creating grounded theory.
REFERENCES

BIS WP2011-01 Analysis of Current Grounded Theory Method Practices

