

Unit 27

Qualitative Data Formatting and Processing

Contents

- 27.1 Introduction
- 27.2 Qualitative Data Processing and Analysis
- 27.3 Description
- 27.4 Classification
- 27.5 Making Connections
- 27.6 Theoretical Coding
- 27.7 Qualitative Content Analysis
- 27.8 Conclusion



Learning Objectives

It is expected that after reading Unit 27, you would be able to proceed with data formatting and processing of the field material gathered by you on the topic of your mini research project in terms of the Circular Process of Qualitative Analysis as formulated by Ian Dey (1993).

27.1 Introduction

The previous units of this block have familiarised you with some of the key methods of gathering data employed by qualitative researchers. The process of discovery and learning embarked upon whilst gathering data is no doubt an exciting and memorable one. However, there comes a point when field notes and interview pages gathered by you must be formatted, processed and analysed. Some of the questions that confront the researcher at this point are: how can I make sense of all this material? How can I organise it to make it meaningful to others and myself? How do I put it all together so as to present a concise and thoughtful formulation of the topic under study?

Hardly any researcher would initiate the process of investigation without a at least a mental checklist of the type of material that is to be gathered. A systematic listing of the topics to be covered is always very useful. Ellen (1984: 275) has mentioned the following three forms of checklists.

- ❖ Checklists encoded in questionnaire forms
- ❖ Checklists employed in the context of semi-informal interviews
- ❖ Background checklists for occasional reference and to provide guidelines for research in general

The idea of mentioning these checklists in the introduction of Unit 27 is to remind you that in order to format and process your field material you need to refer to these checklists to find out if you have actually

completed what you had set out to find out. Only after this initial exercise are you to proceed with qualitative data processing and analysis.

27.2 Qualitative Data Processing and Analysis

The key word that this process entails is analysis. Qualitative analysis requires dialectic[®] between ideas and qualitative data. You may ask: What are qualitative data? For an answer to this question see Box 27.1.

Box 27.1 Definition of Qualitative Data

Qualitative data are materials gathered using field research methods already discussed in Unit 26. Openness and inclusiveness characterise these methods. A researcher applies such methods with the objective of capturing people's lived experiences of the world and the meanings they attach to these experiences from their own worldviews. More often than not collection of qualitative data entails a variety of methods and techniques rather than a single one. The result is that data types include in-depth or unstructured interviews, field notes, unstructured field diaries, personal documents, photographs and so on. Qualitative research in its initial stages means producing a large mass of data even though a researcher may have used a relatively small sample size.

We cannot analyse our qualitative data without ideas, but our ideas must be shaped and tested by the data we are analysing. You cannot make an omelette without breaking and then beating together eggs. Analysis also involves breaking down data into bits and then beating the bits together, so that the data is resolved into its components, and its characteristic elements are revealed. In this sense you can say that we split data processing into two activities, namely,

- ❖ Checking and converting the data (breaking eggs) and
- ❖ Generating metadata (beating together).

The first activity comprises a) checking out the completeness and quality of data, the relationship between data (for example interviews, field notes, audio/ visual recordings etc) and anonymisation[®] and b) converting data or transferring data to a format that is appropriate for dissemination. At this time apart from checking out the completeness and quality (in terms of its physical condition, readability/ audibility, re-usability) a researcher needs to bring to light any problems relating to confidentiality/ anonymity, re-using, suitability for digitisation, etc.

We have mentioned above the term anonymisation. Let us explain what this term means. It refers to maintaining the confidentiality of the respondent or any other person or entity. It is important to discuss the level of anonymisation you arrive at in your research. In some cases, it is not easy to disguise the identity of the subjects of research without bringing about an unacceptable distortion to the data. This implies that the particular data can hardly be re-used for any purpose. The level of anonymisation you use for a dataset depends on the nature of the study and each case has its own unique set of concerns. Here come in issues of

an ethical and legal nature with respect of maintaining confidentiality where requested. In practical terms, all this means that you remove all identifiers and replace them with pseudonyms where appropriate. Care should be taken to use the same pseudonyms throughout. It is strongly advised to delete from the data any slanderous or libellous comments.

The second processing activity of generating metadata refers to the contextual information that a researcher obtains during processing, for example you may create lists of data giving biographical inputs that would make it easier to identify transcripts or make sure that interviewee and interviewer names and questions/ topic guide headers are put in place. The main function of this exercise is to make it easier for the researcher to locate transcripts or particular items in a data set (see below Section 27.6 on theoretical coding). You will find a definition of metadata[®] types in Box 27.2. Further, please note that about digitisation of data you will read in detail in the units of Block 8 and we will not include this factor in the scope of Unit 27.

Box 27.2 Definition of Metadata types

The researcher has to wade through the sea of raw data. In order to make the process of wading easier, it is better to prepare i) a data list and ii) a catalogue of records. In addition, one may also prepare iii) a user guide so that anyone can use the data. These three types of data are called metadata. It is always a good research practice to document one's research and the material produced during field research from its beginning and throughout analysis. The data list provides the key characteristics of the data. It helps the researcher to identify specific types of interviews or transcripts. For example, in the case of data which are based on a sample of interviews with individuals, metadata listing would include their date of birth, gender, employment, geographical location and any other key feature that the researcher has defined for sampling.

How do we carry out the above two processes without getting into a mechanical mode of processing? We need to proceed to analysing our data. The aim of analysis is not just description, but interpretation, explanation, understanding, and possibly, prediction. Description provides the basis for analysis, which then lays the base for further description. We break down our data in order to draw concepts from it, which we then use to classify the data. We draw connections between different concepts and these connections form the basis of further description. Ian Dey (1993) presents the following diagram, which succinctly represents the circular processes involved in qualitative analysis.

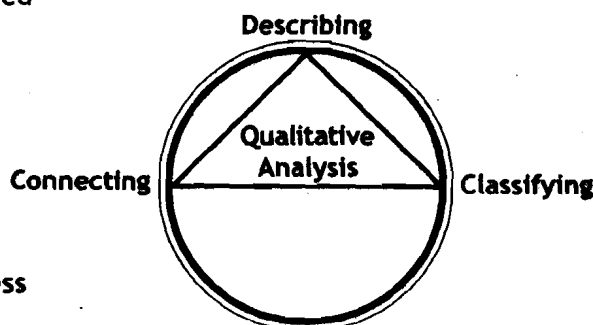


Figure 27.1 Circular Process of Qualitative Analysis

Let us briefly consider each of the processes mentioned on the previous page.

27.3 Description

The first step in qualitative analysis is to develop comprehensive descriptions of the phenomenon under study. Geertz (1973b) coined the term “thick description”. Unlike “thin description”, which merely states facts, thick description includes information about a) the context of an act, b) intentions and meaning attributed by the actor to the action and c) the process in which the action is embedded.

Let us briefly dwell on these points.

- ❖ **Context:** By “context” we mean situating the action within the social, cultural and historical backdrop against which it takes place. Contexts are a key to meaning since meaning is conveyed “correctly” only if the context is also understood.
- ❖ **Intentions:** In qualitative analysis there is a strong emphasis in understanding the subjective meanings imbued by actors to the way in which action is organised.
- ❖ **Process:** The idea of process is linked to that of change. In focusing on process, we shift our attention from context and intentions to the consequences of the action. For example, you are probably familiar with the policy of “perestroika”[®] unleashed by Mikhael Gorbachev in the erstwhile Soviet Union in the mid-1980s. The context was growing disenchantment with and ineffectiveness of the state controls on economy and society. The intentions of Gorbachev were probably the creation of a more just, liberal order. But the consequences were the disintegration of the Soviet Union and the movement towards a unipolar world order governed by the U.S.A.

27.4 Classification

This involves sorting out our mass of data into “classes” based on certain characteristics, which will then aid us in the development of a conceptual framework through which actions and events can be rendered intelligible. You have surely done jigsaw puzzles. Consider your data as hundreds of little bits of the jigsaw, which must be carefully put together so that the resultant picture gives an accurate representation of the social reality you have studied. How do you assemble the picture? By grouping the bits into like classes. Perhaps all the “blue bits” together will form the sky, the “green bits” the forest, and the brown ones the earth. Organising the data into groups based on certain characteristics is essentially the process of categorisation. At the stage of writing up your study the same categories may take the form of index of your book (for indexing see Box 27.3). Categories are the organising tools, which help us to sort out the heap of bits of the jigsaw puzzle according to relevant

characteristics. Classification and categorisations should always be guided by the research objectives. The points mentioned above will be taken up in greater detail when we talk about “coding”[®].

Box 27.3 Indexing

There are three elements found in an index, i) mnemonics, ii) reference and iii) structure. Mnemonics are short forms for information. References show where you are to find the information and the structure is the arrangement for organising mnemonics. Most often you may find that categories in mnemonics overlap. In such a situation ideally you would need to put the overlapping pieces of information in both categories. With regard to references you would require a fair degree of precision for identifying the location of information. You may take recourse to modern information technology (see the units in Block 8). The structure of mnemonics is usually in alphabetical order with substructures. Many categories mention cross-references to take care of overlapping information. Making such indexes complete in the above three respects is of course useful at the stage of data analysis and of writing up the research material. It is also useful if you or someone else was to re-study the same topic in the same region.

Let us now discuss briefly about making connections, the third process of Ian Dey’s diagram (Figure 27.1).

27.5 Making Connections

Describing and classifying are not ends in themselves but serve a more important purpose, namely, to produce an account of our analysis. The concepts we develop are like building blocks, which must be connected together with the mortar of ideas. We have to look for associations between different variables and try to see the patterns within the data, so that we can discern regularities and also variation and exceptions.

We can now discuss the subject of “coding”. But before proceeding to this discussion let us first complete Reflection and Action 27.1.

Reflection and Action 27.1

Read the following excerpt and work out what kinds of data does one require to understand the rituals in their social context and to grasp their significance for the practitioners of those rituals?

Religious behaviour does not exclusively depend on religious contexts but it is generally a human form of behaviour which is realised under the stimulus not of transcendental objects but of their motivations... . Even in its autonomy, religious life contains elements that are not specifically religious but social... only when (these elements) are isolated by means of the sociological method, will they show what within the whole complex of religious behaviour may legitimately be considered purely religious, that is, independent of anything social (Simmel 1950: 15).

You can take help of Victor Turner’s (1967: 181-204) article on “Aspects of Saora Ritual and Shamanism” for answering the questions raised above. Here, Turner has shown the kind of data required and the connections that need to be made in order to understand and explain the “phenomena of mysticism, asceticism, conversion and holy mendicancy in the higher religions”.

27.6 Theoretical Coding

In order to analyse our data, we have to read it in an interactive way, and constantly ask the questions “Who”? “What”? “When”? “Where”? and “Why”? This will open up our data for us and help us think about it in a creative way. The processing of field material, once it has been gathered, invariably implies hard work for researchers who are inexperienced in handling field data. Many information and communication technology aids are now available for them and we have discussed them in Block 8. Here we will be mainly concerned with the theoretical coding of field material.

The key function of this exercise is to convert the material on the schedules into suitable code form. Coding is considered a tedious and demanding job. Further, this is an area where methods of analysis of quantitative and qualitative data complement and supplement each other. For example, Turner (1957 and 1961) has used the quantitative material about village composition in his description of Ndembu[®] social structure.

Let us now look at some techniques and procedures, which will help make sense of qualitative data. The first procedure discussed is “theoretical coding” in order to develop a “grounded theory”[®]. This procedure was introduced by Glaser and Strauss (1967), and further elaborated by Glaser (1978), Strauss (1987) and Strauss and Corbin (1990). Coding represents the operations by which data are broken down, conceptualised and put back together in new ways. A grounded theory is, “a rich, tightly woven, explanatory theory that closely approximates the reality it represents” (Strauss and Corbin 1990: 57). Analysis in grounded theory comprises three major kinds of coding, viz.

- ❖ Open coding
- ❖ Axial coding and
- ❖ Selective coding.

Let us briefly review each of them. But before you get into details of coding in grounded theory, it is a good idea to read in Box 27.4 a bit more about grounded theory.

Box 27.4 Grounded Theory According to Glaser

The following is a discussion on Grounded Theory (GT) according to Glaser (for more details on GT see glossary that is placed at the end of Book 3 of MSO 002).

Goals of grounded theory

The goal of a GT is to formulate hypotheses based on conceptual ideas that others may try to verify. The hypotheses are generated by constantly comparing conceptualised data on different levels of abstraction, and these comparisons contain deductive steps. GT does not aim for the “truth” but to conceptualise “what’s going on” using empirical data. GT is thus a systematic generation of theory from data that contains both inductive and deductive thinking. In a way GT resembles what many researchers do when retrospectively formulating new hypotheses to fit data. However, in GT the researcher does not pretend to have

formulated the hypotheses in advance since preformed hypotheses are prohibited (Glaser & Strauss 1967).

In most research endeavors persons or patients are units of analysis, whereas in GT the unit of analysis is the incident (Glaser & Strauss 1967). The number often amounts to several hundred in a GT study since every participant normally reports many incidents. When comparing many incidents in a certain area, the emerging concepts and their relationships are in reality probability statements. Consequently, GT is not a qualitative method but a general method that can use any kind of data, according to Glaser (2001). However, although working with probabilities, most GT studies are considered as qualitative since statistical methods are not used, and figures not presented. The results of GT is hence not reporting of facts but probability statements about the relationship between concepts, or an integrated set of conceptual hypotheses developed from empirical data (Glaser 1998). Validity in its traditional sense is consequently not an issue in GT, which instead should be judged by fit, relevance, workability, and modifiability (Glaser & Strauss 1967, Glaser 1978, Glaser 1998).

Fit has to do with how closely concepts fit with the incidents they are representing, and this is related to how thoroughly the constant comparison of incidents to concepts was done.

Relevance. A relevant study deals with the real concern of participants, evokes "grab" (captures the attention) and is not only of academic interest.

Workability. The theory works when it explains how the problem is being solved with much variation.

Modifiability. A modifiable theory can be altered when new relevant data is compared to existing data. A GT is never right or wrong, it just has more or less fit, relevance, workability and modifiability, and readers of Paper V are asked to judge its quality according to these principles.

The goal of a GT is to discover the participants' main concern and how they continually try to resolve it. The questions you keep on asking in GT are "What's going on?" and "What is the main problem of the participants and how are they trying to solve it?" These questions will be answered by the core variable and its sub-cores and properties in due course. If your research goal is accurate description then another method should be chosen since GT is not a descriptive method. Instead it has the goal of generating concepts that explain people's actions regardless of time and place. The descriptive parts of a GT are there mainly to illustrate the concepts.

We can now go back to our discussion of the three kinds of coding in grounded theory.

❖ Open coding

Open coding refers to close examination of the data so that phenomena may be named and categorised. An observation, a sentence, a paragraph from an interview transcript is taken apart and given a name which stands for or represents the phenomenon. We ask questions, like "What is this?" "What does it represent?" On the way, we make comparisons so that similar phenomena may be given the same name. Suppose, you are doing a study of how children play together. You see one child pulling away a toy from another and you label it as "grabbing". You may then observe another child "hiding" her toy, a third "avoiding" interacting

with the others in order to protect his toy. Now you can go on endlessly labelling, but this will get you nowhere. You have to start grouping your concepts together as categories. You may ask yourself what the “grabbing”, “hiding” and “avoiding” represent, and come to the conclusion that these are all “strategies to avoid sharing a toy”. After generating categories, Strauss and Corbin (1990: 70) recommend that their properties be identified and then dimensionalised, i.e. located along a continuum in order to define the content more precisely. Let us look at the category of “colour”. Its properties include shade, intensity, hue and so forth. Each of these properties can be dimensionalised; that is, they vary along the continuum. This colour can vary in intensity from high to low; in hue from darker to lighter; and so forth.

There are various ways of doing open coding. Strauss and Corbin recommend analysing the first interview, line-by-line, so that concepts and categories are freely generated. Subsequently, this can be done paragraph-wise or in terms of an entire document or case. It is important not to lose touch with the aims of coding, namely, breaking down and understanding a text in order to generate categories, which can be used for comparing. The result of open coding should be a list of codes and categories written alongside the text itself, along with “code notes” that explain the content of the codes. “Memos” which contain observations on the material and your thoughts about it also go a long way in developing grounded theory.

❖ Axial coding

The next step is to refine and differentiate the categories generated in open coding. Those categories are selected which hold out promise for further development. Strauss and Corbin (1990: .99) suggest a coding paradigm, which is given in the figure below:

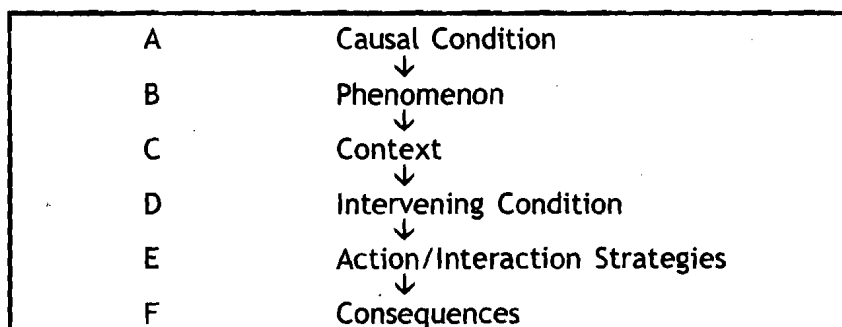


Figure 27.2 Coding Paradigm as per Strauss and Corbin (1999: 99)

In axial coding, categories are developed in terms of the causal conditions that give rise to the phenomenon, location of the phenomenon in terms of its properties, the context, the action/ interactional strategies used to handle, manage, respond to the phenomenon with regard to the context and the consequences of any action/ interaction that is taken.

In axial coding, the categories that are most relevant to the research question are selected from the developed codes and the related code notes. Many different passages in the text are then sought as evidence of these relevant codes.

❖ Selective coding

The third step, selective coding, continues the axial coding at a higher level of abstraction. It aims at laying bare the core category around which the other categories can be arranged. In other words, it reveals the “story of the case”. The story of the case is to be set down briefly, before developing the story line (see Box 27.5 on identifying a story). You will note that the coding approach is essentially an inductive one. From studying the content and meaning of the text, the researcher uses his/her interpretative abilities to formulate code and categories, takes them to a higher level of abstraction and then constructs a story or an account, which is applicable to the whole data. The researcher is able to say, “under these conditions, this is what occurs” (Strauss and Corbin 1990: 131).

Box 27.5 Identifying the Story (from Strauss & Corbin 1990: 119-120)

The following example has been taken from Strauss and Corbin (1990), using data collected by Corbin. Corbin studied how 20 women with chronic illnesses managed their pregnancies. The women she studied had illnesses like diabetes, heart disease, kidney disease and lupus erythematosus[®]. Whilst interviewing these women, Corbin came to realise what an active role they played in managing their high-risk pregnancies. The “story” she identified is as follows.

“The main story seems to be about how women with pregnancies complicated by chronic illness manage the risks they perceive to be associated with their pregnancies. Each pregnancy/ illness can be said to be on-course, indicating that the risks are being managed, or off-course, indicating that they are not. Women are managing the perceived risks in order to have a healthy baby. This desired outcome seems to be the primary force motivating them to do whatever is necessary to minimise the risks. However, they are not passive recipients of care but play a very important role in the management process. They not only are responsible for monitoring their illnesses and pregnancies at home, but also make very active decisions about the regimens they are told to follow. In the latter case they consider the harm that might come to the baby from procedures like amniocentesis or from taking high doses of certain medications while pregnant. They carefully weigh the risks and make judgments about the right thing to do. If they think the doctor is wrong, then they do what they (the women) think should be done.”

Let us now move on to another technique of formatting and processing data, namely, qualitative content analysis[®]. Before reading about qualitative content analysis let us complete Reflection and Action 27.2.

Reflection and Action 27.2

Glaser and Strauss’s (1967) grounded theory supports a generalist approach to social research and does not like any preconceptions about theoretical formulation of research. Such an approach has a clear process of coding. You need to read Section 27.6 carefully and provide examples of the three types of coding from your own mini research project.

27.7 Qualitative Content Analysis

Content analysis is one of the classical procedures for analysing textual

material, be it interviews or media products. It differs from coding in the sense that rather than generating categories from the data themselves, it uses categories brought from the outside and “fits” the empirical material against them, of course, modifying the categories wherever necessary. The aim of qualitative content analysis is to reduce the material so as to make it manageable. The following material drawn from Clive Flick’s (1998) description of the techniques of qualitative content analysis is based on the work of Peter Mayring (1983) on the experience of teachers regarding teaching.

- ❖ **Summarising content analysis:** Here, the material is paraphrased, that is, omitting less relevant passages or those having the same meaning reduces it (first reduction). Similar paraphrases are summarised together (second reduction). Thus the material becomes more coherent, economical and manageable. It is a kind of editing of the material in order to draw out its essence.
- ❖ **Explicative content analysis:** This works in the opposite way. Here, statements which are puzzling, contradictory or unclear, are explained or clarified by keeping in mind their context, or by looking for clues in other parts of the text that would help make their meaning clear.
- ❖ **Structuring content analysis:** Here, the researcher looks for types or formal structures in the material by applying categories that emerged at the stage of formulating the research question itself, and organising the material accordingly. Thus, qualitative content analysis helps the researcher to reduce to manageable level large masses of text, using a uniform scheme of categories, which also helps in the comparison of different cases to which it is applied throughout.

See Box 27.6 for examples of qualitative analysis.

Box 27.6 Examples of Qualitative Content Analysis (from Flick 1998: 194)

The following examples are quoted in Flick (1998) and pertain to data gathered by Peter Mayring (1983) on the experiences of teachers regarding teaching practice.

Example: Summarising content analysis

From an interview with an unemployed teacher, the statement ‘and actually, quite the reverse, I was well very-very keen on finally teaching for the first time’ (Mayring 1983: 104) is paraphrased as ‘quite the reverse, very keen on practice’ and generalised as ‘rather looking forward to practice’ (1983: 59). The statement ‘therefore, I have already waited for it, to go to the seminar school, until I finally could teach there for the first time’ (1983: 104) is paraphrased as ‘waited to teach finally’ and generalised as ‘looking forward to practice’. Owing to the similarity of the two generalisations, the second one then is skipped and reduced with the other statement to ‘practice not experienced as shock but as big fun’ (1983: 59). Thus, skipping those statements that overlap at the levels of the generalisation reduces the source text.

Example: Explicative content analysis

In an interview, a teacher expresses her difficulties in teaching by stating that

she – unlike successful colleagues.—.was ‘no’ entertainer type (1983: 109). In order to find out what she wishes to express by using this concept, first the varied definitions of ‘entertainer’ are assembled from two dictionaries. Then the features of the teacher who fit this description are sought from statements made by the teacher in the interview. Further passage is consulted. Based on the description of such colleagues included in these passage, an ‘explicating paraphrase can be formulated: an entertainer type is somebody who play the part of an extroverted, spirited, sparkling and self-assured human being’ (1983: 74). This explication is assessed again by applying it to the direct context in which the concept was used.

Before we end Unit 27, we need to complete the last Reflection and Action exercise in order to check if you have fully grasped the idea of content analysis.

Reflection and Action 27.3

According to Sarantakos (1998: 280-81), content analysis entails similar steps as any other method of research. This means that it includes selecting the area of research, designing the research, gathering data and analysing them. The content of each step separates content analysis from other methods of research. Content analysis analyses the content (it may be qualitative or quantitative) of documents, books, journals, and other kinds of written text. The content in such an analysis may be explicit or implicit. An example of content analysis is the study of television serials to find out why certain categories of people continue to watch particular shows without missing a single episode. Can you give at least five more examples of studies which can use the method of content analysis to discover attitudes, motives and values of subjects of research?

27.8 Conclusion

Unit 27 has sought to acquaint you with some important techniques of analysing and interpreting your data. However rich and interesting your data, they will make sense only if they are rigorously analysed and cogently presented. The approach of Strauss and Corbian exemplifies how we can proceed step by step to integrating masses of data into a tightly woven grounded theory which will help to explain and predict reality. Qualitative content analysis demonstrates how pre-existing categories can be imposed upon data in order to select and edit what seems irrelevant and unnecessary so that the key research questions may be addressed and understood.

Further Reading

Silverman, David. 1993. *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*. Sage Publications: New Delhi (for various dimensions and logic of qualitative research and for techniques of analysis of texts and interview data)

Singleton, Royce A. and Bruce C. Straits 1999. *Approaches to Social Research*. Oxford University Press: New York (about the set of questions a researcher has the interest in finding answers of)

