Qualitative Data Analysis

By Patricia Sachs Chess and Brian Adams-Thies

Essential Questions

1. What is qualitative data analysis?
2. What steps does a researcher take to prepare for qualitative data analysis?
3. What are the main steps in thematic analysis, and what processes are used in each step?
4. What are the pros and cons of using a qualitative software program?
5. What are common pitfalls of qualitative research?

Introduction

The purpose of this chapter is to provide an overview of general qualitative data analysis from both foundational and practical perspectives. Qualitative data analysis is a systematic endeavor that involves a set of practices or techniques used in several disciplines. However, the form of data analysis doctoral learners select for their studies must be aligned to the research design. For example, a phenomenological study will use a data analysis process (phenomenological reduction) which is very different from the analysis for a descriptive design; a **grounded theory** design will use a data analysis process (similar to inductive thematic analysis but with subsequent steps to produce the theory), which is very different from that of a case study design; etc. This chapter focuses on the simplest and most effective means of analyzing qualitative data: thematic analysis. This form of data analysis can be used in the following designs with some adaptation for each design: **qualitative descriptive** design, narrative design, and case study design. The qualitative design does not require adaptation, the narrative design should focus on stories and their contexts in the narrative, and the case study should pull together analysis to describe the case overall. The other qualitative designs require specific forms of data analysis, which will be addressed in later courses.

What is Qualitative Data Analysis?

Qualitative data analysis is a systematic approach to understand commonalities across a dataset. Qualitative datasets are complex, including many, many pages of transcripts, and depending upon the research purpose, might also include photographs, drawings, audio and video recordings, social media, archival records, documents-in-use, observational records (field notes), demographic information, and questionnaire responses. These may exist as paper documents or digital documents. The vast majority of qualitative dissertations at GCU rely upon recorded and transcribed interviews, recorded and transcribed focus groups, and responses to open-ended questionnaires. Even though these data sources might appear limited, many times learners will have more than 100 pages of transcribed data (the more data the better!). This chapter will help learners to make sense of a process whereby those 100 pages can be distilled into themes, which directly answer the research questions in the study.

In essence, qualitative data analysis is about discovering shared meanings, as demonstrated in the data sources, held by a specific group of participants (the sample) in the study. Keep in mind, the sample is comprised of a clearly defined group of people with shared experiences of a specific phenomenon, in a specific context, at a specific time; therefore, in order to produce good qualitative research, the researcher needs to provide as much detail as possible of the group of people (the sample), the location of the study, the historical context of the phenomenon, and the data collection and data analysis process. The more the researcher describes, the more trustworthy the study. The more data one collects and clearly analyzes, the more trustworthy the study.

Fortunately, thematic analysis is an extremely versatile approach that researchers apply to address and interpret all the data collected through interviews, focus groups, questionnaires, and any other data sources. Thematic analysis involves reviewing data to identify and discover patterns of thinking, action, performance, meaning, insight, and experience. If a researcher masters thematic analysis, he or she will have a strong foundation in many other qualitative analysis processes. Thematic analysis offers a strong foundation in the skills necessary for taking a **corpus** of data and both **reducing** reflecting up and interpreting those patterns within the data to produce themes (shared meanings across the data including pertinent context) (Braun & Clarke, 2006). What follows are emergent patterns within the data presentation and discussion of preparing for and conducting thematic analysis.

Preparing for Data Analysis

Assuming a Scientist's Attitude

Data analysis is an iterative process that involves looking at data from the perspective of a scientist. Assume a scientist's attitude. What does it mean to look at data from the perspective of a scientist? A researcher must try to control preconceived notions of what the data might say. For example, one might be doing research in the very same location in which one works. Though this research context might allow a researcher to have intimate knowledge of the context of the research, the position held by the researcher in relation to the location and the research might compromise the data analysis. The researcher might have a preconceived idea of what the data might say; however, in all research, one must put aside these preconceived notions and let the data speak for itself.

There are many mental processes that might be used to maintain a scientific perspective. One of the most popular, in recent years, is called **bracketing**. Bracketing is an idea that was originally developed in philosophy by Husserl (1913/1982). Essentially, bracketing is a process of reflection that seeks to bring to one's conscious mind those preconceived notions one might have about the data, then set those aside while engaging in the data analysis process.

Put another way, the purpose of bracketing is to help a researcher set aside his or her preconceived understandings of the phenomenon under study and to let the data speak for themselves. Bracketing is achieved by a constant process of reflection and writing that encourages researchers to think critically about whether they are letting their own experiences and preconceptions inform the data. It is important for qualitative researchers to be aware of their own perceptions, beliefs, and positions during the analytic process. The researcher's awareness of his or her own preconceptions during analysis allows the researcher to "bracket" these preconceptions so that the analysis can be systematic, scientific, and without interference.

Other means researchers might use to maintain a scientific perspective are reflective journaling or reviewing the data analysis with colleagues and mentors. Every researcher needs to find a way of maintaining a scientific perspective when approaching data analysis. A researcher must let the data speak rather than forcing the data into his or her preconceptions.

Regardless of the means one might choose to maintain a scientific perspective, the important point here is to let the data speak for itself. Researchers should not assign the meanings they want to the data; they want the data to speak to them. For example, imagine a researcher is doing a study on male children from divorced

homes who are raised only by their mothers. Perhaps the researcher comes from a divorced home, and the experience without a father in the household was less than ideal. The researcher will want to reflect on the personal experience and be sure that experience is not being projected onto the data. If this is not done, the researcher might analyze the data and reach a preconceived conclusion: male children from divorced homes who were raised by just their mothers have a hard time. However, if the researcher brackets preconceived notions, the researcher might discover, through a more scientific data analysis, male children from divorced homes who were raised by just their mothers have better empathy or are more responsible. A researcher must remove preconceptions, agendas, and expectations to maintain a scientific perspective.

Transcribe All Recordings

The information from the sources of data needs to be put into a format that can be analyzed. Interviews and focus groups, two of the most common sources of data for qualitative studies, need to be recorded. These recordings need to be audio but can also be video, depending on how the researcher is conducting the interviews. All recordings will need to be uploaded for the dissertation committee to review, so be sure to record; be sure to have a back-up recording device and be sure recordings can be digitized.

All recordings must be transcribed. For example, researchers cannot transcribe just a few of the interviews they conducted and claim to have found all the themes. Instead, every interview needs to be transcribed and analyzed. It is important to transcribe both the researcher's questions as well as the participants' responses because the researcher may pose questions slightly differently in each interview and follow-up probes may vary slightly as well.

The rephrasing of questions and the variation in follow-up probes is acceptable practice in **qualitative research** because the goal is to ensure the participant understands the question and the researcher has collected quality data that explores the phenomenon holistically. Because participants may respond to questions using the words or terms provided by the researcher, it is important to know what the researcher actually said in order to determine whether there is an emergent pattern. For example, if the researcher says, “Share with me what you think servant leadership might mean in the context of the elementary school classroom,” the participants may use the term *servant leadership* in their responses. Though this may be a code in the first round of **coding**, the researcher will want to think critically as to whether or not this is a code that appears because of the question/prompt or if this is a code that arose from the sample spontaneously. As with all qualitative research, context is everything. In this case, the context of the question the researcher might ask and how the participant responds to the question is extremely important.

The transcriptions of the data should also be verified. Read the transcriptions along with listening to the recordings. Are the transcriptions correct? If one is doing member-checking of the transcripts, be sure those transcripts are verified before being sent to participants. Also, be sure to be clear about what the participants need to do. Do they need to verify this is what they said? Or do they need to expand upon, rephrase, or otherwise change what they said? All of this needs to be completed before data analysis begins, including identifying the sources of data and organizing them.

Organize Each Piece of Data

In the proposal, the doctoral learner will identify the data sources and discuss how these sources were developed. Throughout the data collection process, the researcher will be organized and systematic in where the transcripts, questionnaires, and other sources of data are placed. Though there might be some variation in how a researcher prepares the data for analysis, the simplest method is to organize by type of data. For example, if one used a descriptive design with semistructured interviews and two focus groups and had all participants fill out a questionnaire, then there are three data sources/three types of data. The researcher will want to put all interview data together, all focus group data together, and all questionnaire data together. Put all interviews in a file; put all focus groups in a file; put all questionnaires in a file; etc. Double check that all data sources for each participant are included. This phase of organization, which should be an extension of the organizational process put into place during data collection, makes it easier to code the data so that it can be

compared across type and participant. When the researcher analyzes the data across participants, or data sources, or types, it will be easier to begin to see a pattern of codes by age, gender, role, or other attribute relevant to the study.

Qualitative datasets are often voluminous and, although data may be digital, physical, or both, the vast majority of data will consist of transcriptions created in a word-processing program. Some researchers find it useful to print out the transcripts and code the actual physical transcriptions on paper. Other researchers find it useful to just work with the transcriptions on the computer. Either way is acceptable; however, one must remember to document all aspects of data analysis.

These suggestions provide guidance on how to organize data; however, most qualitative researchers will develop their own specific means of preparing the data and engaging in the analysis. These specific means develop with experience and over time. Again, simple is better at the dissertation phase.

Analyzing the Data

Data analysis in qualitative research is both an art and a science. The science of data analysis occurs by both approaching the process in a systematic manner. The entire process of data analysis needs to be described by the researcher within the manuscript. The process needs to be transparent in order to support the trustworthiness of the study. On the other hand, the art of data analysis means approaching the process in a reflective manner that is connected to the meaning and nature of the data. A qualitative researcher must reflect upon the data analysis process in a continual and sustained manner, be able to communicate the logic of the data analysis to others, and be flexible enough to approach the entire process as an iterative process.

Data analysis is not, and should not be, mechanistic. The process cannot be reduced to word frequency counts, word clouds, or many of the other manners in which various qualitative data analysis programs might organize the data. Those data analysis programs are only tools for researchers to reflect upon their own data and organize/envision the data in various manners. Data analysis must be a complex, iterative, and logical process supporting the texture and subtlety of language, action, and meaning shared by and elicited from the participants. A researcher must think of data analysis as a process that is humanistic, highly contextual, and should be conducted, as Glaser and Strauss (1967) noted, with intelligence and ingenuity.

Getting Started in Data Analysis

There are numerous ways to start the process of data analysis. For example, if a researcher wanted to use printed transcripts, he or she might lay out double-spaced printed interview transcripts on the left half or left two-thirds of the page, keeping a wide right-hand margin for writing codes and notes. Or one might not choose to print the transcripts and instead use bubble comments within a word processing program to indicate the codes. Generally, one has the raw data (the transcripts) on one side of a page and the codes and notes on the side of the page. Keep the codes aligned with the transcription so that they are easily located.

If one is feeling overwhelmed by the amount of data, it is sometimes helpful to separate a transcription into paragraphs or question and answer couplets. This way, the question and answer provided by the participant are visible. Go question and answer couplet by question and answer couplet. This way, one is not faced with all of the data at once and instead can take time, think critically and reflectively on each couplet, and take the data one couplet at a

time. Each researcher finds his or her own means of breaking up the data analysis process so as not to feel overwhelmed. However, the process needs to be logical and explainable and needs to be communicated within the dissertation manuscript.

Code the Data: Codes to Categories to Themes

Once the data have been organized, the analytic process begins with coding the data, which is a way of identifying meaning units within the data collected. The process of coding breaks the larger dataset into smaller units (meaning units). Good qualitative data are rich, meaning such data includes long and reflective thoughts described by participants in their own words from interviews, focus groups, and such. Because of the richness of qualitative datasets, one might think making sense of complex datasets is a daunting task; however, data analysis is the fun part. Data analysis is the culmination of all of the efforts of a researcher and allows the researcher to answer the research questions clearly and concisely.

Thematic Analysis

For the purposes of this chapter, the focus is on general thematic analysis, which is truly the basis of most all other qualitative data analysis processes. Once one masters thematic analysis, other data analysis processes make more sense. The purpose of thematic analysis is to identify themes that will answer the research questions. Themes are the result of synthesizing the outcomes of coding and the clustering of codes into categories. Saldaña (2016) noted, "Like coding, thematic analysis or the search for themes in the data is a strategic choice as part of the research design that includes the primary questions, goals, conceptual framework, and literature review" (p. 200).

Thematic analysis is useful for all qualitative designs.

While there are many different ways that thematic analysis can be accomplished, the goal of this chapter is to provide the novice researcher with a general overview of three main phases in the process: **first-cycle coding** (codes), **second-cycle coding** (categories), and development of themes. The first phase involves initial, open, or first-cycle coding (finding the smallest units of meaning within a body of data), followed by second-cycle coding (moving from codes to categories), and then, the development of themes (moving from categories to themes). The process of moving from codes to categories to themes allows for a systematic approach that is relatively easy to follow and will result, if done from both an art and science perspective that is transparent and logical, in clear answers to the research questions.

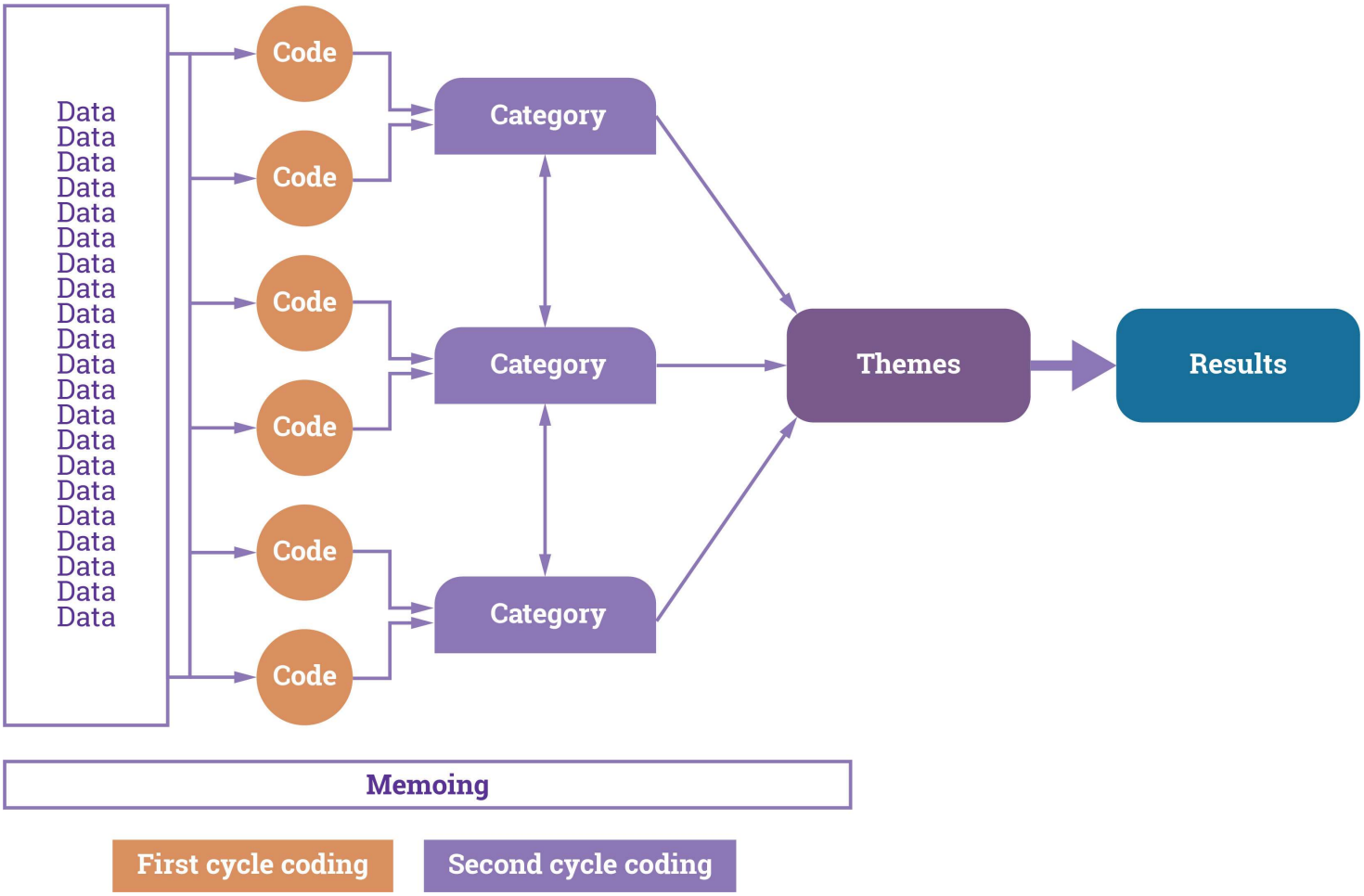
There is not total agreement in the scholarly literature about the use of the terms **code**, **category**, and **theme**. The reason for this is that qualitative research is conducted in several disciplines and for a number of different research designs. Particular disciplines tend to use specific terms. Some disciplines and approaches use the terms code and category interchangeably, while others use category and theme interchangeably. What is important to keep in mind is that, while differences do exist and are important to understand in relation to one's specific discipline, there are also commonalities across disciplinary boundaries. Qualitative researchers need to be proficient in the literature from their own discipline as to how data analysis is talked about, presented, and which terms are used in that discipline.

Figure 7.1 provides a simplified illustration of the data analysis process. The actual process of reading, coding, recoding, categorizing, and theming can be much longer and repetitive than what is pictured. For example, one might reach a category and realize the category does not really align to a group of codes, which may require a return to coding. The process of qualitative data analysis is iterative, which requires the researcher to go back and forth in this process in order to fully excavate the data for a proper analysis. Nonetheless, it is useful to visualize the process in order to understand the relationship of the steps to one another and the ultimate goal

of the analysis. The doctoral learner must also be ready to fully, transparently, and completely describe the data analysis process within the dissertation manuscript. This description of how the data were analyzed allows the researcher to claim the work as scientifically rigorous and systematic. So, as one is conducting data analysis, he or she must be sure to record the data analysis process in detail.

Figure 7.1

Overview of the Process of Data Analysis



Note. Adapted from *The Coding Manual for Qualitative Researchers*, by J. Saldaña, 2016, p. 14. Copyright 2016 by SAGE.

First-Cycle Coding and Identifying Initial Codes

The first phase of thematic analysis involves initial, open, or first-cycle coding (finding the smallest units of meaning within a body of data). The segments of data that are coded reflect what is in the data. For example, if the data are observations of some event or activity, the codes might focus on kinds of social interactions. If the data are verbal reflections by the participant, such as responses to interview questions, then codes might focus on patterns of ideas, stories, phrases, or terms. There are different methods one can use to code the data during the first cycle, or initial coding process. Two common processes are described in this section.

Inductive or Bottom-Up Coding

Inductive or bottom-up coding is also described as inductive thematic analysis. One can think of this as an emergent or unstructured approach to coding in which the structure emerges from patterns identified in the data. A researcher might use inductive thematic analysis to understand commonalities across a dataset, such as commonalities in how research participants describe an experience, or commonalities in how participants think about a topic. Inductive thematic analysis is very useful for finding these commonalities across a large corpus of data. Think of inductive thematic analysis as a process of letting meanings "bubble-up" from the corpus of data. The researcher goes into the process of inductive thematic analysis with no expectations of

what the data might contain. The researcher might use inductive analysis for narrative inquiry identifying stories, stories-within-stories, and versions of stories. A researcher might identify patterns or commonalities across the dataset and identify the participants in terms that are pertinent to the case, such as the roles individuals play within the defined case and what they say. This contextualizes a pattern or story by clearly identifying the speaker or the participants.

Imagine a researcher is attempting to understand how young adults (ages 18–25) who were bullied while attending public high schools in Denver, Colorado describe their high school experiences. In this case, the emphasis is on understanding the commonalities within the experiences of the participants. The researcher wants the themes to come directly from the experience of the participants; therefore, the researcher will use inductive thematic analysis to understand these common themes across the corpus of data. The researcher moves from codes to categories to themes, letting the data lead the analysis.

Deductive or Top-Down Coding

Top-down coding is known as deductive thematic analysis, **typological**, theoretical, or preset coding. Researchers must develop a list of codes prior to collecting data based on the theoretical framework of the study or research goals. These are called **a priori codes** and must be included in a codebook. Researchers use deductive thematic analysis when they want to find elements of a theoretical framework within the corpus, similarities between a previous study and their study, or key concepts from the literature in their own corpus of data. Think of deductive thematic analysis as a process of imposing meaning upon the corpus of data. When using deductive thematic analysis, the researcher goes into the process of analysis looking for specific concepts.

Imagine a researcher wants to understand how aspects of post-traumatic stress disorder (PTSD) might be found in the descriptions of young adults (18–25) who were bulled during high school in Denver, Colorado. The researcher will review the literature on post-traumatic stress disorder and develop a priori codes from the literature. For example, perhaps the following codes arise from the literature (again, this is just an example): depression, lack of sleep, lack of interest, social isolation, and acting out in contexts outside of school. The researcher will then look for these codes within the corpus of data he or she collected in order to demonstrate how PTSD might be found in the experiences of victims of bullying during high school in Denver, Colorado. In this case, the researcher moves from codes to categories to themes, letting the literature (or theoretical perspective) dictate the analysis.

Other Examples of First-Cycle Coding

Some other examples of first-cycle coding are:

- **Open coding** – uses a coding process whereby the researcher denotes units of meaning within the corpus. The purpose is to break a corpus of data into the smallest and most numerous units of meaning, which will then be coalesced into categories and themes.
- Attribute coding – focuses on key attributes such as demographic data.
- In vivo coding – uses the participants' own words to describe a code, rather than using terms or ideas generated by the researcher.
- Process coding – identifies actions in textual transcripts by noting the use of gerunds, which are words ending in "ing" that function as nouns (e.g., "reading is important in doctoral work").
- Affective coding – codes that focus on forms of affect, broadly defined, such as emotions, values, connections, or conflicts.

Types of Coding

- Affective Methods

- Attribute Coding
- Axial Coding
- Causation Coding
- Code
- Concept Coding
- Descriptive Coding
- Domain and Taxonomic Coding
- Dramaturgical Coding
- Eclectic Coding
- Elaborative Coding
- Elemental Methods
- Emotion Coding
- Evaluation Coding
- Exploratory Methods
- Focused Coding
- Grammatical Methods
- Holistic Coding
- Hypothesis Coding
- In Vivo Coding
- Initial Coding
- Literary and Language Methods
- Longitudinal Coding
- Magnitude Coding
- Motif Coding
- Narrative Coding
- Outline of Cultural Materials (OCM) Coding
- Pattern Coding

- Procedural Methods
- Process Coding
- Protocol Coding
- Provisional Coding
- Simultaneous Coding
- Structural Coding
- Subcoding
- Theme/Themeing the Data
- Theoretical Coding
- Values Coding
- Verbal Exchange Coding
- Versus Coding (Saldaña, 2016, Appendix A)

The Code Book

The code book is an organized table that displays the initial codes, or first-cycle codes, in use, their definitions, and examples from the transcript. In inductive thematic analysis, the code book is a way to organize the codes that emerge during the coding process, so the researcher can keep track of the codes he or she has identified as those that continue to emerge. In deductive thematic analysis, the code book is developed before the data analysis process begins, and the code book is usually included at the proposal stage of the dissertation/research process. Because coding is an iterative process, the codes may change or become more developed as the researcher begins to compare and contrast codes. Table 7.1 shows a sample code book.

Table 7.1

Sample Code Book

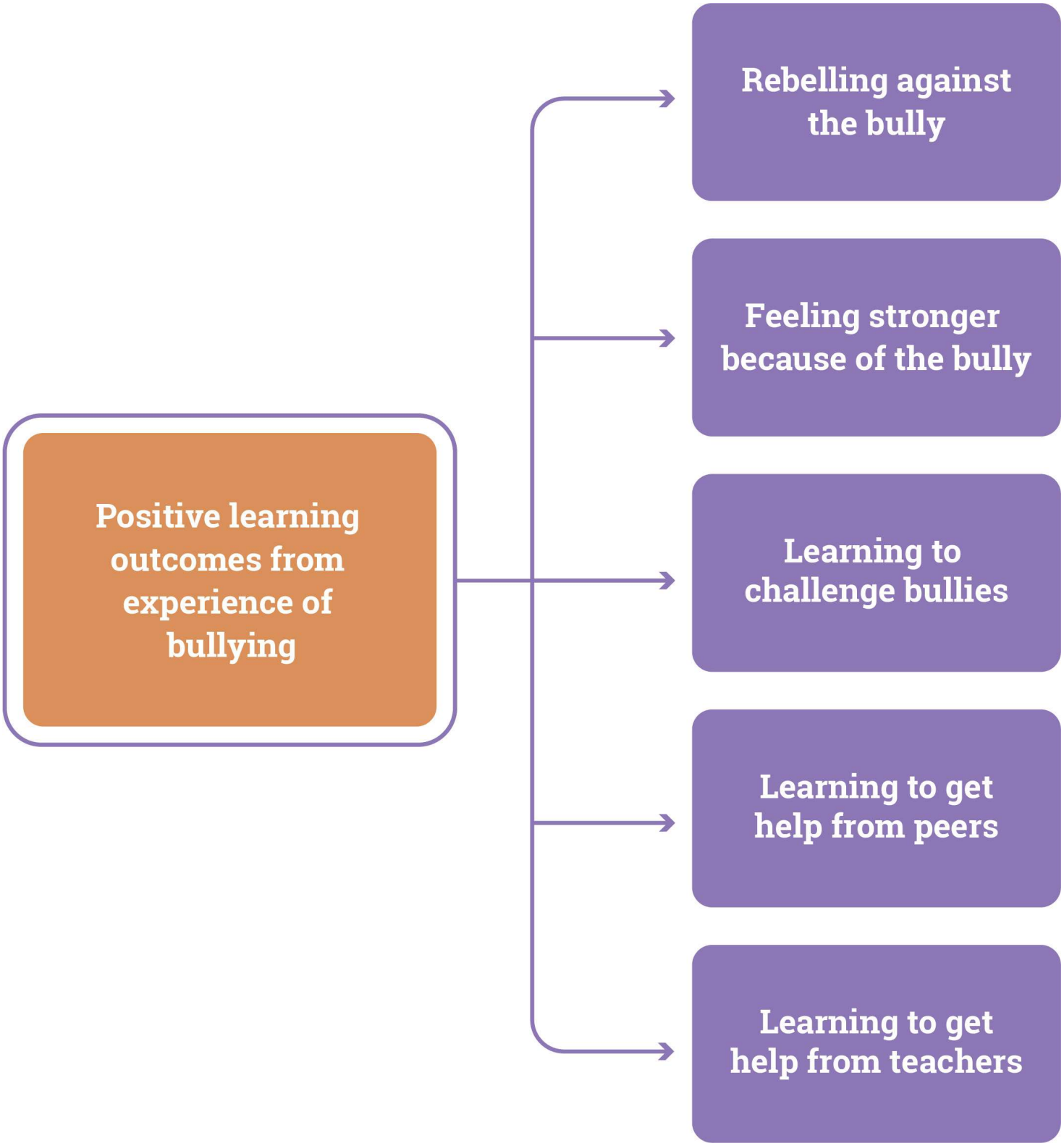
Code	Definition of the Code	Example of Code in Transcript
Learning to Challenge Bully	The process of learning how to challenge the bully; speech or actions that exemplify challenging a bully	Jane Doe: “Once I learned to fight back, not physically but just find ways to talk to them, or, um, put them in their place through words, I really started to get out from under their tyranny.”
Self-Perception	Reflections by the participant regarding the experience of learning to challenge a bully	John Doe: “I felt really bad about myself…like maybe they [the bullies] knew something about me that I didn’t know. I doubted who I was for a long time because I thought maybe they saw something about me I didn’t see myself.”

Second Cycle Coding and the Development of Categories

Once first-cycle coding has been completed, many researchers notice that some segments of data have more than one code or that some codes are similar and may overlap; they find that it makes sense to take another step and clump and/or split those codes and embark on **second-cycle coding**. In short, at this stage of analysis, the research may begin to identify how the data can be further developed into a bigger picture. As Saldaña (2016) noted, “The primary goal during second cycle coding is to develop a sense of categorical, thematic, conceptual, and/or theoretical organization from your array of first cycle codes” (p. 234).

Figure 7.2

Example of Second-Cycle Coding



Second-cycle coding is a process that helps the researcher coalesce codes into categories. Much like first-cycle coding, second-cycle coding, or coalescing the codes into categories, can draw upon different processes of coding. Second-cycle coding tends to produce a shorter list than does first-level coding in which the researcher is identifying the smallest units of meaning within a corpus of data. Some examples of how to move from codes to categories (i.e., second-cycle coding) might include those discussed below.

Pattern Coding

Pattern coding pulls together codes that cluster based on a repeating pattern in the data. For example, a researcher may notice that several participants describe an activity or incident in a similar way, and that description emerged naturally (i.e., it did not mimic terminology or ideas provided by the researcher in an

interview question or comment). Thus, a pattern might be a kind of story that emerges. Researchers "pull together a lot of material from first cycle coding into more meaningful and parsimonious units of analysis" (Saldaña, 2016, p. 236).

Axial Coding

Axial coding refers to one way of coalescing codes into categories by looking only at the codes that were developed in the first round of coding. Put another way, this form of coding identifies relationships between the codes developed in the first cycle of coding. The development of codes and the clustering into categories, and later into themes, is in service of answering the research questions. Remember: codes categories themes. In this case, the researcher looks only at the codes from the first round (not back at the transcripts or corpus) and develops categories by coalescing only the codes. While axial coding is an approach discussed by Strauss and Corbin (1998) in relation to grounded theory; it can be used in other research designs as well. In fact, axial coding is one of the simplest means of moving from codes to categories. If codes have been marked clearly in the first round of coding, then coalescing them into categories using axial coding is highly effective.

Axial coding moves the data analysis from codes to categories. For example, imagine a researcher is engaging in inductive thematic analysis of the data collected during the above-mentioned study on high school bullying. The researcher has completed a first round of coding using open coding and developed a very large list of codes. Five codes stand out from the corpus of data as the researcher starts to engage in axial coding: rebelling against the bully, feeling stronger because of the bully, learning to challenge bullies, learning to get help from peers, and learning to get help from teachers. The researcher coalesces these codes into a category: positive individual learning outcomes from experience of bullying. Of course, the researcher must demonstrate logic in data analysis as to how he or she coalesced the five codes into the category. The researcher will want to revisit the category over and over again during the data analysis process asking whether the category makes sense, aligns to the codes, or needs to be rephrased to encapsulate all of the ideas present in the codes.

Theoretical, Selective, or Conceptual Coding

Theoretical, selective, or conceptual coding is a form of coding that utilized insights from the theoretical framework and/or the literature in order to move codes into categories. This level of coding is related to a set of dominant or core concepts that the researcher wants to verify or explore within a corpus of data.

Second-cycle coding helps the researcher coalesce codes into categories. Second-cycle coding is a means of preliminary organization of the findings or results that emerge from the analysis. Of course, as in all scientific research, one must keep track of the logic used to coalesce codes into categories. The researcher may find **memoing** to be a useful strategy in recording logic and subsequently reflecting upon that logic of coalescence. Memoing is discussed in more detail in the next section.

Developing Themes

The goal in all research is to answer the research questions using a transparent, logical, and robust data analysis. Social scientists use a variety of terms to describe elements or units in the data in relation to the analytic framework, what Ryan and Bernard (2003) referred to as the relationship of themes and their expressions in the data. Doctoral researchers should be mindful of the particular analytic approach they use to guide the terminology and language they use in their studies, and they should provide citations to key authors in the relevant scholarly literature. Whether the specific terms used are themes, labels, categories, or concepts, these analytic units are a way to organize and present the results of the analysis in a coherent and comprehensible manner. However, for the purposes of this chapter, the term *themes* will be used. Remember, codes are coalesced into categories and categories are coalesced into themes, which should directly answer the research questions.

A thought experiment usually helps to be sure the themes answer the research questions (RQs). Imagine a friend invites a doctoral learner for coffee. During the course of the conversation, the friend asks the doctoral learner one of the research questions for the dissertation study. The response to the question is one of the

theme(s). If the friend looks at the doctoral learner in a funny manner or has to ask a plethora of clarifying questions, then the doctoral learner has not answered the question. For example, imagine one of the RQs is: How do young adults who were bullied during high school describe the experience of being bullied? So a friend is going to ask the doctoral learner this question. Imagine the doctoral learner's response is: Positive bully. The friend is going to look at the doctoral learner funny because he or she has not answered the question. Imagine, instead, the response is: Young adults who were bullied during high school describe learning how to confront the bully without physical violence. This response is a better theme because the doctoral learner has answered the question.

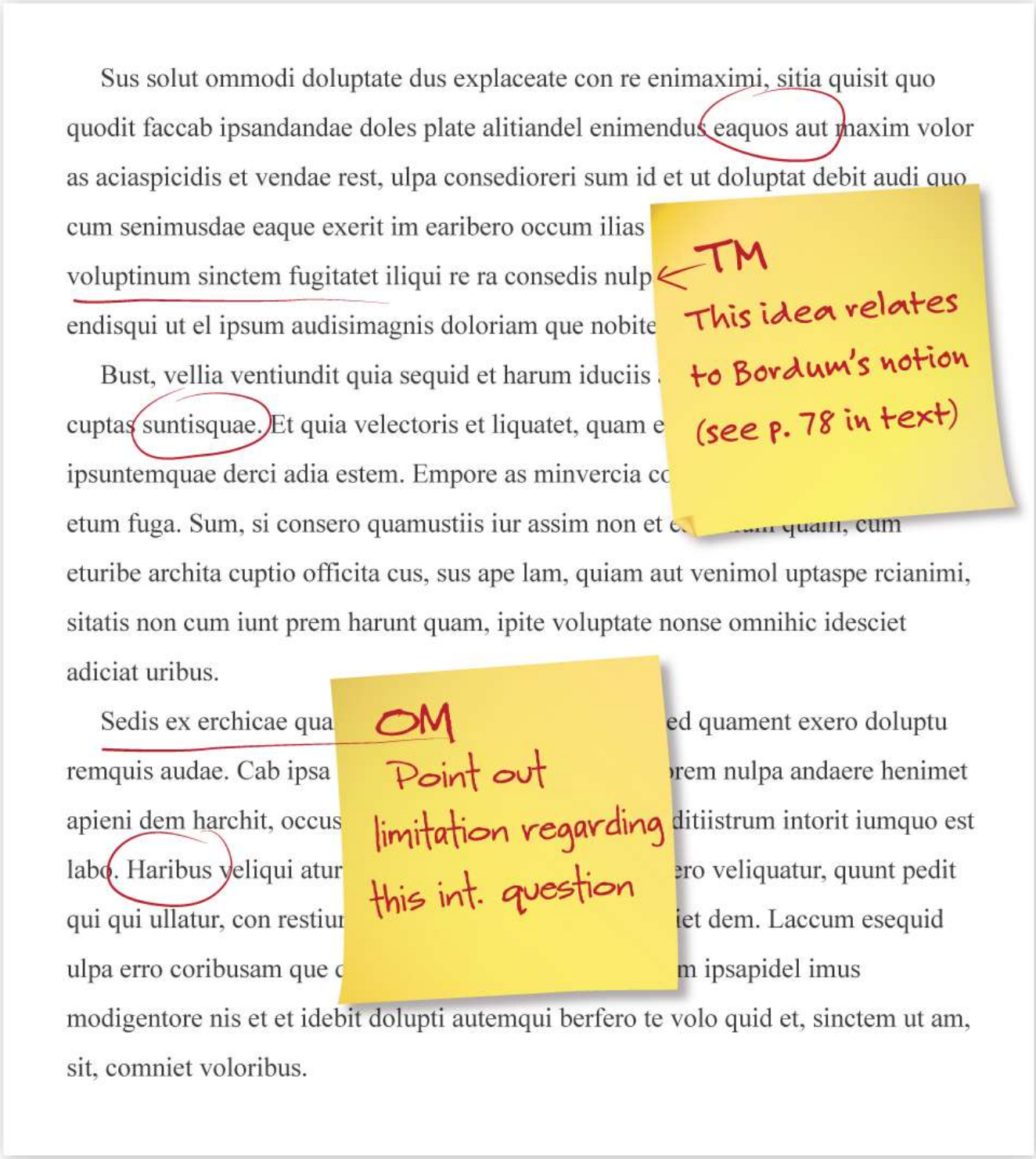
For a simpler example, imagine a friend asks a researcher what is his or her favorite flavor of ice cream. The researcher probably will not respond with one word, and probably not just with a word like: ice (theme). So, the interchange would look like this: "What is your favorite flavor of ice cream?" The answer is "Ice." Though ice might be related to the question (RQ), this is not an answer to the question, right? So, how might the question be better answered? Probably with a response similar to the following: "My favorite flavor of ice cream is mint chocolate chip." With this response, the research question is answered. Many novice researchers develop themes that do not directly answer the RQs or are only slightly related to the RQs. The themes need to be revised and reworded (without changing their meanings) so as to provide a clear one-sentence answer (theme) to the RQ. Remember, for the purposes of this chapter, codes are coalesced into categories; categories are coalesced into themes; themes are filtered through the research questions (Do the themes answer the research questions?); and then the themes are reworded without changing their meanings to demonstrate clear alignment between the theme and the research questions.

Memoing as a Technique to Develop Themes

One important analytic technique to use when developing themes is memoing. Memoing means writing memos or notes to oneself while engaging in the process of data analysis. Memoing can be used throughout the research process; however, this chapter addresses qualitative data analysis, and the focus is on analytical memoing. Memoing is a technique to record the process of data analysis and capture one's thoughts when an idea arises during the data analysis process. One can envision reading through a stack of transcript printouts with pen, highlighter, and sticky notes at hand. Imagine several codes have been developed, and patterns have started to become apparent. The researcher starts the process of coalescing the codes into categories and records this process—the logic of why the researcher coalesced certain codes into certain categories—and reflects on other manners of coalescence of codes to categories. This way, the researcher is constantly thinking and rethinking the data analysis process. Remember, in the final product of the research (the dissertation), the doctoral learner will need to demonstrate and discuss the logic of moving from codes to categories and from categories to themes. Memoing is a means of both recording and practicing explaining one's logic of coalescence of codes to categories and the logic of coalescence of categories to themes. Refer to Barisich (2014, p. 176) for an example of memoing (see Figure 7.3) on a page of hand-coded text with memoing. Qualitative software programs contain memoing features as well, allowing for search and export of information.

Figure 7.3

Example of Handwritten Notes on Hard Copy of Transcript



Organizing and Reporting Analytic Procedures and Results

With the analyses of the data completed, the researcher must write up the results in the dissertation. There are three sections in Chapter 4, the results chapter, of the dissertation. These are "Descriptive Data," "Data Analysis Procedures," and "Results."

Descriptive Data

The "Descriptive Data" section of Chapter 4 describes basic information about the target population and sample and the analysis is written in narrative form. As the name suggests, the section describes the dataset, which includes the number of participants, demographic data, the number of interviews, observations, the duration of the interviews, and the number of transcribed pages that are single-spaced, using 12-point Times New Roman font. The dissertation template outlines the required contents of this section. In addition, it includes graphic organizers that illustrate the dataset. Table 7.2 is an example of a data display as a table that summarizes the data collected. The column headings should reflect the particular sources of data, methods, and participants pertinent to the researcher's study. The descriptive data section can also include a summary description of the participants.

Table 7.2

Summary of Data

	Interview Duration	# Pages Transcribed Single-Spaced	Focus Group Duration	# Pages Transcribed Single-Spaced
Pseudonym 1	60 minutes	15	N/A	N/A
Pseudonym 2	45 minutes	10	85 minutes	38
Pseudonym 3	90 minutes	22	85 minutes	30

Data Analysis Procedures

The "Data Analysis Procedures" section of the dissertation focuses on the processes and procedures used to analyze the data. Tables 7.3–7.5 are examples of tables that display coding, category development, and thematic development. In addition to describing the processes of coding and further analyses in the text narrative, the tables show the evidence of how the analysis was conducted and the basis for the discussion of the results.

Table 7.3

Initial Codes and Emerging Categories

Initial Code	Example from Data	Analytic Notes on Emerging Categories
Write down the name of the code here.	Paste in quotation from the data where the code occurred; provide note regarding the source (e.g., interview, observation, focus group, participant pseudonym).	Briefly note how this code may start to be clustered with other codes. This displays emerging categories.
Write down the name of the code here.	Paste in quotation from data; provide note regarding the source (e.g., interview, observation, focus group, participant pseudonym).	Briefly note how this code may start to be clustered with other codes. This displays emerging categories.
Keep writing down initial codes until they have all been written down.	Paste in quotation from data; provide note regarding the source (e.g., interview, observation, focus group, participant pseudonym).	Briefly note how this code may start to be clustered with other codes. This displays emerging categories.

Table 7.4

Categories and Emerging Themes

Initial Code	Secondary Code/Category	Analytic Notes on Emerging Themes
Write down the name of the initial codes here.	Category into which the codes have been coalesced.	Provide a brief note explaining the logic for grouping/coalescing the codes into this category.
Write down the name of the initial codes here.	Category into which the codes have been coalesced.	Provide a brief note explaining the logic for grouping/coalescing the codes into this category.
Keep writing down initial codes until you have written them all down.	Category into which the codes have been coalesced.	Provide a brief note explaining the logic for grouping/coalescing the codes into this category.

Table 7.5

Themes

Write the categories.	Theme into which the categories have been coalesced.	Provide a brief note explaining the logic for grouping/coalescing the codes into this theme.
Write the categories.	Theme into which the categories have been coalesced.	Provide a brief note explaining the logic for grouping/coalescing the codes into this theme.

Results

The "Results" section of the dissertation involves writing a substantive discussion of the themes that were developed from the analysis and that answer the research questions. Researchers commonly organize results by research question and the themes that answer the specific research question. Some themes might answer multiple research questions. Many researchers still list the RQ and then all of the attendant themes (themes that answer the RQ).

It is important to provide evidence from the thematically organized data when reporting the results. The graphic displays in the data analysis section usually include a few modest segments of transcript. By contrast, the discussion of results needs to include significant portions of transcript or visual media. See Barisich (2014) and Frawley (2016) for examples of the inclusion of transcription and the use of graphic displays in the discussion of results.

Synthesized Discussion of the Results

With the results of the analysis presented at length and in detail, the next step is to provide a synthesized discussion of the results. There is more than one method to accomplish this depending on the research design. The synthesized discussion may include any of the following:

1. Case study—Summarize and/or visualize the components that define the case. Remember, a case study design looks at a phenomenon from three different perspectives. Present those three perspectives and then compare and contrast. This is called a case study report.
2. Grounded theory—Present the themes and then present the model that was created as a result of the themes.
3. **Phenomenology**—In this design, the researcher will have used a phenomenological reduction as a means of analyzing the data (Giorgi, 2018). Phenomenology comes from philosophy but has been adopted largely by psychology as a means of understanding how people make sense of the world around them. The result of a phenomenological reduction is a constituent. A constituent provides insight into the consciousness/meaning/unique perspective of the sample.
4. Qualitative descriptive—Present the themes and discuss in detail how each theme answers which research question.
5. Narrative design—Depending on the purpose of the narrative design, one will have various means of presenting the results. For example, if one elicited stories from the sample, one will probably take those stories and engage in a thematic analysis of some sort. The result will be themes. The themes need to be discussed in relation to the research questions. Sometimes, the researcher will need to "re-story" the narratives, depending on the purpose and RQs. This means the researcher might present all of the common features of the narratives told in addition to the themes.

Table 7.6 outlines the focus of thematic analysis and synthesis across research questions for the GCU Core Designs.

Table 7.6

Comparison of Thematic Analyses for the Core Designs

Design	Focus of Thematic Analysis	Analytic Approach	Synthesis Across Research Question
Phenomenology	Describe the lived experiences of the participants (how participants experienced/felt about particular moments or events in their lives).	• In-depth interpretive analysis of the latent symbolic meaning of individual's experiences. • Requires general familiarity with and understanding of the culture and language of the individuals' identity group (e.g., ethnicity, gender, social class, age, profession, etc.), community or organization. • Data analysis approach may involve thematic analysis and coding from a phenomenological perspective. Interpretive interview summaries (2–3 pages long, single-spaced) done systematically to produce main themes may substitute for coding procedures; however, such an analytic approach requires expertise.	Synthesize the description of the phenomenon across the research questions
Narrative	Describe the stories and kinds of stores that include characters, plot, interactions, emotions, turn of events, suspense, action, sense of consciousness, imagination, and values.	• Thematic analysis for narrative inquiry (See Saldaña (2016)).	Develop a "re-story" based on the synthesized elements including the characters, plot, interactions, emotions, turn of events, suspense, action, sense of consciousness, imagination, and values.
Grounded Theory	Identify the concepts or components that will comprise the theory, model, or process.	• Inductive thematic analysis (open, axial, and selective coding)	Synthesize and visualize the model, theory, or process.

Case Study	Identify patterns that may provide explanations or rival explanations of the phenomenon.	• Ideally inductive thematic analysis with triangulation (across data sources); deductive analysis can be used as well. If deductive analysis is used, justify. Descriptive statistics of quantitative instruments, such as Likert-scale questionnaires. • To complete the study, learners must provide a triangulation-style analysis in Chapter 4, showing how analytic results **from each data source** are woven together to produce a holistic account. Note that for studies that mix qualitative and quantitative instruments, there must be a demonstration of how the quantitative data (e.g., Likert-scale results) help contribute to the final holistic account.	Summarize and/or visualize the components and variables that define the case and possibly rival explanations.
Qualitative Descriptive	Describe the topic being researched.	• Basic, generic, thematic analysis. Can be inductive, deductive, or both. Preferably, start with inductive. For providing the results, the themes are applied to answering the RQs in descriptive fashion (no additional triangulation or master narrative is required).	Develop a summary of the research findings across the research questions.

Qualitative Software Programs

The volume of qualitative data can prove to be a challenge to researchers. Analyzing multiple codes, overlapping codes, and the related examples is challenging. As Gurdial Singh and Jones (2007) noted: “Traditionally, qualitative data analysis has always been associated with sifting through bulky and often large volumes of transcripts, case notes, articles and even pictures which can be quite tedious and messy” (p. 3). For this reason, qualitative software programs can be very useful. However, it is often worthwhile to begin the coding process by hand coding, which involves printing transcripts and other relevant sources of data, reading them, and noting patterns based on the coding practices described above.

Hand Coding and Digital Coding

Hand coding involves printing transcripts and other relevant sources of data, reading them, and noting patterns based on the coding practices described above. One can easily imagine a stack of transcripts on the desk, a notebook, pens, highlighters, and sticky pads at hand. While reading the transcripts, the researcher will begin to notice aspects within the data such as similar stories, patterns, or ideas. As the researcher notices these, he or she will underline them. A phrase or a story might bring a particular article to mind, in which case the researcher will procure a sticky note and write on it, “TN” to indicate a theoretical note. The researcher will then note how that phrase or story seems to exemplify something in the theoretical framework or a significant empirical article discussed in Chapter 2 of the dissertation. After a while, the researcher will have underlined, circled, and highlighted various passages. As the researcher observes patterns, those patterns are named. These are initial codes or open codes. The researcher writes down the code in the notebook and identifies what the code represents. This initial hand-coding process puts the researcher in a close and engaged relationship to the data.

Once the researcher has identified a set of initial codes, the digital files can then be uploaded to any qualitative software program, generically known as computer assisted qualitative data analysis software (CAQDAS) or qualitative data analysis (QDA). Once the digital files are uploaded to the software program, one can look at the various manners the program presents the data. The ways in which the program presents the data can help the researcher to think and rethink the coding. For example, a program might provide one with word frequency counts. Frequency counts are not codes; however, if the word "empowered" appears very frequently, one might go back to the corpus and look at where, when, and how this word is being used and whether one has encapsulated "empowerment" within the codes. In other words, these programs are there to help one think about the data analysis process; they are not there to do the data analysis for the researcher.

Pros and Cons to Using Qualitative Software Programs

There are pros and cons to using qualitative software programs. Advantages include being able to code the data within the program (not having to print and use highlighters etc.), including using overlapping codes, and then having the program export the codes in context, that is, with examples of the code in the context of a paragraph. Programs have the capability to identify researcher-coded excerpts by participant, role, age, or gender to facilitate analysis.

A distinct disadvantage of such programs occurs when the researcher relies on the program to do the thinking. For example, CAQDAS programs can generate a list of frequently used words. While handy, those words are decontextualized from phrases, paragraphs, or participants, and are likely to be unrelated to the key issues and meaning relevant to the study. As described in the coding section above, it is incumbent upon the researcher to identify codes as meaningful units in relation to the study, problem statement, research questions, and/or theoretical framework so that the analysis is meaningful and integral to the study.

Researchers should think of these programs as tools for displaying data in various manners. These programs can display data in various manners, which help one to think about the data. The various displays and presentations of data are designed to give the researcher insights into the data that might carry over into the analysis process. Researchers do not want to think of these programs as doing the coding for them. Remember, data analysis is an art and a science (see above). As both an art and a science, a human needs to do the data analysis thinking and explanation. The program is a tool for spurring thought and gaining perspectives on the data.

There are a number of qualitative software programs described as CAQDAS and QDA. Some programs are available for a fee, such as NVivo, MAXQDA, ATLAS.ti, QDA Miner, HyperRESEARCH, and Dedoose. Others are available as open source programs. Many programs are available for both Mac and PC platforms, and some work on both. Check the software producers’ websites for current updates on particular products.

Qualitative Software Programs

Websites for Qualitative Software Analysis

- MAXQDA: **http://www.maxqda.com/ (http://www.maxqda.com/)**
- NVivo: **http://www.qsrinternational.com/ (http://www.qsrinternational.com/)**
- Dedoose: **http://www.dedoose.com (http://www.dedoose.com)**
- ATLAS.ti: **http://atlasti.com/qualitative-data-analysis-software/ (http://atlasti.com/qualitative-data-analysis-software/)**
- QDA Miner: **http://provalisresearch.com/products/qualitative-data-analysis-software/ (http://provalisresearch.com/products/qualitative-data-analysis-software/)**
- HyperRESEARCH: **http://www.researchware.com/products/hyperresearch.html (http://www.researchware.com/products/hyperresearch.html)**

Common Pitfalls of Qualitative Research

The purpose of qualitative research is to discover how and why participants from a target population perceive, experience, act, and/or reflect in a manner that illuminates one's understanding of a phenomenon. Research questions that ask how or why something occurs seek information that requires a "deep dive" during data collection and produces a large volume of data. As one might expect, there are some common pitfalls associated with conducting this sort of research.

Collecting Too Little Data

What constitutes enough data? As a general guideline, semistructured interviews should last 60–90 minutes. This duration allows the researcher to fully explore a phenomenon with a participant. The researcher will ask follow-up questions, clarifying questions, and attempt to get the participant to explore his or her thoughts and descriptions in as much detail as possible. Although a 60–90-minute duration does not ensure the quality of data, this time duration usually allows for a more full and complete exploration of the phenomenon.

In order to address quality of the data, a researcher might consider doing numerous field tests before beginning data collection. The purpose of a field test is to see what kinds of data one might get with the interview questions and/or focus group prompts that were developed. In a field test, the researcher engages in an interview with someone from the target population just as if collecting actual data. This means the interview is recorded and transcribed. The researcher reflects not only on the quality of data (will these data answer the RQs?) but also the quantity of data (does the interview last 60–90 minutes? Does the interview result in a sufficient number of pages?). Doctoral learners might consider working with their chairs and/or methodologists to look critically at the results of the field test and make the necessary changes to interview questions and/or focus group prompts. The researcher also might consider practicing data analysis process using the data from the field test. Remember, the data from the field test cannot be used for the final analysis. The purpose of the field test is not to collect data, but to hone the data source (e.g., the interview questions, the focus group prompts).

Data from a field test cannot be used for the final analysis.

Letting the Software Do the Thinking

Qualitative software programs are powerful and useful in data analysis. Most, if not all, programs have numerous features that allow the user to display the data in various manners (e.g., word clouds, word frequency charts). The software is only there to spur one's thinking. Look at the various displays of data and engage in various searches for words or concepts in data, but keep in mind that these tools are to help one think about the data in new ways. Computer software is much like an obedient dog: It does what it is asked—it does not do the thinking.

Finding the Answer That the Researcher Wants Rather than Discovering What Really Exists

Qualitative research aims to explore and discover meaning. After all, there is no point in conducting the study if the results are known in advance. The point is to design a study to understand what is not known. To this end, the researcher must ask open-ended questions, use multiple sources of data when appropriate, aim to elicit a large volume of data, triangulate during the analysis, and be open to learning what the analysis reveals.

Additional Qualitative Data Analysis Resources

- Boyatzis, R. (1998). *Transforming qualitative information: Thematic analysis and code development*. SAGE.
- Cohen, D., & Crabtree, B. (2006). *Qualitative research guidelines project*. http://www.qualres.org/index.html
- Crabtree, B., & Miller, W. (1999). A template approach to text analysis: Developing and using codebooks. In B. Crabtree & W. Miller (Eds.), *Doing qualitative research* (pp. 163–177) SAGE.
- Denzin, N. K. (1978). *Sociological methods*. McGraw-Hill.
- Dwyer, S. C., & Buckle, J. L. (2009). The space between: On being an insider-outsider in qualitative research. *International Journal of Qualitative Methods, 8*(1), 54–63.
- Evans, G. L. (2013). A novice researcher's first walk through the maze of grounded theory: Rationalization for classic grounded theory. *Grounded Theory Review: An International Journal, 1*. http://groundedtheoryreview.com/2013/06/22/a-novice-researchers-first-walk-through-the-maze-of-grounded-theory-rationalization-for-classical-grounded-theory/
- Grand Canyon University (Ed.). (2015). *GCU doctoral research: Foundations and theories*. http://www.gcumedia.com/digital-resources/grand-canyon-university/2015/gcu-doctoral-research_foundations-and-theories_ebook_1e.php
- Groenewald, T. (2004). A phenomenological research design illustrated. *International Journal of Qualitative Methods, 3*(1), 15–16.
- Lin, C. S. (2013). Revealing the "essence" of things: Using phenomenology in LIS research. *Qualitative and Quantitative Methods in Libraries (QQML) 4*, 469–478.
- Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *HSR: Health Services Research. 34*(5), 1189–1208.
- Schutz, A. (1967). *The phenomenology of the social world* (G. Walsh & F. Lehnert, Trans.). North Western University Press. (Original work published 1932.)
- Vogt, W. P., Vogt, E. R., Gardner, D. C., & Haeffele, L. M. (2014). *Selecting the right analyses for your data: Quantitative, qualitative, and mixed methods*. Guilford.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobstrusive measures: Nonreactive research in the social sciences*. Rand McNally.

Checks for Understanding

1. True or false: Qualitative data analysis is subjective and personal.
2. True or false: Each core design must have one specific kind of analysis.

3. True or false: Both inductive and deductive coding can be included in the analysis of data.
4. True or false: Coding data is never a good idea.
5. True or false: Qualitative software reduces the work load of the researcher.

Answers

1. False
2. False
3. True
4. False
5. False

References

Barisich, D. G. (2014). *Enhancing new teachers' performance through mentoring: A phenomenological study* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3667770)

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa.hdl:2027.42/138221. S2CID 10075179

Frawley, L. (2016). *Understanding interscholastic athletic administrator experiences and challenges: The importance of mentoring and collaboration* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 10111647)

Giorgi, A. (2018). *Reflections on certain qualitative and phenomenological research methods*. University Professors Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine Transaction.

Gurdial Singh, R. K., & Jones, M. L. (2007). Qualitative data analysis: Making new discoveries and aligning old strategies. *Proceedings of the 4th International Qualitative Research Convention Doing Qualitative Research: Processes, Issues, and Challenges* (pp. 1–22).

Husserl, E. (1982). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy - first book: General introduction to a pure phenomenology* (F. Kersten, Trans). Nijhoff (Original work published 1913)

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods, 15*(1), 85–109.

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.) [Kindle edition]. SAGE.

Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). SAGE.

GRAND CANYON UNIVERSITY