



# 10

## Analyzing qualitative data

Introduction	100
Coding and categorizing	100
Analysis of conversations, discourse and documents	104

### Chapter objectives

After reading this chapter, you should

- have a first orientation about the major approaches in analysing qualitative data;
- see the specific issues concerning research designs in using these methods; and
- have some orientation about when to use which methods of analysis in qualitative research.

### Introduction

The methods briefly presented in the preceding chapters are focused on collecting data (like interviews or focus groups). Or they consist of a more integrated approach (a link between a specific sort of data and a way of analyzing it – visual data or discourse analysis, for example). In this penultimate chapter, we will turn to approaches in which the analysis of qualitative data is in the foreground. Here we will find two different ways – coding and categorizing, which can be applied to qualitative data in general, whereas later a more specified analysis of data will be presented. At the end of this chapter we will try to bring the different issue of this book – the design issues and the methods – together for some kind of outlook.

### Coding and categorizing

Coding and categorizing are ways of analyzing that can be applied to all sorts of data and are not focused on a specific method of data collection. This is not the

only way of analyzing data, but it is the most prominent one, if the data result from interviews, focus groups or observations. In particular, if computers are used for analyzing qualitative data, one or the other form of coding will be applied. The main activities are to search for relevant parts of the data and to analyze them by comparing them with other data and by naming and classifying them. Through this process, a structure in the data is developed as a step towards a comprehensive understanding of the issue, the field, and last but not least the data themselves. It is suggested to interlink data collection with their analysis in order to profit from insights from the analysis for the process of collecting further data.

### ***Research perspective and theory***

The background of coding and categorizing can be realist or constructivist (Gibbs, 2007). The aim is often to develop a theory and therefore the categories for coding the material are developed from the material rather than from existing theories, although this way is possible and usual as well. The approaches need materials at hand – existing documents, transcriptions of interviews, focus groups or interactions. They can also be applied to visual material or Internet material.

### ***Research questions***

These approaches are open to all sorts of research topics. They are less appropriate where formal structures (like organizations of specific forms of talk as in conversation analysis) are the main issue. Although narratives can be analyzed with coding and categories as well (see Gibbs, 2007, chap. 5), an analysis interested in the internal structure and the gestalt of the narrative is more difficult with coding.

### ***Sampling***

When this form of analysis starts, sampling of cases and materials is often already completed. However, sampling in qualitative research can and maybe even should be based on the progress of the analysis of the data collected so far. Therefore, this form of analysis can have an impact on the sampling of cases and materials. During the analysis, sampling in the case and in the material is an important element.

### ***Comparison***

Here we can plan comparisons on three levels: Within a category – what do we find in different interviews, for example, as relevant for a specific category?

Within a case – what does the interviewee say about different issues, how consistent or contradictory are statements across several categories? Between cases – how different or similar are the responses of various interviewees on the level of one topic/category or on the level of the whole interview? Comparisons on the basis of focus groups or observations can be made on similar levels. Comparison here can be facilitated by ordering and structuring – like building a hierarchy between categories, by developing tables for displaying several cases and categories allowing case-by-case or chronological comparisons (see Gibbs, 2007).

### ***Generalization***

Constant comparison of materials is an important step on the way from the single case to more (or less) general statements drawn from analyzing the data. To avoid over-generalization, researchers should carefully reflect the boundaries of their data and the sample of people (or materials) they are based on.

### ***Triangulation***

Coding and categorizing can be combined with quantitative analyses of standardized data. It may also refer to different sorts and sources of qualitative data. Also, Gibbs (2007) argues for triangulating the perspectives of different researchers on the data as well as seeking for respondent validation – to integrate participants' perspectives on the data into the final analysis.

### ***Quality***

Central for the quality of data analysis, according to Gibbs (2007), is that the researchers are reflexive in their practice with the data by critically assessing their own roles as researchers as well as the data and the findings and conclusions drawn from them. Reliability can be increased by rechecking the transcripts and by cross-checking the codes.

### ***Writing***

In this context, writing is not only referring to presenting results in a report to audiences. All sorts of research materials like memos, protocols, field notes and research diaries can become relevant for the analysis and are products of the researchers' writing processes. In reports about the research, a good relation between categories that are presented, analysis and conclusions based on them, and excerpts from the 'raw' material is an important goal.

### **Basic design**

In most cases, coding and categorizing will be based on a comparative design. It can be applied in retrospective or longitudinal design as well as in snapshots, because this is decided more by the sort of data (and their focus and collection) that is used for the study.

### **Resources and stepping-stones**

Analyzing qualitative data in most cases is a time-consuming step in the research, which needs most of the resources of a project. A good transcription will make this step easier to handle, but you will need to spend time or money on this step, if it is to be a good basis for interpretation. Using computers can be helpful for supporting qualitative analysis in two cases in particular: if you are already experienced with the software you want to use and if you have a bigger corpus of data. A major problem to avoid is to have an unclear research question and not to have focused it again when you begin your analysis. Another problem may arise when you find out that your data actually do not include those statements or issues you discover as important for your analysis in later stages.

### **Ethics**

In the analysis of qualitative data, anonymity and confidentiality are central issues from the angle of ethics – in transcription, in analysis itself, and most of all in presenting results and excerpts from the data. For example, researchers should be sure that those who do the transcription keep up standards of confidentiality. If you feed your results back to participants, you should take care that they do not feel embarrassed or hurt by your interpretation, without giving them a chance to reflect with you on what was disturbing them. Interpretations should be fair to the research participants (and maybe the institutions) in the research.

### **Example**

In our project on homeless adolescents, we used thematic coding (Flick, 2006) for data analysis, beginning with all statements referring to an area (e.g. meaning of health) for every interviewee. Comparative dimensions are defined across cases for finding commonalities and differences between the various interviews. Cases are grouped along these dimensions and analyzed for specific combinations of features. Contrasting cases allow comparing the cases in one group for similarities and comparing cases across the groups for existing differences among them. Interpretive and practice patterns can be found and analyzed in this way. Stepping-stones are missing statements by one or more participants about specific issues.

### ***Conclusion***

Using codes and categories for analyzing qualitative data is quite flexible in the sort of data to be used and how the study is designed. It is mostly planned with the aim of comparing data and cases. Triangulation should include different researchers working on the same material for extending the perspectives. Presentations should take the relation between original material and categories or dimensions into account so that the analysis becomes transparent.

## **Analysis of conversations, discourse and documents**

In these forms of qualitative analysis, data collection is often limited to selecting material (like documents or newspaper articles) and to documenting everyday routines (like tape-recording a conversation between doctor and patient). Here, research interests often focus more on formal aspects – how is a conversation started, continued and ended, what are the structural aspects of a document like a patient record – than on content. In discourse analysis, the stress is more on content, however, compared to conversation analysis. In discourse analysis we also find the use of interviews, observations and focus groups as sources for materials to analyze.

### ***Research perspective and theory***

The theoretical background is in most cases ethnomethodology and discourse theory and the interest is in how communication and practices are constructed in everyday life in concrete circumstances. Therefore, the single actor is less focused than interaction processes. On the epistemological level, these approaches are based on social constructionism and theories of (written or spoken) language use in practical contexts.

### ***Research questions***

The primary research question of these kinds of research is: how is a specific issue constructed in some sort of communication and which 'methods' do participants in this communication use for this construction? Thus, documents are always produced by someone and for some audience, for some purpose and by using some communicative devices in order to create a certain format of information. This can be a specific way of communicating with a patient about her illness in a doctor–patient communication, which aims at having a diagnosis, a treatment and a plan for proceeding further at the end of the session. This can also be a record produced about this patient, her history as a patient, diagnosis

treatments and prognosis or a public discussion about this specific illness in the newspaper with the purpose of creating a moral discourse about the disease and the people concerned with it. In each case the methods of producing the specific document or result are in the focus of this sort of research.

### ***Sampling***

As Rapley (2007) holds, the strategy here is building an archive of materials for analyzing the discourse or documents. Selection or sampling then refers to the documents to choose for this archive. Sampling in this kind of research first aims at constructing a corpus of material, which then can be a starting point for sampling inside the materials in this corpus.

### ***Comparison***

In conversation analysis, comparison is in many cases oriented towards a more general model (of how a conversation or a telephone call works), which is juxtaposed to the concrete case that is being studied. Often you will set up an inventory of several exemplars (like a number of beginnings of a counselling conversation) and then compare them with each other to find regular and structural principles in them. Comparison of documents is either focusing the contents or the structure of a series of examples.

### ***Generalization***

Discourse analysis is more interested in case studies, so that we can mostly find internal generalization here, whereas conversation analysis in most cases is interested in identifying general principles of talk and conversations. In this aim, a generalization from case studies to comparing different cases to general models is implied.

### ***Triangulation***

Rapley (2007) mentions the combination of analyzing documents and conversations in this context. A different way is to combine conversation analysis with interviewing, for example for analyzing professional practices by analyzing conversations and professional knowledge by doing and analyzing interviews (see below).

### ***Quality***

In studying conversation, discourse and documents, Rapley (2007) holds that there is no claim about truth in interpretations that can be justified. Rather,

researchers should try to convince the reader that their interpretations are plausible and credible by showing how they are based on materials and their analysis. Another approach to quality here is to validate documents in team meetings of the research group. In general it should become visible in the results and their presentations, how the researchers have checked and rechecked their conclusions against the material and especially against cases and materials that are deviant from their findings, conclusions and generalizations (as in a typology: see Flick, 2007, for managing diversity in qualitative analysis).

### ***Writing***

Depending on the research questions and the materials used in the study, different forms of writing can be appropriate here. An important issue here is how to demonstrate how conclusions are based in the material, and in particular in conversation analysis with its sometimes very detailed transcriptions, how to present excerpts from the material in an accessible and readable way. Too detailed a transcription can obstruct the perspective on the content and context of what is said; too simplified presentations of transcripts may undermine (the impression of) exactness. In discourse analysis in particular, writing and analysis are very closely linked to each other and writing up has an impact on the analysis (new ideas, categories or clarifications).

### ***Basic design***

Most research in these areas is providing a snapshot – analyzing materials in the here and now for what contents it includes or how it is structured. Conversation analysis is mostly based on comparing different examples (of counselling interaction, for example), whereas many discourse analyses are based on case studies. In document analysis we often find a retrospective approach (for example, how has the documentation of diagnoses of mental illness changed over the years?).

### ***Resources and stepping-stones***

In conversation analysis a major resource you need is time (and money) for transcription of what has been recorded before (with very good equipment if possible – another necessary resource). Finding and accessing the relevant material for describing the discourse about an issue can be difficult and time-consuming as well. This will be the case, in particular, if you want to have (all) the relevant material and not just a single example or an awkward selection. In conversation and document analysis, a danger is to forget about the contents and their meaning in analyzing basically the structure of the material.

### **Ethics**

If you base your research on (audio or video) recording for producing data, you will have to make sure that the participants know about and have consented to being recorded and that they have the chance to make you stop your recording if necessary. With the material you have to take care of anonymity issues – that nobody is identified in or through your research and that you do not talk openly about your participants. Here again, such issues become more difficult once you work with vulnerable people or sensitive topics (see Rapley, 2007, chap. 3).

### **Example**

In a study in the field of community psychiatry, I could show, for counselling conversations, how starting into conversation is organized so that a consultation begins and not something different. In the conversations analyzed, such a start may be designed rather open-ended (e.g. ‘What made you come to us?’ or ‘And what is it about?’ or ‘What is your desire?’). In other cases, they name the (given) topic for the counselling, or specific characteristics in the way the counselling conversation came into being. These openings, which begin the actual counselling relationship and delimit it against other forms of talk, are sometimes linked to explanations about the way the conversation came about. These explanations are specific for the situation (e.g. ‘So, your brother gave me a call’).

In analyses of the ending of first contacts in counselling processes, I could show two tasks to be achieved. A timely ending of the conversation has to be ensured. At the same time, the counsellor has to guarantee the continuation of the relation. This analysis could show which formal steps counselling conversations ran through more or less regularly. It could also show how these steps not only built up the conversation in itself, but also were influential in processing the clients and their cases – regardless of the specific content of their problems. So, the analysis was more formal than content-oriented, but shows the construction of cases in the conversations. Sampling was oriented towards cases of conversation and then to formal parts (beginnings, endings) in it; comparison of these parts was the aim of the design and analysis. Ethical concerns are about having the informed consent of clients and the counsellors and to maintain the privacy of both. This conversation analysis was triangulated with interviewing the counsellor, whose consultation was analyzed.

### **Conclusion**

These analyses often focus formal parts of verbal exchange more than the content, although discourse analysis is interested in both – form and content. Basic designs are comparative snapshots in most cases, although documents are sometimes studied with a retrospective focus. From an ethical point of view, anonymity and the sometimes confrontational character of the analysis should be taken into account,



in particular if you use the results (and examples) for training or other practical purposes.

### Key points

- These approaches to analyzing qualitative data are either quite open in the theoretical approach (as in coding) or come with a background (ethnomethodology in conversation analysis).
- They stress either content or formal structures.
- In both cases, ethics demand a special care for the anonymity of the participants in the analysis.

### Further reading

The following books will go into more detail about the methods outlined in this chapter:

Gibbs, G. R. (2007) *Analyzing Qualitative Data* (Book 6 of *The SAGE Qualitative Research Kit*). London: Sage.

Rapley, T. (2007) *Doing Conversation, Discourse and Document Analysis* (Book 7 of *The SAGE Qualitative Research Kit*). London: Sage.