



Tips for Analyzing Qualitative Data

Analyzing qualitative data involves reading through the interview or focus group transcripts and other data, developing your codes, coding the data, and drawing connections between discrete pieces of data.

Reviewing your data

Begin data analysis as soon as possible after you begin data collection. Read the data with an eye for themes, categories, patterns, and relationships. Write analytic memos [example?] and/or hold analytic meetings with other CLIP members to capture your initial thinking and tentative ideas about the data. These preliminary reviews may reveal areas that are being overlooked in the interviews and prompt you to allow time to address these in future interviews. The initial themes and categories that you see in the data also inform the codes that you will use for a systematic analysis of your data.

Organizing your data for analysis

The most common forms of qualitative data that CLIPs collect are transcripts from individual and group interviews (including focus groups) and responses to open ended questions from questionnaires.

How you organize your data for analysis depends on whether the data are grouped question-by-question and are woven throughout an interview.

If you cannot easily separate the data question-by-question, you will need to analyze the whole interview for common themes, categories, and patterns. See the section below on developing your codes, coding your data, finding themes, patterns, and relationships, and summarizing your data.

Instruct the person who is transcribing the tapes to indicate along the left side of the transcript who is speaking, i.e., the respondent or interviewer. This is relatively easy to do for individual interviews but may be difficult for a focus group. Detailed notes of the focus group that include the speakers' names or initials will help you to correct the transcript as needed. Leave a wide right margin for coding the transcript and adding comments as you analyze the data.

When you have qualitative questionnaire responses as well as responses to very structured interviews, you can organize responses question-by-question. This makes it easier to analyze the data. The following is one approach to organizing the data in a word processing program.

Example - Questionnaire data¹

Analysis of qualitative data can be facilitated by organizing the data in tables that can be sorted by respondent, question, and other characteristics. Begin by creating a three- column table.

¹Although this process can be used for structured interviews, in this example we only refer to questionnaire data.

List all those to whom questionnaires have been sent. Sort the list alphabetically, and assign a *respondent number* to each person (starting with “1” though the total number of respondents). Use the third column to note when a completed questionnaire has been returned. As each questionnaire is returned, write the person’s *respondent identification number* at the top of the first page of the questionnaire.

Respondent ID #	Respondent Name	Received
1	Allison, David	
2	Baker, Susan	
3	Cairns, Mary	
4	Callahan, Jennifer	
5	Cooper, Richard	
6		

Next create a four-column table, making the first two columns narrow, the third column wide, and the fourth column narrow. Column 1 is for the *respondent ID number*. If you have respondents from various populations you might use other indicators such as **T** for teacher and **S** for student to allow further sorting of the data. Column 2 is for the *question number*. Column 3 is for the *response*. Column 4 is for the code you assign during analysis.

ID#	Q#	Response	Code

To begin entering the responses, place your cursor in the first cell of Column 1. Enter the *respondent ID number*. Let’s say that Jennifer Callahan is the first person to return her questionnaire. Put a 4 in the first cell of Column 1. Then in the first cell of Column 2, put the first *question number*, in this case “1”. In the first cell of Column 3 put her *response* to this question. The Code column will be used later in the analysis.

ID#	Q#	Response	Code
4	1	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	

Continue entering all of Jennifer’s responses. For each response, first enter her *respondent ID number*, then the *question number*, and then the *response*. If she has left a question unanswered, enter “No response,” “NR” or some other placeholder in the *response* column to

indicate that her response to this question was not overlooked and she, in fact, did not respond to this question.

ID#	Q#	Response	Code
4	1	xx	
4	2	No response	
4	3	xxxxxxxxxxxxxxx	
4	4	xxxxxxxxxxxxxxxxxxxxxxxxxxxxx	
4	5	xxxx	

Continue in this fashion until all responses for all returned questionnaires have been entered. You can enter them as they come in or wait until you have a large number of them and enter them at the same time. The order in which you enter the individual questionnaires is not important as they can be sorted by P# or Q# later.

ID#	Q#	Response	Code
4	1	xx	
4	2	No response	
4	3	xxxxxxxxxxxxxxx	
4	4	xxxxxxxxxxxxxxxxxxxxxxxxxxxxx	
4	5	xxxx	
2	1	xx	
2	2	xxxxxxxxxxxxxxxxxxxxx	
2	3	xxxxxxxxxxx	
2	4	xxxxxxxxxxxxxxxxxxxxxxx	
2	5	xx	
5	1	xx	
5	2	xxxxxxxxxxxxxxxxxxxxxxxxxxxxx	
5	3	xx	
5	4	xxxxxxxxxxxxxxxxxxxxxxx	
5	5	xxx	

When all questionnaire responses have been entered, sort them so that the information will be most useful to you. For example, first sort on column 1 (*respondent ID number*) to put all the respondents in alphabetical order.

ID#	Q#	Response	Code
2	1	XX	
2	2	XXXXXXXXXXXXXXXXXXXX	
2	3	XXXXXXXXXX	
2	4	XXXXXXXXXXXXXXXXXXXX	
2	5	XX	

ID#	Q#	Response	Code
4	1	XX	
4	2	No response	
4	3	XXXXXXXXXXXX	
4	4	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	
4	5	XXXX	
5	1	XX	
5	2	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	
5	3	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	
5	4	XXXXXXXXXXXXXXXXXXXX	
5	5	XXX	

[Why don't you just sort 2 sets? All these seem unnecessary. – JAK]

Then sort on column 2 (*question number*) so that you have all the responses to each question together.

ID#	Q#	Response	Code
2	1	XX	
4	1	XX	
5	1	XX	
2	2	XXXXXXXXXXXXXXXXXXXX	
4	2	No response	
5	2	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	
2	3	XXXXXXXXXX	
4	3	XXXXXXXXXXXXXXXXXXXX	
5	3	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	
2	4	XXXXXXXXXXXXXXXXXXXX	
4	4	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	
5	4	XXXXXXXXXXXXXXXXXXXX	
2	5	XX	
4	5	XXXX	
5	5	XXX	

You now are ready to code and analyze the data question by question.

Developing your codes

Coding is a process for categorizing your data. Develop a set of codes using both codes that you predefine and ones that emerge from the data. Predefined codes are categories and themes that you expect to see based on your prior knowledge. For example, when inquiring about faculty experiences with the college's revised schedule, you may be aware of faculty concerns about time to cover course material adequately and of students' concerns about the availability of

courses. Create a two-column table for listing the categories you anticipate and the code you will use for that category.

Code	Description of Code
TC	Time to cover course material
CA	Course availability

Emergent codes are those that become apparent as you review the data. Identify these, add them to your table, and create a code for each. This type of coding allows for the emergence of crucial but previously uncovered issues and connections. Also be alert to specific terms that occur in the data and use these as codes when appropriate.

Coding your data

Closely review and code your data. If possible, have more than one person code the data to allow for different perspectives on the data.

As you proceed you may find that your initial codes are too broad. Create subcategories of your codes as needed. Or you may find that you have created codes that are too detailed and that attempt to capture every possible idea. In that case consider how you can pull categories together into a broader idea. For example, if you had one code for student concerns about courses not being available on certain days of the week; another for concerns about courses not being offered in the day; and another for courses not being offered in the evening, you would need to combine those into one category about time of course offerings.

Remember that you are coding in order to gain an understanding of the inquiry issue, how respondents perceive the issue under review, and the nature and types of relationships involved. Coding is a process of reducing the data into smaller groupings so they are more manageable. The process also helps you to begin to see relationships between these categories and patterns of interaction.

Continue to write analytic memos and hold analytic meetings as you proceed with the coding. Such memos and meetings provide an opportunity to reflect on your initial ideas and examine them in relation to new information, and pose ideas about developing relationships and patterns.

Finding themes, patterns, and relationships

Step back from the detailed work of coding your data and look for the themes, patterns, and relationships that are emerging across your data. Look for similarities and differences in different sets of data and see what different groups are saying.

At times, you may find yourself so lost in the data that you need to find a way to look at it from another perspective. One way to do that is to hold an analytic meeting with your CLIP

members and engage them in a discussion of the data that does not fit the emerging patterns.² Focus on the “outlying” data and look for explanations of why they do not fit the patterns. The following process is outlined as steps but people likely will move back and forth and begin looking for exceptions, contradictions, and surprises around a particular theme in the data.

- Develop a summary of the results that is relatively short but complete enough to contain the data that you are puzzled by. Do not draw conclusions. Send this summary out a week before the meeting.
- When the group meets, ask people to identify what usually or generally is seen in the data. List these. Then list an exception for each of the items that usually is seen. (usually ... but ...) Then ask people to explain how both can exist.
- Identify contradictions in the data. Use the phrase, “on the one hand... but on the other hand...” Encourage them to discuss anything that does not “fit.”
- Ask for surprises in the data. “I’d expected to see ... in the data but it wasn’t there.” “I didn’t expect to see ... in the data but”
- Ask if there are any remaining puzzles in the data and why.

Summarizing your data

After you have coded a set of data, such as transcripts of interviews with faculty or questionnaire responses, write a summary of what you are learning. Summarize questionnaires question-by-question to illustrate key themes in each question. Similarly, summarize the key themes that emerge across a set of interview transcripts. When available, include quotations that illustrate the themes.

With your data coded and summarized you are ready to look across the various summaries and synthesize your findings across multiple data sources.

² This process has been adapted from the work of Bob Williams. See <http://users.actrix.co.nz/bobwill> for additional information and resources.