Data analysis: comparison between the different methodological perspectives of the Grounded Theory

José Luís Guedes dos Santos¹, Kamylla Cunha², Edlamar Kátia Adamy³, Marli Terezinha Stein Backes¹, Joséte Luzia Leite⁴, Francisca Georgina Macedo de Sousa⁵

ABSTRACT

Grounded Theory (GT) has undergone many transformations since its proposition as a research method in the 1960s. As a result of these changes, three different GT approaches have been outlined: classic, straussian and constructivist ground theory, each with its peculiarities, but mainly concerning data analysis. Thus, this study aimed to describe and compare the data analysis process adopted by the different methodological perspectives of GT. The text is organized into two topics. Firstly, we present the common and differentiating characteristics of the three methodological perspectives of GT. In sequence, the data analysis system adopted in each of the GT methodological perspectives is described and exemplified, varying between two and three stages, and may or may not include the use of the paradigmatic model. Thus, this study provides support for understanding the different data analysis systems adopted in GT, which may contribute to the rigor and scientific quality of nursing research adopting this method.

DESCRIPTORS

Qualitative Research; Nursing Research; Nursing Methodology Research.

¹ Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil.
² Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Enfermagem, Florianópolis, SC, Brazil.
³ Universidade do Estado de Santa Catarina, Chapecó, SC, Brazil.
⁴ Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.
⁵ Universidade Federal do Maranhão, São Luís, MA, Brazil.
INTRODUCTION

Grounded Theory, abbreviated as GT, is one of the most used methods in Nursing qualitative research. It enables generating explanations from understanding the actions of individuals and/or groups in a given context in facing problems or experienced social situations. Its use is mainly indicated when the topic of interest has not been previously studied or the scientific production on it is scarce\(^1\)\(^-\)\(^4\).

The method was developed by the sociologists Barney Glaser and Anselm Strauss in the United States as an alternative to the hypothetico-deductive tradition of qualitative research at the time. The first work on GT was published by the authors in 1967, and is entitled *The Discovery of Grounded Theory*. The book was considered innovative as it proposed the development of theories from data obtained through research, rather than the deduction of analyzable hypotheses by means of existing theories\(^5\)\(^-\)\(^8\).

After publication of the original work on GT, the idealizers of the method began to diverge on the methodological procedures of the method, adopting independent lines of work. Glaser continued to defend GT’s original approach, becoming the main exponent of the Classic or Graserian perspective of the method. Strauss, in partnership with Juliet Corbin, incorporated new instruments of analysis and stages for developing the theory, founding the Straussian or relativist perspective of GT. In the 2000s, Glaser’s former student Katy Charmaz introduced her own version of GT, starting the constructivist perspective of the method\(^9\)\(^-\)\(^8\).

Thus, it is considered that the three main methodological perspectives of GT are: classic, straussanian and constructivist. The constitution of these different approaches of GT mainly occurred as a result of evolving scientific thought and the paradigms that guide qualitative research\(^1\)\(^,\)\(^3\)\(^,\)\(^7\)\(^-\)\(^8\). One of the main differences between them is the data analysis system, which brings particularities according to each methodological perspective.

In Brazil, several studies have discussed conceptual and operational aspects of GT application in nursing research\(^1\)\(^,\)\(^3\)\(^,\)\(^4\)\(^-\)\(^9\)\(^,\)\(^1\(^2\). However, by consulting some of the main libraries and online databases such as the Scientific Electronic Library Online (SciELO), the Base de Dados em Enfermagem (BDENF), the Latin American & Caribbean Health Sciences Literature (LILACS) and the National Library of Medicine (PubMed), no Brazilian publications specifically discussing GT data analysis process have been identified.

In addition, an integrative review on the method application in nursing research in Brazil highlighted the need for discussing the methodological aspects of GT, seeking to use the method with greater scientificity and rigor\(^4\)\(^-\)\(^5\). Therefore, it is believed that this study brings relevance in order to provide subsidies for researchers interested in using the method in their research.

Thus, the objective of the study was to describe and compare data analysis systems adopted by the different methodological perspectives of GT.

The text is didactically structured into two topics. Before discussing GT’s data analysis systems, a description of the common and differentiating characteristics of the method is initially presented according to each of its methodological perspectives.

COMMON AND DIFFERENTIATING CHARACTERISTICS OF THE METHODOLOGICAL PERSPECTIVES OF GT

The three main methodological perspectives of GT have four characteristics in common: (1) theoretical sampling; (2) constant comparative analysis of data; (3) elaboration of memos; and (4) differences between substantive theory and formal theory. These aspects can be considered inherent principles to the GT method\(^3\)\(^,\)\(^8\).

Theoretical sampling is one of the GT differentials in relation to other qualitative research designs. It refers to the process of data collection with the objective of searching for places, people or events that potentiate identifying modifications between concepts, as well as the category density, their properties and dimensions according to information needs identified in developing the research\(^8\)\(^,\)\(^9\).

Achieving the theoretical sampling starts with data collection from people and/or data sources considered relevant to answer the research question and the research objectives. As the first data collected are analyzed, the next subjects or data sources can be listed according to specific need to deepen the knowledge or the gaps to be filled, where it is possible to change the characteristics of subjects, situations or events\(^8\)\(^,\)\(^9\). One of the strategies to obtain theoretical sampling is to conduct sampling composed of groups with different participants, but with relevant experiences in relation to the research phenomenon.

Thus, the sample is not *a priori* defined, but rather along the course of the study the construction of hypotheses allows for developing and deepening of concepts to fill in the gaps of the emerging theory. This is possible due to the cyclical nature of the method, since data are collected and analyzed concomitantly until reaching theoretical saturation\(^8\)\(^,\)\(^9\). Therefore, data are at the same time the products and the producers of new data through a dynamic process of deduction, induction and verification. Deduction allows the construction of hypotheses, while induction allows the apprehension of implications arising from these hypotheses to qualify or deny them. Thus, the data are constantly subjected to questioning, making the theoretical explanation increasingly denser.

Data analysis in GT is based on a process of constant data comparison, known as constant comparative analysis. First, the collected data are meticulously analyzed word-by-word, line-by-line or incident-by-incident in order to generate conceptual codes. These codes are grouped into categories, denoting higher-level concepts. Because the steps of data collection, analysis and categorization in GT are simultaneous, there are three levels of constant comparisons: codes with codes, codes with emerging categories and categories with categories\(^3\)\(^,\)\(^1\(^3\)\(^-\)\(^1\(^4\).
In the context of constant comparative analysis, the preparation of memos is another feature that prevails, regardless of the methodological perspective of GT\(^8\). As concepts begin to emerge through the process of constant analysis and comparisons, the researcher reflects on the data. Such reflections are recorded in the form of memos, which contribute to illustrate the development of ideas and codes that will aid in developing the theory.

The difference between substantive theory and formal theory was established by Glaser and Strauss in the published work that gave rise to the method. According to them, when GT is generated from a specific context, a theory only applying to the investigated field emerges, which is called substantive theory. In turn, formal theory requires in-depth study, involving the generation of abstract concepts that can be applied in a generalized way to a broader reality\(^{8,13}\). Therefore, substantive theory is the foundation for a formal theory. For example, a study that focuses on parents coping with the early diagnosis of hearing loss of their child corresponds to a substantive theory; while a research focusing on understanding parental confrontation with any trauma or general coping will require the development of a formal theory. It is important that the researcher focuses on generating one or another type of theory in their study, and that they are clear about the differences between them\(^{15}\). All three methodological perspectives of GT have the same distinction in relation to substantive theory and formal theory\(^8\).

Regarding the differentiating characteristics of the GT methodological perspectives, three aspects stand out: (1) philosophical basis; (2) use of literature; and (3) the data analysis system\(^3,8-9\).

There is a broad discussion in the scientific literature about the philosophical basis underlying each of GT’s methodological perspectives, both among authors and among scholars of the method. In his work, Glaser does not directly address the philosophical basis that supports classic GT, classifying it as a general method that can be used to collect all types of data, regardless of the theoretical reference\(^8,13\). However, positivism is considered the philosophical basis of classic GT due to Glaser’s connection to this philosophical current and of the importance he attaches to neutrality and objectivity in developing the theory\(^9\).

Strauss and Corbin make the link to post-positivism clear in their books, in which Symbolic Interactionism and Pragmatism permeate their proposed methodology. Charmaz endorsed the principles of Symbolic Interactionism and Pragmatism, but criticized Strauss’ post-positivist expression in his systematic coding procedure. Thus, she proposes the recovery of the pragmatic and symbolic interactionist emphasis on the meaning, the language, the interpretation and the interaction, linking GT to the constructivist paradigm as an interpretative methodology\(^8\).

Due to the influence of these philosophical assumptions, the indication for using the literature also differs according to the methodological aspect of the GT. Glaser’s classic GT recommends that the researcher begins data collection without any knowledge of the preexisting literature, with the aim of keeping an “open mind”, free of external influences\(^{13}\). In contrast, Strauss and Corbin suggest appropriate use of literature throughout all the research phases, pointing out the difference between an “empty head” and an “open mind”\(^6,16-17\). Charmaz endorses Strauss and Corbin’s point of view, however she recommends compiling the literature after data analysis. She believes that this strategy enables understanding the existing scientific production on the researched topic and it helps to develop the researcher’s argumentative potential without compromising their creativity\(^5\).

The data analysis system is the differentiating aspect that usually generates more questions among the researchers interested in using GT, and even among those who already use it. For this reason, it is the main focus of this study and will be presented in detail in the following section. Chart 1 presents a summary of common and differentiating characteristics of GT to conclude this section.

![Chart 1 – Common and differentiating aspects of GT – Florianópolis, Santa Catarina, Brazil, 2017.](chart1.png)

**DATA ANALYSIS SYSTEMS IN GT**

Coding is the procedure in which the data are separated and conceptualized for data analysis, seeking to define and identify the relationship between them. This is the first step in developing the theory. Due to the influence of the philosophical framework, each methodological aspect of GT supports its own system of codification and analysis, as can be observed in Chart 2. Each of them is presented below.
study developed with 51 nurses (10 in Brazil, 9 in England, 10 in Germany, 10 in Ireland and 12 in Palestine). The central category identified in the research was: “Negotiating reorientation”. This category expresses that upon the imminent death of an ICU patient, the care focus ceases to be healing and becomes a comfortable death for the patient and support for family members. Nurses have a decisive role in the negotiations between professionals and families which involves this reorientation of care, as evidenced in the subcategories of the study: “Seeking consensus” and “Providing emotional support”.

**Straussian perspective**

The Straussian tradition has Anselm Strauss and Juliet Corbin as its main representatives. They defined new stages for GT development with the objective of making the methodology more accessible and didactic. This perspective highlights the active position of the researcher in facing the data and in elaborating the theory, which can seek theoretical support before and during the data collection and analysis. In this approach, the data analysis system is divided into three stages: open coding, axial coding and selective coding.

Open coding is the first analytical step, in which the researcher must focus on the collected data, examining it, comparing it and conceptualizing it with words that convey action. This step happens by analyzing each word line-by-line, meaning that a thorough examination through exhaustive questioning by the researcher regarding the data is necessary: “What is this? What does it represent? What is happening here?” Thus, identification of substantive codes, their properties and dimensions can proceed.

The second step of the analysis is Axial coding, and it is marked by inductive-deductive motion, which demands theoretical sensitivity and reflection by the researcher, who seeks answers to questions such as: Why? In which way? Where? When? And How? At that time, the data that was separated in the open encoding is regrouped in order to form explanations about the researched phenomena and to enable category emergence.

In this step, an analytical tool called the codification paradigm or paradigmatic model is used, which assists in axial coding by systematically ordering data, integrating structure and process, and capturing the evolutionary dynamics of facts. The paradigmatic model consists of components known as the “5 Cs”: context, causal conditions, intervening conditions, strategies and consequences. These components help to establish relationships between categories.
and identify the phenomenon or the central category of the research.

With the evolution of the Straussian approach of GT, the paradigmatic model became to have three components, the “3 Cs”: conditions, actions-interactions and consequence\(^{(6)}\). However, in Brazil, the work available in Portuguese presents the model composed by the “5 Cs”, which makes this paradigm the best known by Brazilian researchers. Such a change indicates the influence of constructivism and contemporary postmodern thinking on this methodological perspective of GT. In their work, the authors themselves acknowledge this influence and express admiration for the work that has been developed by Charmaz\(^{(6)}\). Adoption of the three-component model allows greater flexibility to researchers in adopting this approach.

Chart 3 presents a synthesis of the axial coding models for the Straussian perspective.

<table>
<thead>
<tr>
<th>Model</th>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strauss and Corbin (2008)</td>
<td>Context</td>
<td>Where the phenomenon happens and conditions that enable strategy development.</td>
</tr>
<tr>
<td></td>
<td>Causal conditions</td>
<td>Set of events that trigger or influence the phenomenon development.</td>
</tr>
<tr>
<td></td>
<td>Intervening conditions</td>
<td>Aspects that interfere with or alter the impact and/or the phenomenon development.</td>
</tr>
<tr>
<td></td>
<td>Strategies</td>
<td>Actions and interactions designed and developed to deal with the phenomenon.</td>
</tr>
<tr>
<td></td>
<td>Consequences</td>
<td>Current or potential results of the strategies identified in the study.</td>
</tr>
<tr>
<td>Corbin and Strauss (2015)</td>
<td>Conditions</td>
<td>Reasons given by the informants as to the occurrence of a specific fact, as well as explanations for the reasons why they respond to an action in a given way.</td>
</tr>
<tr>
<td></td>
<td>Actions-interactions</td>
<td>Answer expressed by participants to events or problems.</td>
</tr>
<tr>
<td></td>
<td>Consequences/Results</td>
<td>Refer to the expected or actual results from actions and interactions.</td>
</tr>
</tbody>
</table>

In the selective coding, the last moment of the analytical process corresponds to a refinement of the previously found categories and subcategories, continuously compared and analyzed by integrating the data and thus enabling the researcher to identify the central category or phenomenon\(^{(6)}\). In the most recent Straussian GT work, the selective coding was denominated integration, a term considered more appropriate to the process performed in this step. The main category is a broad and abstract concept that in a nutshell describes what the researcher considers as the main theme of the study\(^{(6)}\). At the end of the coding stages, the generated theory is organized according to the elements of the paradigmatic model.

The Straussian approach also suggests the use of the conditional or consequential matrix as an analytical tool, which is represented by a set of eight levels of integrated circles that are inserted into one another. The more external the circle, the broader the context to which it refers, being: international, national, communitarian, organizational and institutional, suborganizational and subinstitutional, collective, interactional and action. This device assists in identifying relationships and connections between conditions/consequences and actions involving events or incidents throughout the theory development\(^{(6,8,16)}\).

To illustrate the use of the Straussian perspective, a research developed with the objective of “to understand the nurse leadership exercised in the services of Primary Health Care” can be mentioned, in which 30 nurses were divided into four sample groups and were interviewed. Six nurses from the Family Health Strategy, eight nurse coordinators of the local and the regional Health Units, 10 nurses with administrative positions at a managerial level and seven nurse university professors with expertise in the subject. The phenomenon obtained was: “Revealing the nursing leadership practices in the complex context of Primary Health Care,” being composed of nine categories according to paradigmatic model: “The public management of Primary Health Care is considered complex by the nurses” (context); “Highlighting the multiple features of the leader nurse” (causal condition); “Perceiving complementarity between confidence and experience in order to generate links”, “Contrasting the dependence of nurses to the system and their autonomy in relationships, interactions and associations in Primary Health Care” and “Limitations of working conditions and training of health professionals influencing the different demands of nurses” (intervening conditions); “Improving the practice of leadership through planning and support from managers”, “Using communication as a tool that gives dynamism to the interrelations between nurses, staff and community” and “Enhancing skills, and individual and team potential” (strategies); and “Nurses sharing the leadership as opposed to individualism” (consequences)\(^{(20)}\).

**Constructivist perspective**

The constructivist approach considers that the theory is a reciprocal construction between the researcher and the research subjects. In this sense, the research focuses on the meanings attributed by the participants to the research phenomenon, which are contextual, shaped by social interactions and that change over time\(^{(23,5,21–22)}\).
Two main coding steps are recommended: initial coding and focused coding. In initial coding, the data is fragmented and analyzed with the objective of conceptualizing ideas or meanings expressed by the participants, thus transforming them into codes. This process can be carried out word-by-word, line-by-line or incident-by-incident. We emphasize the importance of microanalysis using *in vivo* concepts; analysis of a word and incidents in order to discover relevant dimensions of the categories and the relationship between these and the subcategories in a causal relationship.

To do so, researchers can use sensitive questions (who are the actors involved in the phenomenon); guiding questions (to guide interviews; and change constantly); theoretical questions (which help to see the process, variations and connections between concepts); and structural questions (which help to develop the theory structure). In other words, more explanations are needed to determine the concepts’ properties and dimensions. The researcher strategically uses the comparison from the incidents to move from a description level to an abstraction level.

The codes generated in the initial coding are called provisional, which makes it possible to keep the researcher open to other analytical possibilities, progressively replacing those codes by new ones that better suit the data from the point of view of understanding the meanings and experiences of the research participants. The search for an analytical direction of constructivist GT at a higher level allows for focusing on certain ideas first and conclude, and then returning to unfinished data and analysis in other areas. The coding of each line of data allows insights into what type of data to collect next, refining these data and further directing the research. At this point, it is important for the researcher to be aware whether their personal tendencies, assumptions or beliefs, or those of the informants are interfering with the analysis.

Focused coding is the second coding step in the constructivist perspective, and allows for separating, classifying and synthesizing large amounts of data. At this stage, the elaborated codes are more targeted, selective and conceptual, as they should synthesize and explain larger data segments. Thus, the more significant and/or frequent codes identified in the previous step can also be used based on the definition by the researcher for which initial codes allow better analytical understanding of the data. As certain concepts emerge more frequently and prominently, subcategories and categories are generated, which in turn reveal the phenomenon or central category of the research. The central category represents the most potent central analytical organizing concept. The identification of the central category depends on the researcher’s perception and it represents the most relevant process in the investigated area.

In the coding process of the constructivist approach, Chamaz emphasizes the importance of conjugating the verbs in the gerund (×ing) form to represent the actions that are being codified, for example: searching, realizing, feeling, etc. The purpose of using gerund form is to assist in developing the theoretical sensitivity of the researcher, allowing identification of the concepts and processes being developed.

To exemplify the use of constructivist GT, we can highlight a study carried out with 38 nursing students from Canada with the objective of understanding the resilience process development throughout their academic training. The central category identified was “Struggling,” which emerged from the expression (“pushing through”) used by the participants to describe their effort to overcome difficulties and to not let them (the difficulties) stop them. The students’ resilience process is described in progressive stages that express the trajectory of the participants seeking academic and professional development.

### FINAL CONSIDERATIONS

In this study, we have attempted to describe and compare the coding steps recommended for data analysis in the different methodological perspectives of GT. Based on what has been presented, some readers may be wondering: “How to define which methodological GT perspective to adopt in a study?” and/or “What is the best way to analyze the data for GT development?”.

There is no single answer to these questions, and discussions on these topics will not end with this work. However, it can be pointed out that the definition of the methodological GT perspective should be based on the type of research phenomenon to be investigated, and the approach through which the researcher would like to study it. Among the three methodological GT aspects presented herein, the best one is the one that best suits the researcher’s world view, the theoretical reference of the research and the issue under investigation.

The Straussian perspective can be considered a better option for researchers who are new to the method, since it presents a more systematic data analysis system in relation to the other currents of GT. In adopting the classic and constructivist perspectives, more time may be necessary for developing the research, taking into account the theoretical abstraction necessary for data interpretation and theory elaboration without adopting a paradigmatic model direction. In this sense, the time allocated for collection and analysis is a factor to be mainly considered by students and undergraduate and graduation supervisors.

It is fundamental when using GT that the researcher understands the guiding characteristics of the method, which are common among its three methodological approaches, and above all that they understand the differentiating principles. We hope that this study has contributed to this, and has allowed for understanding the main differences between the data analysis systems adopted in the methodological perspectives of GT.

As a limitation, this theoretical essay had the analysis of the three most widespread methodological perspectives of GT in nursing research as its scope. For this reason, two new GT variations were not included: Postmodern Situational Analysis by Adele Clarke, and Dimensional Analysis by Leonard Schatzman. Future studies discussing the specificities and potentialities of these new approaches in the method could further improve the possibilities of applying GT in nursing research.
RESUMO
A Teoria Fundamentada nos Dados (TFD) tem passado por muitas transformações desde a sua proposição como método de pesquisa, nos anos de 1960. Como resultado dessas mudanças, delinearam-se três diferentes vertentes metodológicas da TFD: clássica, straussiana e constructivista, cada uma com suas particularidades, principalmente no que tange à análise de dados. Assim, este estudo teve como objetivo descrever e comparar o processo de análise de dados adotado pelas diferentes perspectivas metodológicas da TFD. O texto está organizado em dois tópicos. Primeiramente, apresentam-se as características comuns e diferenciadoras das três perspectivas metodológicas da TFD. Na sequência, descreve-se e exemplifica-se o sistema de análise de dados adotado em cada uma das perspectivas metodológicas da TFD, o qual varia de duas a três etapas, podendo incluir ou não o uso do modelo paradigmático. Dessa forma, este estudo fornece subsídios para a compreensão dos diferentes sistemas de análise de dados adotados na TFD, o que pode contribuir para o rigor e a qualidade científica das pesquisas em enfermagem que adotarem esse método.

DESCRIPTORES
Pesquisa Qualitativa; Pesquisa em Enfermagem; Pesquisa Metodológica em Enfermagem.

REFERENCES
Data analysis: comparison between the different methodological perspectives of the Grounded Theory


This is an open-access article distributed under the terms of the Creative Commons Attribution License.