

■ What Is Qualitative Data?

Most would agree that qualitative data generally take the form of words, not numbers. Modern writers include visual, audio, or graphic data in the definition, as well as verbal or textual data. While some argue that qualitative data can be transformed into quantitative data, I think it is those who practice a traditional or fundamentalist paradigm who take this position. If you support a more inclusive position, then almost any data you gather from, by, or about your study can be seen as qualitative data.

Suppose you are interested in studying single-sex classrooms. Some schools have adopted the practice of organizing classrooms by gender. There are merits and disadvantages on both sides. In a study involving British people born in 1958, Sullivan and her colleagues (2010) found this type of schooling positive for girls but neutral for boys. The American Civil Liberties Union opposes the idea, but advocates see that teaching can be targeted and distractions reduced.

You are interested in going beyond the statistical data. Let's look at the kind of qualitative data you might collect: interview data with students, teachers, and parents; observational data recorded in note form of classroom practices and student behaviors; documents related to the issues; shadowing of selected students; photographs or videos of students interacting; your notes regarding your thoughts about the practice; student work products, either on the computer or in hard copy; student chat room comments regarding their feelings about participating in this type of class; and your observations about the classroom's physical appearance and the appearance of the students. All are legitimate types of data. No one form of data is better or more legitimate or more meaningful than another type. You are limited only by your creativity and the available technology.

As you can imagine, you will have an enormous quantity of data in somewhat different forms. You will most likely transcribe your interviews and observational notes into a word processing program. You may have some data already on your computer, taken from chat room discussions or student work products. All your verbal data can be organized in a word processing program; you will have to devise a way to organize your visual or graphic data. Some qualitative software programs incorporate visual and audio data and provide ways to analyze them.

I want to stress that the data will be collected not at one time, but at several times across the life of your research project. In the same way, your analysis should cover the life of the project and should begin as you begin collecting your data. Planning how you will do your analysis should precede actual data collection. Let me emphasize that your plan should represent a general guide and may be modified as necessary, depending on the data you collect and the available tools for analysis.

■ When to Do Your Analysis

I see analysis as an ongoing process throughout the life of a project; however, we write about it in a linear fashion and often researchers get so involved in data collection that they do not begin any serious analysis until all data are collected. It would be ideal if researchers could follow a circular model of gathering and analyzing data. Often, a researcher will enter data into a computer program—a word processing program or qualitative software program—in concert with collecting additional data. Even when a researcher makes a decision not to conduct analyses using a computer, he or she organizes the data on the computer. Having entered the first piece of data—an interview, some field notes, or the current teaching unit—a researcher begins the process of analysis. Some do this informally while others proceed in a more formal manner.

■ Gaining Meaning: Themes or Narratives

In many traditions, there appears to be general agreement that the goal of analyzing the text and words collected is to arrive at common themes. Most procedures involve a process in which the researcher chooses to code words, phrases, segments, or other portions of text. Some people believe that the codes should be determined *a priori*. However, most take the position that the codes emerge from the data via a process of reading and thinking about the text material. Aside from a specific process identified with grounded theory, coding is usually done through a careful reading of the text. I have seen some people read the text and mark large chunks of material with codes. Others work from a microlevel and code text chunks or segments. Whatever the process, and I believe it varies by individual and perhaps even by type of data, the initial goal is to arrive at a manageable number of codes.

I see the process as one of organizing and categorizing. You begin with a large amount of material, for example, the text of an interview. That material is dissected and categorized into codes. Next, you proceed to a second interview. Again, dissect and categorize the data into codes; you can use the previous codes or add new codes. This iterative process continues until you have coded all your interviews. By this time, you have reviewed many interviews and coded them. You can now review your codes and look for ones that overlap or are redundant. You might find that you will rename some of your codes. You will likely generate many codes. These codes can then be organized into hierarchical categories, in which some codes will be subsets of larger categories. You might have 80 to 100 codes that you then organize into 15 to 20 categories and subcategories. These categories can then be organized into five to seven concepts. As a general rule, even large data sets do not reveal more than this small number of central and meaningful concepts about the topic of interest.

There are various ways to conceptualize data analysis in qualitative research. I want to discuss two of them: identifying themes and telling stories. Much of the writing about analysis deals with identifying themes. Here is the idea in a nutshell: You gather a large amount of data. It might come from one individual over time; it might come from several individuals; it might come from one or a number of settings; or it might be derived from other sources (chat rooms, interviews, observation notes). You might collect all the data or several people might collect data and combine the data. All data are gathered in order to answer your research questions. Now, the data are usually so large and disparate that they make no sense without some thought and organization. As a researcher, it is your task to provide those steps. I provide detailed information in the next section on steps you can follow to move from the data to development of themes and ultimately to meaning.

I see this as a process of sorting and sifting. Imagine that you have a large sieve. But it doesn't look like a traditional sieve. Some holes are square, some round, and some irregularly shaped. You put into the sieve a number of objects—some round, some square, and some irregularly shaped. You shake the sieve. The round ones drop through the round holes. The square ones drop through the square holes. Some of those irregularly shaped drop through the odd-shaped holes, while others stay in the sieve. You have sorted your objects based on a system. Some fit well while others do not. Or, think of how you could take items from the pockets of a group of children and put them into a large bin. Your assignment is to organize the items into three piles. Well, you might place edibles in one stack, games or toys into another, and items related to school into a third. Or, shifting

your thinking entirely, you might place disposables into one pile and those that are permanent into another. So, there is no right way to do this. And, you are no more or less expert at doing the task than the next person. Organizing the data you collect takes on the same kind of challenge. There are no clear rules, nor is it obvious how you should begin. As a beginning researcher, I suggest you begin by practicing. By following a process of coding and looking for themes, you can begin to make sense of your data.

One limitation of this type of analysis is that it operates from a reductionist perspective. Do we really believe that we can capture so much of what a person thinks and feels and portray it in five or six basic themes? Some would argue that by doing this, we are trying to move into an analytic mode that is more closely allied to principles of quantitative paradigms. An alternative approach to an analysis that identifies themes is the emphasis on finding the narrative or telling stories (Coffey & Atkinson, 1996). The intention is to examine how such stories can be used as structured or formal ways to transmit information. You can read in greater detail Denzin's (1989) account of interpretive biography. Connelly and Clandinin (1990) have written extensively on the topic. Baumgartner (2000), in her study of HIV-positive adults, shed light on exploring how stories can be used as a source of data. Although Richmond (2002) proposed some specific guidelines for a narrative analysis, most researchers follow their own plan and often do not describe details in their published work. Guy and Montague (2008) analyzed the personal narratives about men's friendships. Zilber, Tuval-Mashiach, and Lieblich (2008) stressed the importance of context in the construction and understanding of life stories. I particularly like Coffey and Atkinson's (1996) admonition:

There are no formulae or recipes for the “best” way to analyze the stories we elicit and collect. . . . Such approaches also enable us to think beyond our data to the ways in which accounts and stories are socially and culturally managed and constructed. That is, the analysis of narratives can provide a critical way of examining not only key actors and events but also cultural conventions and social norms. (p. 80)

Some researchers have experimented with visual narratives. Researchers might rely on narratives generated by participants or those they generate. For details using this approach you can examine a Visual Research Methods course taught at the University of California, Berkeley, during the spring of 2011 (<http://blogs.ischool.berkeley.edu/VizNarrative/potential-readings/>). Others combine multiple types of data: visual, written, and spoken. Keats's (2009) point is well taken: “Studying narrative texts aids the researcher in understanding how participants experience, live, and tell about their world” (p. 181). Li and her colleagues (2011) have provided an excellent example of visual narrative combined with text.

I am suggesting that you can either conduct an analysis in which your goal is to identify themes or conduct an analysis in which your goal is to provide an interpretation of the data by telling or retelling a story. Neither way is “right.” The process you follow to get to the end depends on your goal.

Writing about the analysis process is linear; in contrast, actually doing an analysis is anything but. You will be faced with many questions you need to answer and decisions to make. It is often the case that you know in advance the main types of data you will collect. However, as your project develops, you might discover that additional data become available. You may decide in advance

that you will use a computer software program to analyze your data. However, the program you want may not be readily available, or you may think you can learn how to run a program that turns out to be much more complicated than you anticipated. You may decide that you are going to concentrate on one aspect of a problem and then find that the data you collect lend themselves to exploring totally different arenas. You may decide to incorporate images in your data, but you are not really able to determine how best to include the visual data and how to incorporate them into an analysis. The process may appear to be relatively clear and systematic; however, in reality, you might find yourself getting bogged down in details you did not expect. You might find that you want to capture information from the Internet (e.g., chat room discussions, Listserv comments, the blogosphere, YouTube), but you do not know enough about the logistics to do this effectively. Shulman (personal participation, webinar, 2011) has developed Discover Text that enables the researcher to capture data from Facebook or Twitter in real time and apply preliminary codes.

■ Conducting an Analysis for Central Themes

The goal of qualitative analysis is to take a large amount of data that may be cumbersome and without any clear meaning and interact with it in such a manner that you can make sense of what you gathered. Again, you should not be surprised that there is no right way to do this. In fact, there is less written about the mechanics of doing such analysis than about any other topic in qualitative research. When authors do write about the process, they are quite vague. I propose here a process that I have used over many years. I suggest you think about it as a starting point, rather than a prescription. I hope you will find the ideas useful.

Getting Started

Qualitative research is usually a solo activity. You collect data on your own, analyze it on your own, write it on your own, and are responsible for what you say. But we know that much research benefits from interacting with others, trying your ideas out on others, and learning about the reaction of others to your ideas.

I know that students learn by doing and practicing. I encourage you to work with small groups of students as you embark on looking for meaning in what you have gathered. Here is an exercise from Barbara Kawulich (personal communication, 2008; exercise used with permission of Barbara Kawulich) that she calls “Hot Monkey Sex”: Students are given three Post-it Notes each. On the first, they write their answer to the question “Given all the money you need, where in the world would you like to go for a month’s vacation?” On the second, they write down whom they want to go with. On the third Post-it Note, they indicate what they want to do on the vacation. Students work in pairs to analyze the data by organizing the Post-it Notes in various ways to tell a story the responses generate. Kawulich chose the title for the exercise while working with a class consisting of several young teachers and one older, quiet teacher who typically did not get involved in the class. This activity really got the quiet teacher involved: Her Post-it Notes revealed that she wanted to go to Hawaii with Brad Pitt and have hot monkey sex. “The younger teachers roared with laughter, loving her openness and appreciating the fun-loving side we had never before seen.” Kawulich reported that once she renamed the exercise, students would bring cameras to class because their friends wanted to see what “Hot Monkey Sex” looked like. Exercises like this can help you see how the codes emerge from the data and how no single scheme is better than another.

Preparing and Organizing Your Data

Once you have gathered some data, you need to put it into a format that is useful for analysis. In most cases, you will find a way to transcribe interviews, capture online discussions, or otherwise put words and text in a useful format. You should also think about visuals or audio that is not transcribed. I recommend that you place each item in a separate file, using a word processing package. It is helpful to insert your comments in brackets, and in a different font or color.

Make a folder and label it, for example, My Qualitative Research Project. You will place several files in this folder, depending on how much data you have collected. These files can be individual interviews and/or your observation notes or your researcher journal. At the very least, you will place your data and your journal files in the folder. It is helpful if you label each file in a systematic manner. For example, suppose you have four interviews: two with the same person and two additional ones. You would create the following four files: DonaldInt1, DonaldInt2, DavidInt1, DanielInt1. In a large project, you might have observation data as well: DonaldObs1, DavidObs1, and so on. Of course, your choice of file names depends on the type of data you collect. Some researchers like to incorporate a date in the file name.

In addition to these data files, you will want to create your researcher **journal**. Make another file and label it Researcher Journal. Some people also put information collected from a literature review in this folder. This folder should be created when you begin your research, not when you finish it. You should plan to keep adding to it as you move along.

Make sure you save a copy of this folder in a location other than your hard drive. In today's computer world, keeping copies of your files is somewhat easier than in times past. But I encountered a serious problem and lost an entire book chapter recently when my file became corrupted and I could no longer save it. An external hard drive became the savior.

Reviewing and Recording Your Thoughts

Most people find it helpful to read through all the material in their folders. In keeping with the iterative nature of the process, you can begin by reading a transcript. Add your thoughts and comments to your Researcher Journal file. It is okay to use informal writing here. Remember to date your notes. Your comments might look something like this:

9/15/2008. Read through the transcript of DanielInt1. Daniel certainly had a lot to say. I wish I had asked him more about why he decided to leave the field of teaching. I will need to remember to do that in my next interview and if I go back with him as well.

9/20/2008. Finished my second interview with Daniel. Glad he clarified his thoughts on this topic. Not sure I would have picked this up unless I had read what he said.

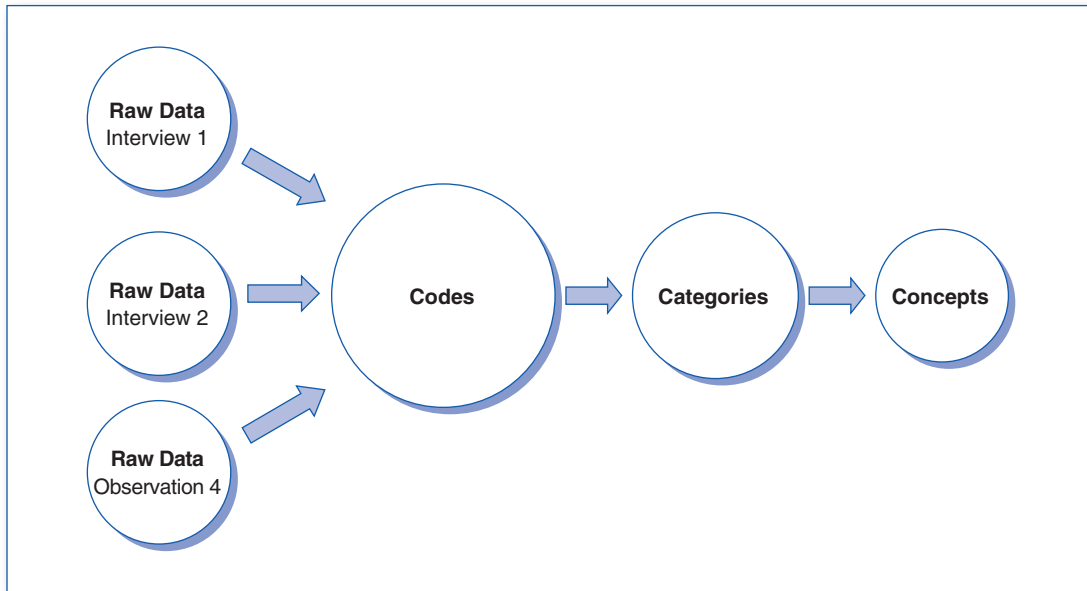
The Three Cs: Coding, Categorizing, and Concepts

You are now at a point where you can see how to move from raw data to meaningful concepts or themes. I call this the three Cs of analysis: from coding to categorizing to concepts.

Coding interview data, observational notes, and text into meaningful chunks is a challenging task. Whether you work with a word processing program or with other software, it is

your responsibility to generate the codes. Do not expect that a computer program will generate codes or organize them; rather, you will need to provide the input. I have broken down this process into six steps (see Figure 12.1).

Figure 12.1 Three Cs of Data Analysis: Codes, Categories, Concepts



The six steps are as follows, and are subsequently explained in detail:

Step 1. Initial coding. Going from responses to summary ideas of the responses

Step 2. Revisiting initial coding

Step 3. Developing an initial list of categories

Step 4. Modifying initial list based on additional rereading

Step 5. Revisiting your categories and subcategories

Step 6. Moving from categories to concepts

Step 1. Initial coding. An initial code can be a word, a phrase, or the respondent's own words. You come to it by a careful reading of the text. In my coding of the Armstrong interview, I used brief phrases for my coding. Even if you have only collected a small amount of data, it is not too early to begin coding. Select any transcript. Read the initial page or two. Use the "Comment" function in your word processing program to insert your initial codes (in Microsoft Word, you will find the function on the "Insert" menu). Enter your initial codes. Continue reading the transcript while

entering different codes. Upon completion of initial coding with one transcript, select another transcript and continue the same process. Box 12.1 is an example that might help you see this more clearly. The researcher's codes are in brackets.

Box 12.1 EXAMPLES OF INITIAL CODING

Transcript 1. Partial Interview: Cross-Gender Friendship

It was sophomore year in college. We knew each other—or at least who each other were—from freshman year. Sophomore year on the first day of classes, we met during some orientation. She was orientating freshmen. I was probably hanging around, looking for something or other to get into around campus. About four months after that we started hanging out constantly. [maintenance as friendship only]

We just—or at least I tried to—stay out of situations where it could have turned, become physical. [physical attraction] And we tried not to talk about it—those kind of things. [evolution into something more] After about eight months or so, things shifted. We both realized our feelings had changed. We tried to hold off as long as we could and keep the friendship as long as we could. But we started going to the next step. [potential problems]

One of the main things is that since I'm not actively seeing another person or actively engaged, is that when I meet a girl, any girl, is potentially more than a friend. [tensions/barriers]

Transcript 2. Partial Interview: Cross-Gender Friendship

Someone that I can talk to intimately. Someone that I can tell just about anything. [intimacy, talk to] ... It's just a spark similar to physical attraction [physical attraction] but it's different. You know, when you talk to them that, you know, you may not agree on things necessarily, but you can understand each other.... natural progression. [something more]

Step 2. Revisiting initial coding. By now, you will have developed a large number of codes. Some of them will be redundant, and you will need to collapse them and rename the codes. I have observed that some people tend to code almost every phrase or sentence, while others code larger chunks of information. You need to choose whatever works best. You may want to modify your codes based on an examination of what you have already collected and new raw data. The brief transcript I provide does not give you a chance to show how you might rename your codes. But you should focus on removing redundancies, renaming synonyms, or clarifying terms. If you coded attributes, you should make them consistent, for example, men, boys, teenage boys, and so on. Choose one term to describe the attribute.

Step 3. Initial listing of categories. Now that you have modified your codes, it is time to organize them into categories. I have found that certain codes become major topics, while others can be grouped under a major topic and become subsets of that topic. In essence, you have moved from one long list of codes into several lists of categories, with related codes as subsets of the categories (see Box 12.2).

Box 12.2 EXAMPLE OF INITIAL CATEGORIES

(Subcategories have been omitted.)

- Maintenance
- Physical attraction
- Intimacy
- Tensions/barriers
- Problems
- Issues with boyfriend/girlfriend
- Meaning of friendship
- Issues of homosexuality

Step 4. Modifying the initial list. At this point, you will need to continue the iterative process. You may decide that some of your categories are less important than others, or you may see that two categories can be combined. Remember that your goal in the three Cs analysis is to move from coding initial data through identification of categories to the recognition of important concepts or themes.

Step 5. Revisiting categories. I suggest that at this point, you revisit your list of categories and see whether you can remove redundancies and identify critical elements. In my experience, most novice researchers tend to see everything as important. They appear reluctant to say that one area might reveal more interesting ideas than another. This is where you can exercise your judgment about what is important and what is not.

Here is an example taken from an entirely different context. Suppose you have 100 books and you want to arrange them into five piles. Well, there are a number of ways you can do this. You can sort by color—all blue-covered books together, all green-covered books together, and so on. You can sort by size—grouping all oversized books together. You can sort by topic—all books on science together, all books on humor together, and so on. Or you can sort by author—all books by Roth together, all books by Faulkner together, and so on. You could arrange by publication date—all books published after 2000 together, all from 1990 to 2000 together, and so on. Obviously, some categories make more sense than others, depending on your purpose. Further, you could place the books in subsets within each of the major categories. To continue with my example, you could place light blue books together, navy blue together, and so on. If you arranged by author, you could put major works together and minor works together.

Step 6. From categories to concepts. The final step in the process is to identify key concepts that reflect the meaning you attach to the data you collect. While there are no definitive rules for the number of concepts you might identify, I believe very strongly that fewer well-developed and supported concepts make for a much richer analysis than many loosely framed ideas. As you read and reread your data, you will see that some ideas appear richer and more powerful than others. It is up to you to determine that. I would suggest, as a rule of thumb, that five to seven concepts should be the maximum number that you can find in a set of data. You might consider even fewer

concepts as you become more sophisticated in conducting analyses. Obviously this number depends on the size of the data and the range of the interview. Some information is unimportant even though it is there.

When organizing your codes into concepts, it is your task to decide the most informative or logical manner of sorting. You need to determine from the data what meaning you think can be found. Sometimes your initial thoughts are quite superficial. You will find that reorganizing and rewriting and rethinking often lead to more powerful ideas. At this point, some writers like to weave information from the available literature together with their new data.

Strengthening the Process

By now, you will have completed the six steps in the movement from codes through categories to concepts. To add texture and depth to your analysis, you may want to return to your documents to look for other things that will enhance your interpretation. One promising area to explore is the use of metaphors. Our language is rich with metaphorical allusions and they often reveal much about what others mean. To what extent were metaphors used? Are there sufficient metaphors to incorporate as part of your written paper? If so, can you code them according to certain criteria that may emerge? You might look for type of language, metaphor chosen, or gender-related metaphors.

You could also explore the use of stories. To what extent were stories used? Are there stories that might lead to epiphanies? Are some better than others? Other kinds of things to look for in your data are the richness of detail, conflicting ideas from the same respondent, unusual or unique experiences, or ideas that contradict current thinking on the topic.

I want to reemphasize that making meaning from qualitative data is a process that moves between questions, data, and meaning. Figure 12.2 provides a summary of the data analysis process when looking for themes. Key elements in the model are that it is iterative, circular, and can be entered at any point. You need to try to think of your own work in this way as well.

Figure 12.2 Relationship Between Questions, Data, and Meaning

