Analysis of Qualitative Data[1]

Demola Akinyoade[2], PhD

---

[1] July 2013

[2] Faculty at the Peace and Conflict Studies Unit, Department of Political Science and International Studies, College of Social and Management Sciences, Afe Babalola University, Ado-Ekiti, Nigeria. +234 805 770 2787. demola.akinyoade@abuad.edu.ng

Abstract

Qualitative data analysis is a distinctive form of analysis in the social research enterprise. It is an approach that is less understood than its counterpart—quantitative analysis. Diversity and flexibility are main features of qualitative data analysis. These features also expose it to the danger of doing it anyhow—a slapdash analysis unbecoming of scientific endeavor. Despite its diversity there are common features to the analysis of qualitative data that beginning researchers or trainee-social scientists, such as undergraduates, should be familiar with. This is the focus of this paper. It focuses on necessary areas in data analysis to help this category of researchers to make sense of their qualitative data. It covers sources and types of qualitative data, basic issues and procedures in qualitative data analysis. It presents a systematic, disciplined, transparent and describable process to the analysis of qualitative data in consonance with the nature of the science and its method.

Introduction

Natural and social scientists alike rely heavily on and have well-developed methods for analyzing qualitative data. Qualitative data may be about people or things. Ecologists working with satellite images of the earth, political scientist analyzing presidential debate, astronomers listening to sound recordings from other galaxies, football analysts reviewing a tournament's goals, doctors listening to heart beats, economist or sociologists poring over transcripts of interviews on consumer behavior or single male/female sexual relationships are all looking for regularities in qualitative data. Looking for and accounting for regularities in qualitative data is what qualitative data analysis is all about in science. Qualitative data analysis is making sense of participants' description of the situation (being studied) noting patterns, themes, categories and regularities. It involves organizing, accounting for and explaining the data. (Cohen, Manion, & Morrison, 2011) The phrase qualitative data analysis could mean 'analysis of qualitative data' of 'qualitative analysis of data'; its usage in this chapter is in the sense of the former meaning—analysis of qualitative data.

To analyze qualitative data, one needs to understand the philosophy and meaning of qualitative data; how and probably why they are generated. Qualitative research (which essentially generates qualitative data) is concerned with understanding meaning from the perspective of the people that are being studied— the respondents or participants. It is rooted in the belief that we can only understand things (phenomena) from the interpretation given to them by individuals involved in those things, that is, from the participants' interpretation of those things. This is the philosophy or rationale behind how to know things in qualitative research. This is it epistemology—subjectivism or humanism. Much of qualitative data analysis is therefore the interpretation of participants' interpretation of their realities.

Qualitative Data

Social (and management) scientists are concerned with people's thoughts, emotions, behaviors, artifacts, and the environmental conditions in which people think, feel, behave and make things. Qualitative data in the social sciences is a way of recording people's attitude, feelings and behaviors in greater depth than simply counting (quantifying) them. It "…provides a rich, detailed picture to be built up about why people act in certain ways, and their feelings about these actions." (Learn Higher and MMU, 2008). We generate qualitative data in many ways including interviews, focus groups, observation, field notes, documents and reports, memos, diaries, emails and online conversations, website data, audio and video materials, advertisements and print media, pictures and photographs, artifacts and open ended items from questionnaires/surveys.

Interviews, focus groups and observation are most frequently used methods of data collection in qualitative research. Interviews are conversation between two or more individuals—the

interviewer(s) and the interviewee(s) where the former ask questions to obtain information from the latter. In focus groups, a group of (between 6-10 typically) people are asked about their views about or attitude towards things—product, service, concept, idea, amongst others—events or people. We do qualitative data analysis when we want to look further than precise numerical evidence; look for categories such as events, descriptions, comments, behavior; look for new or emerging categories; discover and account for regularities, differences, or patterns in data; and develop theories (or evidence-based explanations) from data collected.

From the list of ways of generating qualitative data identified above come five types of qualitative data including physical objects, pictures (still images), sounds, videos (moving images), and text (written words). Text is by far the most frequently used. Table 1 below shows the forms and accessibility qualitative data.

Table 1: Forms and accessibility of qualitative data

S/N	Form	Small		Large	
		Accessibility			
		Public	Private	Public	Private
1	Physical Objects	Park sculptures, street signs, pottery shards, store merchandise	Personal jewelry, pill bottles, blood samples	Archaeological ruins, buildings, houses, universities, skyscrapers	Household garbage, clothing
2	Still Images	Magazine ads, cave art, billboards, Web pages, paintings hung in museums	Doodles, line sketches, family portraits, patient X-rays	Large detailed murals, art exhibits	Family albums, art portfolios, CAT scans
3	Sounds	Jingles, radio ads, intercom announcements, messages you hear while on hold	Memo dictation, answering machine messages, elevator conversations	Political speeches, sports play-by-plays, music albums, focus group tape recordings	Oral histories, demo sound-tracks, in-depth conversations, clinical interviews
4	Moving Images: Video	TV ads, news footage, sitcoms	Home-movie clips	Full-length movies, documentaries, television programs	Long video recordings of family reunions and special events, like weddings
5	Texts	Epitaphs, obituaries, personal ads, political buttons, parking tickets	Thank-you letters, shopping lists, short responses to interview questions, e-mails	Books, manuals, religious tomes, court transcripts, Congressional Record, newspapers	Diaries, detailed correspondence, private chat-room discussions

Source: (Sage)

Key Issues in Qualitative Research and Data Analysis

When thinking of analyzing qualitative data, one must consider some critical issues regarding making qualitative analysis valid and reliable as a scientific endeavor. Scientific analysis must be rigorous and that means being systematic, disciplined, transparent and describable. Anything short of this will make the product of such analysis pseudo-scientific (false science) or unscientific. Establishing trustworthiness of qualitative data analysis requires efficacy check on data by ensuring that the interpretations are credible, transferable, dependable and confirmable. Credibility, that is, believability, in qualitative research is ensured through prolonged engagement with participants. This is hinged on the notion that participants are the only legitimate judge, hence the credibility must be established from their perspective. Credibility can also be established through triangulation in data collection. This means using different methods to collect same data. This is ensures internal validity.

Transferability means the degree to which the results of qualitative research can be generalized or transferred to other context or settings. We can achieve transferability by describing the research context, process and its central assumptions in great details to allow readers to see if the results can be applied to other settings (Trochim, 2006; Learn Higher and MMU, 2008). This is external validity in qualitative research. Dependability requires that the researcher account for the ever-changing context within which the research occurred and how these affected the she approached the study. Confirmability is the degree to which the results could be confirmed or corroborated by others. There are a few ways to enhance confirmability. First document the procedures for checking and rechecking the data throughout the study. Another way is by another researcher taking a "devil's advocate" role with respect to the results, and documenting this process. Yet another way is to actively search for instances that contradict prior observations (negative instances or cases). Finally, after the study, conduct a *data audit* that examines the data collection and analysis procedures and makes judgments about the potential for bias or distortion (Akinyoade, 2012).

Considering the nature of the relationship between data collection and analysis is also important in qualitative data analysis. Data collection and analysis in qualitative research tends to be iterative. That is, there is a back and forth movement between data collection and analysis. In other words the analysis of qualitative data begins immediately after collecting the data and interim analysis influence the next stage of data collection. This continues until the researcher understands the phenomenon being studied. This relationship is most remarkable in a particular qualitative strategy or method—grounded theory. Also considering the goal of the research will influence the data analysis process. Qualitative research has four main goals—exploration, description, comparison and testing models (Sage).

When the goal is exploration, the research aims at discovering themes and patterns to build early models of how complex systems work. Exploration involves following leads and hunches, taking a step forward and then backtracking, uncovering what is there, experiencing the phenomena

being studied and identifying its unique features and those features it shares with other phenomena. (Sage). Description is required in all scientific fields. When we list typical features, idiosyncrasies, and exceptions of phenomena (for instance in case studies in Law, medicine, management, or ethnographic field notes) we are describing it. When conducting a research that has the goal of description, it important to get as much data as possible. This gives more room for maneuvering later in the analyses. However, if we collect less data than required we may not be able (robust enough or adequate) to describe the phenomenon as required. Hence in describing cultural beliefs and practices, the researcher should focus on what people share and what they don't share. Here again, it pays to get as much as possible because one can only generalize from the specifics (Sage).

In qualitative comparison the research focuses on identifying features that phenomena (events, individuals or groups) share and don't share. For instance, collecting observational data on female-male interactions in the evening among students of private and public university is a comparative qualitative research. In testing models, we test hypotheses against observation. It is possible to test hypotheses in such qualitative research (Sage). We can start out with the hypothesis that restrictive rules as to female-male relationships among undergraduates will lead to less intense interactions among undergraduates of the opposite sex. We can then draw conclusion about the relationship between concepts (restrictive rules and intensity of interactions) relevant to the phenomenon—opposite sex evening interaction. Some projects will involve all or a combination of these four goals. These four goals, types of research and examples of their corresponding questions are shown in the Table 2 below.

Table 2: Goals of Qualitative Research

S/N	General Aim	Type	Questions
1	Exploration		What kinds of things are present here? How are these things related to one another? Are there natural groups of things here?
2	Description	Case	What does a case look like?
		Group	What does a set of cases look like? Is a particular kind of thing (A) present or not? How much of that kind of thing (A) is there?
		Cultural	What does the culture look like?
3	Comparison	Case	How is case X different from case Y?
		Group	How is a group of Xs different from a group of Ys?
4	Testing models	Case	To what degree does a particular case conform to the proposed model?
		Group	To what degree does a group of cases conform to the proposed model?

Source: (Sage)

General Activities in Analyzing Qualitative Data

There are common steps in analyzing qualitative research. Many times these steps will not follow one another as presented here. However, they represent general activities involved in the analysis of qualitative data.

Transcription

This is the representation of audible and visual data into written form is the first step in analyzing data. (Bailey, 2008) Data from interviews, observation are usually (not necessarily) transcribed into text. When transcribing an interview, the transcriber listens to the recorded interview and types this in text form. Usually, the transcribers plays and pauses the audio using proper gadgets (tape player, midgets or computer software) in order to type the interviews in text format. Computer audio-visual software packages such as Windows Media Player, VLC, and the likes could be used for transcribing. In addition, some computer packages for transcribing such as Express Scribe are available for transcribing.

There are different levels of details required for various research projects. The level chosen will be determined by the nature of the research. Transcribing is an interpretive act rather than a mere technical procedure. (Bailey, 2008). It involves paying the close observation that transcribing entails can lead to noticing unanticipated phenomena. It is impossible to represent the full complexity of human interaction on a transcript and so listening to and/or watching the 'original' recorded data brings data alive through appreciating the way that things have been said as well as what has been said.

Reviewing the data

This is the next step whether the data are transcribed or not. At this stage, you read and re-read the data (or transcript) with an eye for recurrent ideas, themes, categories, patterns, and relationships (InSites, 2007). You write *memos* to capture your initial thinking and tentative ideas about the data. Reviewing your data this way has some usefulness. First, you may realize that you are overlooking certain important areas in the interviews, which you then address in future interviews. Also, the interim themes and categories that you see in the data inform the *codes* that you will later use for a systematic analysis of your data. (InSites, 2007). Reviewing your data gives you an idea of what people are saying and the general trend of the results. However, there is still need for deeper, more systematic, and more rigorous analysis.

Organizing your data

Once you have reviewed your data, organizing your data becomes easier. You may achieve this by organizing using first, the data collection question on your interview guide, and then the research

questions. Typically, the questions on your interview guide or schedule will be drawn to gather data to answer the research questions (the original quest of your study). Sometimes two or more data collection questions (those on the interview guide or questionnaire) will be developed for each research question. Your data collection and research questions come in handy when organizing your data. This is because the sheer volume of data generated from qualitative research could answer an incredible number of questions. This could become so overwhelming. However, organizing your data using your original questions will help to keep your analysis focused and manageable. This is what O' Connor & Gibson, call *thinking inside the box*. So, in organizing your data for analysis, first bring together all participants' responses to a particular (data collection) question in a tables or summary sheets. Second bring together responses of all data collection questions for a particular research question.

For example, a research question in a research on the role of women in managing Ife-Modakeke Crisis reads 'What roles did Ife and Modakeke women play in conflict resolution and peace building in Ife-Modakeke Crisis?' Relevant data collection questions include 'what role did Ife women play in resolving Ife-Modakeke crisis?', 'what role did Modakeke women play in resolving Ife-Modakeke crisis?', what role did Ife women play in building peace after resolving the Ife-Modakeke crisis', and 'what role did Modakeke women play in building peace after the Ife-Modakeke crisis?' These four data collection question will collect the data necessary to answer the research question. At the stage of analysis, participants' responses to each of these four questions will be grouped together in tables or summary sheets. Then all these responses to each of the four questions will be grouped together as data necessary to answer the research question.

Organizing your data around the original research questions helps you to see your data at a glance. This will be useful later when you are looking for categories or themes. Also, it will help to quickly determine whether you need to seek further clarifications from the participants on that particular question. Because of its flexibility, participants may provide information useful to answer a question while responding to another question. Hence it is important for the data analyst to be open to these surprises in organizing the data during analysis. The idea of organizing data is to bring all relevant and useful data together in preparation for systematic and rigorous analysis later

Clarifying your data

This may be necessary after organizing your data. You may need to go back to seek clarifications on your discovery about the phenomenon being studied. However, this is a process done at any time required in the analysis of qualitative data.

Search and identify ideas, concepts, themes in the data

The analyst goes over the organized data group-by-group or section-by-section to identify salient themes, recurring ideas, words or language, and patterns of belief of participants about the phenomenon. This exercise, in the words of Marshall and Ross (1995, p 114, cited in O'Connor & Gibson) '…is the most intellectually challenging phase of the analysis and one that can integrate the entire [data analysis] endeavor.' It involves paying attention to words or phrases used frequently, ideas from how participants expressed themselves and stories they have told the interviewer. This typically will involve underlining or otherwise marking sections (word, phrase, a whole paragraph) of the data that you find relevant to answering the questions. This is where two of the most important activities in qualitative data analysis—coding and memoing—are most intense. We are now going to look at each of these one after the other.

Coding

This is the process of marking segments of data with symbols, descriptive words, or category names. It is therefore a process of categorizing the data into meaningful segments. A sort of coding is what we do when we mark text we are reading and write words or notes by the margins to describe their meanings. However, in qualitative data analysis, coding is not restricted to only text. We can code still pictures, sounds, moving pictures, especially when using computer applications to analyze our qualitative data. In coding our data, we use predefined codes or those that emerge from the data. Predefined codes are those ones we brought to the analysis based on our knowledge of the phenomena. For instance in the study of the Ife-Modakeke crisis above, codes such as conflict resolution, peacebuilding are already predefined in the research questions before the stage of data analysis. Ascribing these codes to instances of conflict resolution techniques used by women is using predefined codes. Emerging codes are those ones that will be suggested (from frequently used words, phrases or ideas) or become apparent from the responses of the participants. These are codes that the researcher did not think of before the analysis. They 'jumped out' of the data in the process of her close engagement with the data.

You assign a code or category identifier whenever you find a meaningful segment of your data. You do this for all your data until it is completely segmented. This is the initial coding, where you develop and define your (predefined and emergent) codes. It is important you keep a master list (that is, a list of all the codes you have developed and used in a particular analysis). This is automatically done for you if you are using computer application for analyzing qualitative data such as ATLAS.ti. Codes from this master list are reapplied to new segments of data each time the analyst encounters a relevant segment of data.

After the initial coding, you go over the data again for the second coding phase. This is important because by the time you finish going over the whole data at the initial coding phase, you have become so familiar with the data that you are in a position for better coding it. At the second coding phase, you may see reasons to collapse or redefine some codes or rename others. This is because some codes may too broad (needing subcategories), others may be too detailed (needing pulling together in categories). The whole essence is to have well segmented and categorized data. In this sense, therefore, '…coding is a process of reducing the data into smaller groupings so they are more manageable.' (InSites, 2007) This way, the analyst begins to see relationships between categories and patterns of interaction.

Table 3: Coding your data

S/N	Question/topic	Response	Code/category
	History of illness (emotional reaction)		
1	How did you feel when you first know you had TB?	I was <u>scared of what was going to happen to me . . . most people I knew who had TB died</u>. I was also scared of what others would think and what their reaction would be — especially my family. **Also, I am scared — I don't want to make anyone else sick, so I try not to go places where I might give it to other people.**	<u>Concern about not making healthy recovery</u> **Concern over other people's reactions**
2	How has having TB affected your life?	<u>My body has been feeling weak. I do not want to eat much and I have lost a lot of weight.</u> **I had to miss some work, and when my boss found out I had TB, he made me go on leave for a while.** *Also, it is difficult because my family— they have to be careful and avoid touching anything that I touch so that they don't get sick.*	<u>Physical effects</u> **Work/financial effects** *Feeling of being "infectious"*

Source: (O' Connor & Gibson, ?)

We code categories and subcategories identified by reading and reading the data. We code themes or topics, ideas or concepts, terms or phrases, keywords. There are different kinds of coding. Coding typically start with a summary of the data being examined. This is called **descriptive** coding because it is a summary description of what is in the data. (Learn Higher and MMU, 2008). Coding then moves beyond description into categorizing and analyzing. This is **analytic** or **theoretical** coding. Also, we can code in **flat** or **non- hierarchical** manner or in **tree** or **hierarchical** manner. The former is like a list in which there are no sub-code levels, the latter has branching arrangement of sub-codes like a tree. The tables below present examples of each.

Table 4: Flat or non-hierarchical coding

When asked to discuss 'Making use of the computers in the library', the responses included:	Or, when asked 'What kind of computer training have you received?' the responses included:
• European Computer Driving Licence • Library induction tutorial • Genealogy research	• College course • Computer programming • Word processing

• Holiday research • Shopping • Visits to London • Magazines • Books	• School • Introduction to computers • Self taught • Library courses • Online course • ECDL • No training

Source (Learn Higher and MMU, 2008)

Typically, sub-codes in a tree relate to their parents codes by being 'examples of...', or 'contexts for...' or 'causes of...' or 'settings for...' and so on.

Table 5: Tree or Hierarchical coding

For example, Making use of the computers in the library:	And for What kind of computer training have you received?:

• **Services and courses offered by the library.**
    • ECDL
        • Word processing
        • Introduction to databases
    • Library induction tutorial
• **Specific websites and internet based services.**
    • Genealogy research.
        • Family history centre
        • Census website
    • Holiday research
        • Expedia
        • First Choice website
    • Online shopping
• **Other services and resources used (non ICT)**
    • Community leaflets
    • Magazines
    • Book
	• **College course**
	    • Computer programming
	    • Word processing
	• **School**
	    • Introduction to computers
	• **Self taught**
	• **Library courses**
	    • Online course
	        • ECDL
	• **No training**

Source (Learn Higher and MMU, 2008)

We can also classify codes into **factual**, **axial** and **selective**. We do factual coding when we code data according to what each piece of data represents. For example, responses to question on women's role in peacebuilding could include examples of activities such as religious activities, demonstrations, and meetings. Axial coding involves building connections within categories, and between categories and sub-categories. This gives pictures of the relationships between the categories and sub-categories. For example, let's consider these questions: 'what is the ownership type of your university?' and 'describe the regulation of female-male evening interactions in the university? We can compare responses to these questions to see if there are any links between ownership type of university and regulations of female-male evening interactions. Selective coding involves identifying a picture of relationships within a core category by focusing on one aspect of a core category. This way we discover the relationship between it and related categories. From our previous example, we may focus

on the responses of students from private-owned universities and compare their responses to other questions in order to explore their experience more in-depth.

Memoing

This is another important activity that usually takes place throughout the entire process of qualitative data analysis. As you engage the data, some initial thoughts or ideas about the various aspects of the data and their relationships begin to occur to you. You may have to write them down in order not to forget the critical information. This is memoing and what you have written down about various segments of your data are memos. Memoing is the act of recording reflective notes about what you are learning from your data. The gist is to write memos to yourself when you have ideas and insights. These memos become additional data to be analyzed.

Memoing usually goes concurrently with coding. It may be about a word, phrase, segment of the data or about patterns, themes or relationships among concepts, ideas or themes about the subject of inquiry.

A summary of coding and memoing activities by Analyse This!!! by Learn Higher and MMU is useful here:

- Coding involves categorizing and indexing sections or chunks of your data
- Codes can come from theory and explanations 'outside the data' and/or 'emerge from the data'
- Data formats that can be coded ranges from transcribed text to video
- Coding often starts by being descriptive but needs to becomes analytical
- Any new codes created should be applied to the whole data set (previously coded units of data)
- Memos should be used to record your thoughts and ideas about your codes during the process

Finding themes, patterns, and relationships

Having done detailed work of coding your data and writing memos, you may need to look for the themes, patterns, and relationships that are emerging across your data. This involves looking for similarities and differences in different segments and sets of data and seeing what different groups of participants are saying. You will begin to see themes, patterns and relationships within and among categories and subcategories. You may need to focus on data that do not fit into the emerging themes, patterns and relationships. The process involves moving back and forth as the analyst look for exceptions, contradictions, and surprises around a particular theme in the data. You may realize that you can collapse different categories under one main over-arching theme. Also, you begin to see patterns of how themes are related.

Once you have established over-arching themes and their relationships, you may notice that some cases in the data do not comply with the themes. These are "outlying" data or negative cases.

Focus on the "outlying" data and look for explanations of why they do not fit the patterns. Your memos are part of the data to be analyzed at this stage. You may even need to write more memos. These activities are the culmination of qualitative data analysis. It could be overwhelming at times.

Reporting your qualitative analysis

Discussion

After you have done some analysis and you have found interesting themes, patterns and relationships in your data, you may have to start writing your discussion of findings. Qualitative research usually has rich discussion. You embellish your discussion with quoting participants verbatim to illustrate your points.

Conclusion

This is a summation of your analysis and discussions.

Recommendations

At the end of your study, you may want to make recommendations based on the findings of your study or suggest where extra data collection and analysis activities are required.

Conclusion

Analyzing qualitative data involves a whole set of activities with its peculiar epistemological orientation, internal logic and set of process that make it systematic, disciplined, transparent and describable in consonance with the scientific method. It involves key issues such as credibility, transferability, dependability and confirmability. The general steps in qualitative data analysis include: transcribing, reviewing the data, organizing the data, clarifying you data, coding, memoing, finding themes, patterns and relationships, and report writing.

Bibliography

Akinyoade, D. (2012, April 4). introduction to designing qualitative research. Ile-Ife, Osun, Nigeria.

Bailey, J. (2008, February 27). First steps in qualitative data analysis: transcribing. *Family Practice ,
25* (2), 127-131.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education* (7th ed.). Oxon:
Routledge.

InSites. (2007, May). (Module 5)Tips for analyzing qualitative data.

Learn Higher and MMU. (2008). *Analyse This!!! Learn to analyse qualitative data*. Retrieved March
21, 2013, from Analyse This: http://www.learninghigher.ac.uk/analysethis

O' Connor, H., & Gibson, N. (?). A Step-by-Step Guide to Qualitative Data Analysis. *Pimatiziwin: A
Journal of Aboriginal and Indigenous Community Health, 1*(1), 63-90.

Punch, K. F. (1998). *Introduction to social research: quantitative and qualitative approaches.*
London: Sage Publication Ltd.

Sage. (n.d.). Introduction to Tex: Qualitative data analysis.

Seidel, J. V. (1998). *Qualitative Data Analysis.* Retrieved April 4, 2012, from Qualis Research:
http://www.qualiresearch.com

Trochim, W. M. (2006, October 20). *Qualitative Methods*. Retrieved March 22, 2012, from Research
Methods Knowledge Base: http://www.socialresearchmethods.net/kb/