About the author

Hennie Boeije received her PhD from Erasmus University Rotterdam in Health Care Policy and Management in 1994 and a Master's in Sociology. She is currently an Associate Professor with the Department of Methodology and Statistics of the Faculty of Social and Behavioural Sciences at Utrecht University, the Netherlands. Her primary areas of teaching and researching are qualitative research, mixed methods research, and pretesting surveys. Dr Boeije has written extensively in Dutch and internationally in the area of research methodology including the constant comparative method in data analysis and the presence of third persons in interviews. She has published qualitative research on the areas of refugees, nursing homes, informal caregiving and the chronically ill. She is currently working on a number of projects related to trauma issues.

ANALYSIS IN QUALITATIVE RESEARCH

HENNIE BOEIJE

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

anticipate the amount and complexity of the data set by planning for comparisons between the different parts that make up your data.

5. Perhaps you have formulated propositions or conjectures based on the literature review, conversations with experts in the field or a pilot study. Write them down and monitor whether they remain relevant or lose importance, whether they can be answered, and whether the answers in turn lead to new questions and conjectures.

6 DOING QUALITATIVE ANALYSIS

Qualitative analysis entails segmenting and reassembling the data in the light of the problem statement. But how does a researcher approach analysis in a concrete project? Where to start? How is coding done? How can creativity be preserved when using 'routine' procedures such as open, axial and selective coding? How routine are these procedures anyway? How to proceed and remain in control when using software for qualitative data analysis? The spiral of analysis (Figure 5.4) is the foundation for the methods and techniques discussed in this chapter. We start with open coding and axial coding as means to break up the data into smaller parts, and then proceed to selective coding which facilitates reassembly of the data. At the end of the chapter we look back at what we have been doing by transforming the data through coding.

LEARNING AIMS

By the end of this chapter, you will be able to:

- Describe the functions of open, axial and selective coding
- Position the three types of coding in the spiral of analysis
- Employ open, axial and selective coding in a small- or moderate-scale research project
- Identify different sources of codes
- Appreciate the comparison of codes assigned by different researchers
- List the four elements of the coding paradigm
- Recognize a core category in your own research
- Focus the analysis based on the relevant criteria
- Detect possible sources of bias during analysis
- Argue how alternating data collection, analysis and sampling support exploration and verification

Introduction to coding

Analysis consists of segmenting the data and reassembling them with the aim of transforming the data into findings. Findings can consist of descriptions that are more or less theoretical as well as interpretive explanations of the research subject (see Chapter 8). Findings of qualitative research always include interpretations of the empirical data. It is a mistake to consider raw data as findings. In the analysis phase of the research project data are sorted, named, categorized and connected, and all these activities entail interpretation. When you interpret someone's words or actions, you explain what you think it means and how you think it should be understood. Jorgensen summarizes the analytical process as follows:

> The analysis of qualitative data is dialectical: data are disassembled into elements and components; these materials are examined for patterns and relationships, sometimes in connection to ideas derived from literature, existing theories, or hunches that have emerged during fieldwork or perhaps simply commonsense suspicions. With an idea in hand, the data are reassembled, providing an interpretation or explanation of a question or particular problem; this synthesis is then evaluated and critically examined; it may be accepted or rejected entirely or with modifications; and, not uncommonly, this process then is repeated to test further the emergent theoretical conception, expand its generality, or otherwise examine its usefulness. (Jorgensen, 1989: 111)

Interpretive analysis is shaped in many different ways, but we can distinguish two types of analyses: one is oriented towards the themes or categories present in the data, and the other is oriented towards the cases, such as organizations, activities, events, situations or participants. As with many other terms in qualitative research, several different terms are used to describe the two categories: some refer to code-based and case-based analysis (Lee & Fielding, 2004), or cross-case analysis and within-case analysis (Merriam, 1998), while others refer to cross-sectional or categorical indexing and non-cross-sectional indexing (Mason, 2002), or to issue-focused and case-focused analysis (Weiss, 1994), and variable-oriented and case-oriented analysis (Miles & Huberman, 1994).

The start of many, if not all, of these types of analyses comes down to coding. Why? Because everyone has to start with reading the data and then separating the data into meaningful parts. The latter is, in essence, what we know as coding. Coding was developed as a technique in the grounded theory approach, and has been increasingly refined since (see Chapter 1). At first the data may appear to be a bulky, diverse collection of accounts, but coding is a tool with which to create order. In this book coding is presented as the most important tool for qualitative data analysis. The three types of coding that are distinguished, namely open, axial and selective coding, will finally lead to the production of the definitive findings.

Segmenting and reassembling (see Chapter 5) occur largely through the coding process. Charmaz (2006) understands coding to be the process of defining what the data describes. In more detail she conceives it as:

> Coding means categorizing segments of data with a short name that simultaneously summarizes and accounts for each piece of data. Your codes show how you select, separate, and sort data to begin an analytic accounting of them. [...] Coding is the first step in moving beyond concrete statements in the data to making analytic interpretations. We aim to make an interpretative rendering that begins with coding and illuminates studied life. (Charmaz, 2006: 43)

When coding, the researcher distinguishes themes or categories in the research data and names them by attributing a code. A code is a label that depicts the core topic of a segment. While coding, a researcher is looking for descriptions and sometimes for theoretical statements that go beyond the concrete observations in the specific sample. Charmaz, cited above, refers to coding as part of constructing a grounded theory. In that case special demands are made on codes because they need to be developed into concepts that together will constitute a theory. However, not all research, and consequently not all analyses and coding sessions, have to lead to a grounded theory. Lewins and Silver offer a much more pragmatic definition of coding:

> Qualitative coding is the process by which segments of data are identified as relating to, or being an example of, a more general idea, instance, theme or category. Segments of data from across the whole dataset are placed together in order to be retrieved together at a later stage. (2007: 81)

In this case a code may 'represent a deeply theoretical or analytical concept; it could be completely practical or descriptive; or it could simply represent "interesting stuff" or "data I need to think about more"' (Lewins & Silver, 2007: 83). Bear in mind the different uses of coding and codes while conducting your own project or reading someone else's.

Since the use of different terms in qualitative analysis can be quite confusing, frequently used terms as they are applied in this book are described in Box 6.1. They will be further elaborated on throughout this chapter.

BOX 6.1 IMPORTANT TERMS IN QUALITATIVE ANALYSIS

Analytic:	Conceptual – expressed in concepts – or theoretical
Category:	A group or cluster used to sort parts of the data during analysis and designated with a code
Code:	A word or string of words used as a name for a category generated during analysis
Concept:	A term referring to a category and used as a building block in a theory
Interpretation:	An explanation of the meaning of what is observed in empirical data
Pattern:	An orderly sequence consisting of a number of repeated or complementary elements
Theme:	The matter with which the data are mainly concerned
Theory:	A coherent framework that attempts to describe, understand and explain aspects of social life

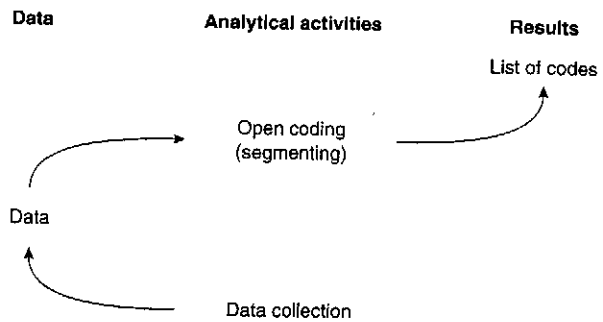

FIGURE 6.1 THE SPIRAL OF ANALYSIS ENLARGED: OPEN CODING

Strauss and Corbin (2007) distinguish three types of coding: open coding, axial coding and selective coding. These three types are expanded on in the following sections. Each section starts with the position of the type of coding in the spiral of analysis. Subsequently, details and frequently asked questions are addressed.

Open coding

Open coding is the process of 'breaking down, examining, comparing, conceptualizing and categorizing data' (Strauss & Corbin, 2007: 61). This means that all data that have been collected up to that point are read very carefully and divided into fragments. The fragments are compared among each other, grouped into categories dealing with the same subject, and labelled with a code. Open coding usually takes place at the beginning of the research project and starts during the collection of the first round of data. Little to no selection is made in terms of relevance of the research material, because it is still largely unpredictable what will be of value and what will not.

A code is a summarizing phrase for a piece of text which expresses the meaning of the fragment. It is by codes that fragments can be compared and filed. 'Qualitative codes take segments of data apart, name them in concise terms, and propose an analytic handle to develop abstract ideas for interpreting each segment of data.' (Charmaz, 2006: 45). Coding can be performed with paper and pencil, writing codes in the margin of a text, but coding can also be employed with the available computer programs. In Table 6.1 an example of open coding is provided from a student research project on dyslexia and choice of study and employment (Dam & Rooij, 2002). It concerns an interview with a male nurse who is dyslexic.

Exploration of the data by open coding constitutes the start of conceptualization of the field of research. Codes provide an analytic handle on the data. Open coding encourages a thematic approach since it forces the analyst to break up the text into pieces, to compare them and to assign them to groups that address the same theme. Open coding contributes to a clear organization of the data as well, since it results in an indexing system that fits the researcher's

TABLE 6.1 EXAMPLE OF OPEN CODING OF AN INTERVIEW WITH A PERSON WITH DYSLEXIA

Code	Interview transcript
	I: How did working as a psychiatric nurse work out for you?
Falling short	P: Especially in psychiatry I fell short, because I was asked to make analyses, in other words reports and behaviour along with it, and I really fell short. I obtained my diploma as a psychiatric nurse, and then I moved on as this area was not the right choice for me.
Importance of language Choosing a career	I: Because you ran into problems with reporting you have moved on to something else?
	P: Yes, I did continue to work in nursing, but in general nursing instead. Reporting is also important there, but less so.
Career change Job match	I: If you would be given the choice, regardless of aspects of language, what would you like to do most?
Preferred job	P: I felt most comfortable working on the ambulances.
	I: What was the role of language there?
Importance of language	P: It was again very important, but the forms had a specified format, reporting had to be accurate and very brief and it was not that important. I do remember that I did okay as long as I kept my sentences short. And that was enough for the ambulance-world. I also had to transfer everything orally, but that was not a problem for me.
Aids	I: But if you had not finished your report on time, how did colleagues respond to that?
	P: Well, it was not a matter of finishing, but I made more spelling errors, and the sentences were illogical. But they thought I didn't know better. The first time they mark it with a red pencil, but then they just read over it, because they know what the essence is, what is meant. [...]
Flaws Convey message	
Discovering dyslexia Reading exercises	I: I: How did you find out you were dyslectic?
	P: I never knew I was dyslectic until I had to write those reports. I used to visit the library three to four times a week. I've read many books, into the middle of the night. That has probably saved me and it has enabled me to choose this career.
Choosing a career	I: How has this been as you've aged?
	P: I have stopped reading, and I fill out the forms in short sentences, and occasionally there is a spelling error, but that washes over my broad back. I do not care at all. If it is really important, I have it checked or use the spell-checker on the computer. But I must admit that since I stopped paid employment, I am increasingly wondering how some things need to be written.
Indifference Aids Reduced writing skills	I: I wonder to what extent you experience your dyslexia as a handicap?
Experiencing no handicap	P: I haven't really experienced it as a handicap. I was just worse at language than others, but other than that it didn't matter. As a nurse, I was just as good as the others. I have never noticed anyone looking down at me for making so many mistakes in my language. I have to add that as an ambulance nurse you have your own responsibilities and you report to a superior and not to your colleagues. If I read a booklet, the sentence structure is completely strange to me. That is all I notice about my dyslexia, but it does not really affect my life. It's a society of reporting, reading and writing, and that is a society which badly needs language and is very fast-paced.
Equality Job match Unaffected life Languaged society	

Source: Dam & Rooij, 2002

analytical needs. A code enables the easy retrieval of the fragments that have been assigned a specific code.

The process of open coding involves the following steps:

1. Read the whole document.
2. Re-read the text line by line and determine the beginning and end of a fragment.
3. Determine why this fragment is a meaningful whole (text which belongs together and deals with mainly one subject).
4. Judge whether the fragment is relevant to the research.
5. Make up an appropriate name for the fragment, i.e. a code.
6. Assign this code to the text fragment.
7. Read the entire document and code all relevant fragments.
8. Compare the different fragments, because it is likely that multiple fragments in a text address the same topic and they should therefore receive the same code.

FAQ 1

Q: Can coding be done using a computer?

A: Yes, that is certainly possible and there are several advantages in using a software package designed for qualitative data analysis. Some researchers prefer using paper and pencil, others use their word processing program, but software packages for qualitative data analysis are becoming more readily available. When using a computer program for qualitative analysis, the researcher goes through the same steps as when coding manually, but the steps are automated instead. Any number of codes can be assigned to a single segment of text of any size, and to overlapping or embedded segments as well. Software also offers opportunities that are not really possible without a computer, or at least not with the speed of a computer, such as code searches, matrices or frequencies of codes (see Chapter 7).

The result of open coding is a list of codes, also referred to as a 'coding scheme'. Within a software program codes can be listed or organized in different ways. There are non-hierarchical and hierarchical coding schemes, in which higher-level codes can contain sub-codes. The codes can be sorted alphabetically or in an order determined by the researcher. When indicating a hierarchical structure of a coding scheme, we speak of the code tree in this book. For the dyslexia example, an alphabetically sorted list of codes may resemble Figure 6.2.

Multiple codes may be assigned to fragments if they contain information of multiple topics which are relevant to the research. However, when retrieving fragments, the same fragment will of course show up for every code that it is assigned to. This is convenient if the researcher really wants to retrieve the fragment for more than one code, but it is not helpful in reducing the amount of material. This

Aids
Importance of language
Choosing a career
Career change
Convey message
Discovering dyslexia
Equality
Experiencing no handicap
Failing short
Flaws
Indifference
Job match
Languaged society
Preferred job
Reading exercises
Reduced writing skills
Unaffected life

FIGURE 6.2 EXAMPLE OF A CODE TREE IN THE DYSLEXIA RESEARCH

is why it is sometimes better to choose to file the fragment with one code only. Such a choice shows an awareness of the relationship between the various emerging categories at this stage of the analysis: the fragment is seen as being part of one category and not the other.

FAQ 2

Q: What should I do with fragments which are irrelevant in my view?

A: Some researchers recommend deleting pieces of text that you believe are not relevant by crossing them out. However, the reverse can apply: bits of text which are not coded will eventually be left out of the next phases of analysis. They have not been labelled, and thus the researcher will not look for them anymore. Some computer programs explicitly offer the option to save pieces on which there is doubt, or for which the researcher has yet to come up with a tag but that do seem interesting. This tool usually includes the word 'free', as in free quotations.

Anyone who starts coding will quickly realize that it is more than simply writing down a word in the margin. The main activity during open coding is to ask questions about the data, such as: What is going on here? What is this about? What is the problem? What is observed here? What is this person trying to tell? What else does this term mean? Which experience is represented here? Asking these kinds of questions about your empirical data will familiarize you with the data and lead to a better understanding of them. In open coding, doing (actually assigning a code) and thinking (coming up with good questions and codes) converge.

FAQ 3

Q: To be honest, I don't understand the difficulty in the whole coding procedure. I've grouped together all the answers received for the first question on my questionnaire, and labelled them 'application'. Then, I clustered all of the answers on the second question, and labelled them 'work satisfaction'. And so on and so forth. Am I doing something wrong?

A: The way in which you approach coding reflects your research purpose and design. You probably knew what you were looking for, so you can be more explicit about the themes or categories to be considered at the outset of the coding process. This reflects a more deductive approach to research, including coding – this is rather more common in applied research, which strives for specific and immediate objective comments or specifically identified outcomes (Lewins & Silver, 2007). Most likely you have used a relatively structured questionnaire with little to no additional questioning beyond what was specified beforehand in the questionnaire. The results of the research will resemble a topical or thematic survey (see Chapter 8), and merely consist of a summary of the answers which were most frequently given.

Even though such a research design is irrevocably qualitative, this chapter – and in fact the whole book – is concerned with a more inductive approach to research and design. This approach identifies and develops new concepts, while existing theoretical concepts will not a priori over-define the analysis (Lewins & Silver, 2007). This implies that data collection takes place in a less structured fashion, and the answers to your questions and 'what you are looking for' may be found scattered through the different documents. In parallel, initial data analysis is less structured, since the frame of analysis has yet to be developed. This way of working forces researchers to find relevant themes and name them, and this is exactly what makes coding a very demanding task.

FAQ 4

Q: What is an appropriate length for a fragment to be coded?

A: There is no simple answer to this question. What is important is that the fragment is a coherent piece of text. This becomes apparent when one reads the fragment and considers its comprehensibility. Most fragments which consist of only one line do not meet this criterion. On the other hand, fragments which exceed half a page may bypass the purpose of open coding, which is to unravel a text, reduce the amount of data and organize the data.

FAQ 5

Q: Is it necessary to code the interviewer's question with the fragment?

A: Many researchers do not code the question, but merely register what the participant had to say. There are arguments in favour of coding the question as well. Doing so enables an evaluation of the answer given, i.e. was it really an answer to the question? Providing the question may improve the interpretation of the answer given, because it gives it context. It may not be necessary to code the questions with every single fragment, but only when it facilitates understanding the text segment.

How do researchers come up with codes? Many words from daily language may be used as a code. A code does not have to be complex or difficult to understand. Specific codes derived from the participants' terminology are known as 'in vivo' codes or field-related concepts (Strauss & Corbin, 2007). These codes then are identified within the data. 'Falling short' is an example of an in vivo code in the study on dyslexia. In vivo codes are not just catchy words; rather they pinpoint exactly what is happening or what the meaning of a certain experience or event is. Charmaz (2006: 55) describes in vivo codes as:

1. General terms everyone 'knows' that flag condensed and important meanings.
2. Terms made up by participants that capture meanings or their experiences.
3. Insider shorthand terms specific to a particular group that reflect their perspective.

Analysts must develop their theoretical sensitivity in order to notice useful in vivo codes for their research material (see Chapter 5).

Another source of codes is the concepts derived from social theories that researchers have come across during disciplinary education and while reading the literature. These kinds of codes are commonly known as 'theoretical concepts' or 'constructed codes' (Flick, 2006). They are, after all, the constructions of social scientists. Examples of theoretical concepts in the dyslexia research can be 'equality' and 'languaged society'. Of key importance are the sensitizing concepts which can be used as a code (see Chapter 2).

FAQ 6

Q: I have yet to start collecting my data, but I already have several codes. Is that possible?

A: Yes, that is possible. In this case codes may be derived from the professional arena as well as the literature (see above). When you already have several

(Continued)

(Continued)

codes beforehand it suggests that you are using a more deductive approach to coding. This might either stem from an applied research with an immediate aim for practical understanding (see FAQ 3) or it might stem from a clear theoretical frame that you want to verify on new cases. The research will therefore start with some predefined, higher-level areas of interest which are explicitly looked for in the data (Lewins & Silver, 2007).

An example of a deductive project that applied existing theoretical ideas is the earlier example on football hooligans (Spaaij, 2006). The author wished to verify the proposition that the nature of football hooliganism depends on the fault lines within a society by applying social identity theory. In addition to his empirical work, he wished to add the local influences on hooliganism to the existing theory. The possibility of changes over time in hooliganism was also investigated. This part of the research had a more explorative character. Within his deductive approach he allowed for emergent themes to be incorporated.

FAQ 7

Q: I am using the computer for my analysis. Can I import the codes I already have before I start my first round of data collection and coding?

A: Yes, codes can be generated at any point during the research independently of data (Lewins & Silver, 2007). Specialized software programs offer more than one way to work with codes. One is to generate a list of codes prior to commencing field work and attaching fragments to them later on. Another way is first importing the text files and then coding them. Each time a new code is needed, it is added to the code tree. Often researchers use both methods: they generate a small code tree before coding the data and adjust it and add codes to it when they actually start analysing their empirical material.

FAQ 8

Q: Can abbreviations be used as codes?

A: Yes, a researcher may use codes and abbreviate them. An advantage is that this provides even shorter names. A disadvantage to using abbreviations is that you have to remember what each abbreviation stands for, something which may prove to be difficult. To my knowledge very few researchers use abbreviations, even though Miles and Huberman (1994) use them frequently in their book on analysis.

Both descriptive and interpretive phrases may be used as codes (Miles & Huberman, 1994). It is quite possible when first entering the field that codes are needed to describe what is seen and heard. In that case, predominantly descriptive codes are used to label the observational reports. In this stage of the research, field workers particularly pay attention to the people involved (numbers, groups, activities, appearances, types of interaction), the physical environment (rooms, furniture, sounds, lighting) and the sequence in which events recur (work, play, breaks, discussions).

Figure 6.3 provides an example of a code tree that was constructed as part of a study on opportunities and obstacles of informal interactions in organizations

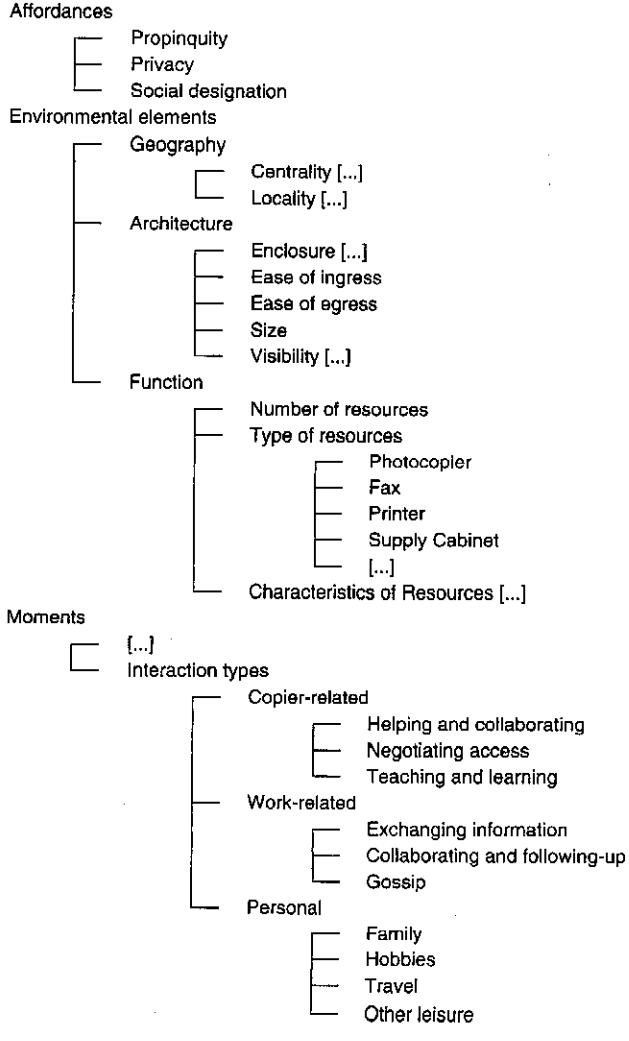

FIGURE 6.3 EXAMPLE OF DESCRIPTIVE CODES FOR OBSERVATIONAL NOTES (FAYARD & WEEKS, 2007)

(Fayard & Weeks, 2007). The authors chose photocopier rooms in order to examine if these rooms encouraged social interactions between employees, of what kind these interactions were, and what conditions provoked interactions. Note that only a part of the code tree is represented here.

After this initial phase, it is important to discern the meaning of the observations. This is where interpretive codes come into play. If it is exclusively 'facts' (activities, places) that are coded, then the meaning of the text would be lost. Experiences, perspectives, impressions and views are important in qualitative research (see Chapter 1). The researcher needs to gain experience in theoretical sensitive coding, which means that skills need to be gained to understand the key point in a piece of text, grasp what it is all about, notice meaningful expressions, and read between the lines. In the study on dyslexia, a remark such as 'I have never noticed anyone looking down at me for making so many mistakes in my language' (see Table 6.1) is very likely to harbour a lot of information on inferiority, superiority and unequal treatment. It is the researcher's task to be alert to these kinds of statements and to interpret them. Another example of this is given below, taken from an interview on binge drinking:

Interviewer: Why do you engage in binge drinking?
Interviewee: Well, when I go out and meet my friends we immediately start drinking. You have to do something you know. And with alcohol the chances are high that something exciting is going to happen. This is permanently lacking throughout the rest of the week.

One could code the entire fragment as 'reasons for binging', and simultaneously code parts of the fragment as 'social contacts' and 'excitement'. The key issue is that only assigning the code 'reasons for binging' results in loss of the main message that is being communicated. That message could very well be that life is dull and boring and that the weekend should compensate somehow. This meaning would be more apparent if we were to assign the fragment with the code 'compensating boredom' or 'evoking excitement'.

Charmaz (2006) particularly advises coding for action and process. This is made easier by coding with gerunds (verbs ending in '-ing', e.g. eating, swimming). You gain a stronger sense of action and process when you read 'choosing a career' instead of 'choice of career' and 'discovery of dyslexia' is more static than 'discovering dyslexia' which implies that an activity is involved. So you should keep your codes active and close to the data. Examples of questions that can help are: What process is at issue here? How does it develop? How does the participant act while involved in the process? When, why and how does the process change? What are the consequences of this process? (Charmaz, 2006: 51).

FAQ 9

Q: Sometimes I would like to see the broader context of the text fragment that I have coded. Is this possible?
A: Yes, with most computer programs you can go directly to the position in the document from which the fragment was taken. When analysing with paper and pencil, this is certainly more labour intensive. A summary of each document can be helpful, providing a quick résumé of the most important findings of an interview or observation. Context information is valuable for interpreting the fragment against the broader background of the entire document.

FAQ 10

Q: I have formulated a number of sensitizing concepts based on my literature study. Yet I still feel that I cannot properly use them in open coding. Why does this happen?
A: During open coding, the researcher starts analysing at a fairly detailed level in order to form a foundation of the analysis in the data. The data describe daily events and experiences and in order to name them appropriately, concepts of a low level of abstraction are needed to match the data. Sensitizing concepts often have a general character, and still need to be applied to the field of research (see Chapter 2). They possess a level of abstraction which transcends that of regular fragments. It is therefore no surprise that these concepts may not be usable in the open coding phase. In later phases of coding, in which analysis takes place at a higher level of abstraction, these codes may turn out to be useful.

In open coding, researchers should not globally examine the text but do so in detail. In research on dyslexia and job choice, for example, codes such as 'motivation', 'discrimination' and 'language' are not specific enough at this stage of the analysis. Such codes will come to harbour a large number of very diverse fragments and the disassembly of the data will be insufficient. The result is that fragments which have received the same code will have to be differentiated later on. In short, this means postponing the analysis. However, a coding scheme which is too detailed is not recommended either. When every distinguishable fragment in a text receives a different code, the researcher goes beyond comparison and categorization. This approach may easily result in hundreds of codes.

FAQ 11

Q: Is there a usual number of codes for the open coding phase?
A: This is impossible to say. A coding system is uniquely developed for each separate research project, and is related to the modus operandi of the researcher. It is important for researchers to develop their own coding system. For a small-scale study (with a narrow research question, a limited number of participants and

(Continued)

(Continued)

restricted data collection), 50 codes may be sufficient. For a large-scale investigation, there may be as many as 100 to 200 codes. Keep in mind that this is a rough estimate only.

An indication of whether an 'appropriate' number of codes has been awarded is the number of fragments hosted by one code. If there is only one fragment assigned to a single code, this code should probably be discarded or merged with another code, and the fragment filed under this other code. If there are more than 20 fragments assigned to a certain code, the diversity of these fragments should be evaluated. If the diversity is deemed too large, the fragments may have to be divided into two or more categories. More specific sub-codes can then be assigned to these smaller groups of fragments. Another indicator is the ratio between main codes and sub-codes: 10 to 20 main codes seem to be an acceptable starting point. The number of sub-codes can vary strongly between researchers, as it depends on how detailed the researcher's coding is (see next section).

FAQ 12

Q: Is it necessary to code all sorts of personal information of the interviewees, such as gender, age and study or profession?

A: No. Codes such as 'age', 'place of residence' and 'profession' are useless because they will not become a part of later descriptions nor of the theoretical model. This information about participants is best kept in a separate file about the composition of the sample. Some computer programs offer the option of assigning variables or attributes to a document, for example, an interview could receive variables representing age and gender of the interviewee. This is a wise thing to do for reasons of comparison, for instance, whether female binge drinkers engage in disorderly behaviour and if so whether this behaviour is different from that of males. Then again, it is not the demographic variable itself that is of interest per se, but the mechanism behind it. For instance, female binge drinkers run into different types of trouble than male binge drinkers, and it is relevant to describe the differences and the causes.

In this phase of a research project, it is recommended that researchers work in a group instead of on their own. Having others to confer with contributes to a well-developed coding system, thereby ensuring that certain fragments are systematically awarded the 'correct' code. This is known as 'inter-rater reliability'. Asserting inter-rater reliability does not tell us anything about the adequacy of coding, but it does cover the systematic approach to coding. Every fragment dealing with a certain topic

will end up with the same code; they will not be awarded one code at one time and another code the next. In the end, this procedure does give an indication of the adequacy of the coding scheme, since it is assumed that the members of the group discuss the interpretation of the texts, exchange their views and come to an agreement.

FAQ 13

Q: Can software for qualitative data analysis calculate the inter-rater reliability?

A: According to Lewins and Silver (2007) who did an extensive comparison of different software packages, only N6 and CISAID can calculate a reliability rating. In other programs the evaluation of the level of reliability is not so much based on assigning the exact same code between independent analysts as on discussing how parts of the texts should be interpreted. MAXQDA2007, for instance, offers the possibility for different analysts to code the same text. The coding of other raters can be turned off so as to not be influenced by them. Later on, when all codes are made visible, a comparison can be easily conducted.

Eventually, all relevant data must be covered with the generated codes. Judging what is relevant is difficult at the start of the research. The research questions are the most important guideline for doing so. The phase of open coding can be ended if no new codes are necessary. Another term for this is 'saturation': the information from a new case can be separated into fragments that can be covered with one of the already existing codes. The entire process may repeat itself until the second or third round of data collection (see Figure 5.4, the spiral of analysis). Every time new observations provide a reason for generating a new code, open coding is resumed. Box 6.2 contains a summary of the most important elements of open coding.

BOX 6.2 SUMMARY OF OPEN CODING

Purpose:	Exploration of the field, coverage of the field with codes, manageability of data files, familiarity with data
Phase:	Especially in the beginning of the research
Activities:	Reading and re-reading (close reading), asking questions about the data, comparing data with data, assigning codes to data fragments
Starting point:	Data
Results:	List of codes, memos
Validation:	Saturation, meaning that no new codes are needed to label fragments that appeared in the data up until now

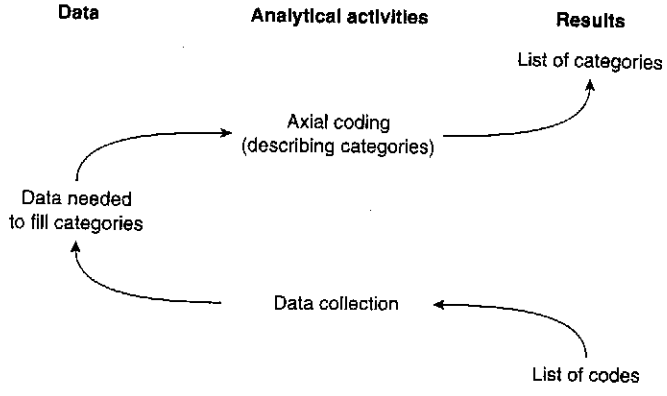

FIGURE 6.4 THE SPIRAL OF ANALYSIS ENLARGED: AXIAL CODING

Axial coding

The term 'axial coding' refers to 'a set of procedures whereby data are put back together in new ways after open coding, by making connections between categories' (Strauss & Corbin, 2007: 96). Axial coding is a more abstract process and consists of coding around several single categories or axes (see Figure 6.4). The term is somewhat confusing and personally I like 'focused coding' (Charmaz, 2006) better, but axial coding is the more commonly used term. 'Axial coding relates categories to subcategories, specifies the properties and dimensions of a category, and reassembles the data you have fractured during initial coding to give coherence to the emerging analysis' (Charmaz, 2006: 60). When researchers employ axial coding, the reasoning moves predominantly from codes to data, whereas in open coding the reasoning moves in the opposite direction, from data to codes. While in this phase of the analysis, new ideas are generated by conducting another round of data collection. Categories and propositions generated in the previous phase are tested by confronting them with the new material. In axial coding the distinction is made between the categories for which everything indicates that they might fulfil an important role in the definitive findings and the categories that can be gathered around these particular categories. The relationships between salient categories (axes!) and subcategories can be generated, modified, refined, elaborated or even rejected throughout axial coding.

Axial coding involves the following steps:

- Determine whether the codes developed thus far cover the data sufficiently and create new ones when the data provide incentives to do so.
- Check whether each fragment has been coded properly, or if it should be assigned a different code.
- Decide which code is most suitable if synonyms have been used to create two equal categories, and merge the categories.

- Look at the overview of fragments assigned to a certain code. Consider their similarities and differences.
- Subdivide categories if necessary.
- Look for evidence for distinguishing main codes and sub-codes and assign the sub-codes to the main code.
- See whether a sufficiently detailed description of a category can be derived from the assigned fragments and if not, decide to collect new data to fill the gap.
- Keep thinking about the data and the coding.

The primary purpose of axial coding is to determine which elements in the research are the dominant ones and which are the less important ones. As insights into the field increase and ideas about the observed social phenomena develop, confidence grows in making choices among the codes and the connections between them. The second purpose of axial coding is to reduce and reorganize the data set: synonyms are crossed out, redundant codes are removed and the best representative codes are selected. All activities are employed to gradually focus the research.

Sensitizing concepts come into play again during the phase of axial coding. As described in Chapter 2, sensitizing concepts have a rather general content at the beginning of the research. They give you ideas of directions to pursue and sensitize you to ask particular kinds of questions about your topic (Charmaz, 2006). The sensitizing concepts can be filled in during the analytic phase of axial coding to get their definitive content that fits in with how they are used by the participants in the field (Jorgensen, 1989). During axial coding, one tries to determine the properties of the categories. In doing so it becomes clear by which indicators a category can be recognized in the data. Note that with a definitive concept it works the other way around: its fixed content is reflected by its measure, i.e. the indicators that stand for the concept. Qualitative research is well suited to discover the contents of a concept when these contents are unclear. This function is also used in a particular type of mixed methods research that aims at the development of a measuring instrument (see Chapter 8).

In the dyslexia study (Dam & Rooij, 2002), the following main codes were awarded after several rounds of data collection: previous history, experiencing problems, choosing education, choosing a career, compensating, own responsibility and environment. The other codes were considered to be sub-codes of these; that is to say, they were considered instances, specifications, parts or stages of the main codes. The code tree is depicted in Figure 6.5.

The hierarchical sequence in the tree structure is simply a clear way to sort and organize the subjects, making it easy to retrieve the data pertaining to the subjects later on. It is possible that there are no fragments assigned to a main code, like 'experiencing problems', 'choosing a career' or 'strategies' in Figure 6.5. Such a code then functions as a hanger for others. But the main code can also contain fragments which clarify what is meant by the main category. For example, 'past history' can contain fragments that have to do with significant events in the lives of participants leading up to the discovery of dyslexia and 'compensating' can hold fragments in which participants explain what they mean by this term and when they feel they need to even out their flaws.

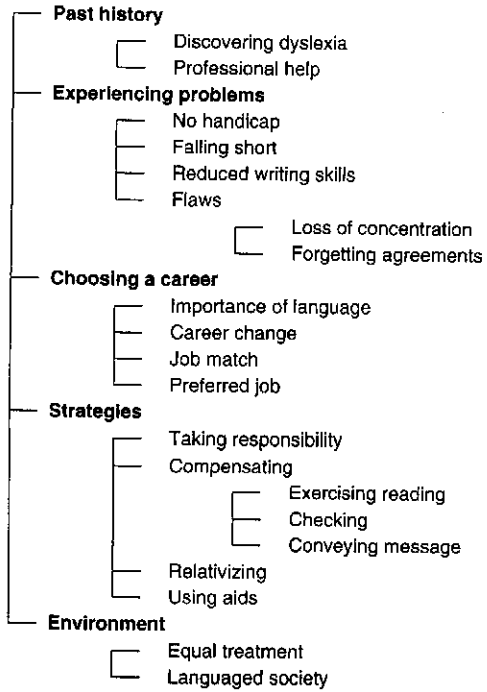

FIGURE 6.5 EXAMPLE OF A CODE TREE IN THE DYSLEXIA RESEARCH AFTER AXIAL CODING (DAM & ROOIJ, 2002)

FAQ 14

Q: In the beginning I felt comfortable using the hierarchical ordering of the list of codes. But now that I see connections between the codes I wonder whether I can use the software for linking the codes in different ways?

A: Different software packages offer different ways of escaping the structures of the main code listing (Lewins & Silver, 2007). In this respect you can think of codes to be retrieved in different combinations independent of the main structure. You can bring the different elements of your project – codes, memos, segments – into a mapping device; you can link fragments within and between documents, and you can regroup the documents and codes in sets or families leaving the original code structure intact (see Chapter 7).

At this point during the analysis it is important to remember that codes can be used for different purposes, whether more descriptive or analytical (see the beginning of this chapter). In the list of codes in Figure 6.5 you can see that some of the codes are more

descriptive, such as 'past history' and 'flaws', and that others may be more analytical, like 'choosing a career' and 'strategies'. There is no set way in which to use a code; they must simply be meaningful for you or the research team in the sense that they indicate the nature of the data grouped by that code in some way (Lewins & Silver, 2007).

In the case of analytical use of codes, thinking about inter-rater reliability is important. When codes are to capture important properties of a concept, it is essential that different researchers agree when these properties appear in the data, i.e. what fragments are seen as evidence for these properties. Subsequently, they have to be coded with a concise and accurate label. When the codes are used in a more practical or descriptive way, the value of inter-rater reliability diminishes and can even become a hindrance. When you use a code as a 'box' to put in interesting data to revisit and think about, then it is not that important how you label the box. Spending a lot of time on agreeing on the correct code is a waste of time and a hindrance for creative labelling and may cause unnecessary delays to the project.

FAQ 15

Q: I only have limited time for my qualitative research and my sponsor has already provided a number of relevant themes. Can I go straight into axial coding and skip open coding?

A: In commissioned research the exploration of the field during open coding is sometimes seen as a waste of time, because so much detail is unnecessary and because 'you do not have to start from scratch'. Much depends on the purpose of the investigation. For example, when the purpose of the research is to make an inventory of a number of predetermined topics, data will often be collected with a reasonably structured measurement instrument. In this case, the need for open coding is minimal. The data can be assigned to anticipated categories, leaving some room for adding categories by means of open coding (see also FAQ 3).

In the axial coding phase, the codes become increasingly disconnected from the data although the data never entirely disappear from sight. In the open coding phase the data were central and evidence was built-up to be used later on. Open coding logically paves the way for axial coding, as the basic data is transformed into a more abstract framework through – let's not forget – the researcher's hard work. Nothing emerges without effort. This organizing structure is grounded in the data, preventing friction between the data and the framework. By systematically segmenting the fragments assigned to a code the researcher gains the ability to describe the category at a more abstract level. An example may clarify this issue. In the study on the quality of care in nursing homes (see Chapter 5), staff members often spoke of the hospitalization of residents (Boeije, 1994). The fragments labelled with the code 'hospitalization' are listed:

- 'Being hospitalized is being used to an institution, that everything is done for you, and that you will see yourself as not being able to care for yourself anymore.'
- 'The process of hospitalization contributes to a decline in interest; everything that happens outside the institution is not of concern.'
- 'Hospitalization is the loss of autonomy, going along with the rhythm, not wanting to do anything else, adaptation to rules. Initiatives are no longer necessary.'
- 'Those people become very passive, they stop thinking for themselves.'
- 'Hospitalization means that people are fully self-directed, like "I should get my porridge now", and "I have to go to the bathroom now", and "I have to go to bed". Everything needs to be done at its own specific hour and time.'

The various fragments are compared to each other in order to extract the core of the hospitalization process as it is viewed by the staff. Hospitalization is depicted as a process that takes place over time. It is the process by which residents lose their autonomy and stop thinking for themselves, take no responsibility anymore and leave everything to others. The consequence is that their world is reduced to the confines of the nursing home. Hospitalization is an unwanted phenomenon with respect to self-worth. In the hospitalization process residents give up their 'selves', which is at odds with striving for self-worth. When residents are 'de-hospitalized' their self-worth increases again.

In order to ensure theoretical sampling (see Chapter 2), comparison cases were sought (other nursing homes with other staff members) for a new round of data collection. Emphasis was put on locating settings in which residents were not hospitalized. Also, cases were sought in which hospitalization was valued positively. This led to the discovery that some staff members thought it to be effective when residents had little say and went along with the working routines established on the ward. Note that hospitalization had become an important category in the research, because of its relationship with resident's self-worth and the quality of care.

FAQ 16

Q: I have performed open coding with the support of a software package. How can I use the computer in axial coding?

A: The computer is a great tool for creating a hierarchical code tree and for retrieving fragments that have been assigned a certain code. Both these tasks are relevant in the stage of axial coding. Software packages are able to trace the conceptual development of the data together with the various phases and processes of the entire project. During axial coding visual tools and various search options can assist in the discovery of themes, patterns and connections between the different categories (see Chapter 7).

We have addressed the role of the coding paradigm in the discussion about 'forcing' versus 'emerging' in Chapter 5. Remember that the coding paradigm was used as a mould for organizing the data and consists of four elements, namely context, conditions, interactions/strategies and consequences. The coding paradigm is meant to be applied during the axial coding phase (Strauss & Corbin, 2007). Application to the above-mentioned research could, for example, lead to the following distinctions:

- What is the *context* in which hospitalization takes place?
  - For instance, placement in a relatively closed institution due to severe physical and/or mental limitations of the elderly; wider nursing home policy including staffing.
- Under which *conditions* does hospitalization take place?
  - Fixed daily schedule; structures, routines and rituals; absence of variety of stimuli or impossibility of processing these stimuli.
- Which *interactions* take place?
  - Staff and residents meet when a deviation from the daily schedule upsets residents; staff and residents negotiate possible deviations; residents delegate activities to staff members; staff members urge residents to adjust to the valid, same care pattern every day; caregivers urge residents to take initiatives.
- What are the *consequences*?
  - For residents: passivity, loss of self-worth, rest, surrender. For staff: control, boredom, detachments, dissatisfaction, frustration.

Salient codes are raised to categories. They represent important themes in the data and can lead to more abstract, theoretical ideas. In the axial coding phase, both the definition and the properties of a category become clear. When coding, it can be useful to think about properties in terms of dimensions (Strauss & Corbin, 2007). For instance, 'equality' could be appropriate when studying dyslectics in a social setting. Dimensional categories of 'equality' could be 'equal treatment' and 'disadvantaged'. A different property of the social context may be 'nature', which could be dimensionalized as 'languaged' or 'physical'. It can be proposed that people with dyslexia tend to withdraw from social environments that are perceived as 'languaged'. It then becomes easier to argue what comparison cases are needed for further data collection to test this proposition. Thinking in dimensions contributes to theoretical selection and constant comparison (see Chapters 2 and 5). Dimensionalizing is useful, but only when done in moderation otherwise the researcher is in danger of losing track of the coding system because it becomes forced and unwieldy.

The logic of going back to the data and moving forward into analysis is that missing knowledge often comes to light when glossing over the categories that have been developed. If you detect a code that has only one fragment assigned to it while glancing over the tabular summaries of code frequencies, consider whether it is best to delete the code (known as 'uncoding') or to fill in the gap and gather more data about that particular topic. When your categories are still only briefly described and are not convincing, you should return to the field and gather more data to enhance the descriptions. Focused data collection is needed to fill conceptual gaps. Ultimately, simultaneous data collection and analysis help you to gain more insight into the research problem.

In this phase of the analysis, regularities or patterns emerge which rise above the level of a single text fragment. Usually such a pattern is difficult to capture in a

code, although Miles and Huberman (1994) speak of pattern codes in this way. A pattern can be defined as 'an orderly sequence consisting of a number of repeated or complementary elements' (Fredericks & Miller, 1997). Patterns have an overarching character and cannot be placed within one category. In her study on family relationships, Mason gives 'reciprocity' as an example of such a overarching topic and remarks 'that it is unlikely that such a process of reciprocity will be neatly bundled into small chunks of interview text ready for the reader to categorize and index' (2002: 119).

In other words: the encompassing topics in which the researcher is interested are not readily observable in a piece of text. Ideas, patterns, explanations and types which do not 'fit' into codes are quickly forgotten. In order to remember and develop them, a researcher would write memos (see Chapter 4). Gradually, coding becomes aimed at explaining larger parts of the data and at bringing the different parts back together. Axial coding can be finalized when the distinction between main codes and sub-codes is clearly established and the contents of the categories are known. This means that the preliminary findings up till then do not change substantially when confronted with newly gathered data. The test again is saturation: newly collected data from comparison cases do not change the outcomes so far.

BOX 6.3 SUMMARY OF AXIAL CODING

Purpose:	Describe and delineate categories, determine relevance of categories, increase level of conceptual abstraction
Phase:	Halfway through
Activities:	Retrieve and compare fragments assigned to a certain code, define the category, determine relationships between main categories and subcategories, add, refine and check preliminary ideas and conjectures with newly added data
Starting point:	Codes developed during open coding
Result:	Categories are described and distinctions are made between main categories and subcategories, memo file containing ideas and verified assertions
Validation:	Saturation, meaning that the definition and properties of each category (axis) are clear and that no further adjustment is needed

Selective coding

Selective coding refers to looking for connections between the categories in order to make sense of what is happening in the field. Selective coding is aimed at integrating the loose pieces of your earlier coding efforts and can be considered a logical step after the segmenting of the data:

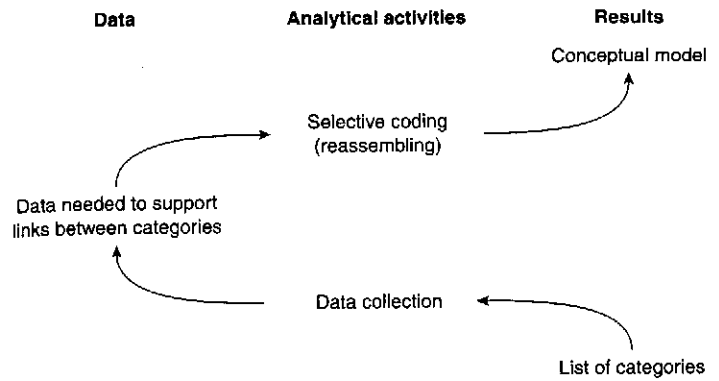

FIGURE 6.6 THE SPIRAL OF ANALYSIS ENLARGED: SELECTIVE CODING

> Efforts to code data will lead to sorting, sifting, organizing, and reorganizing these materials, usually into larger units and components. Sometimes this involves flashes of insight about how things fit together, while at other times it depends on less dramatic hunches, or simply hard work. Is a pattern or type discernible? Is some sequence or process apparent? Can you ascertain connections or relationships among concepts? [...] As you sort, sift, arrange, and rearrange the data and analytic labels and comments about them, it will be increasingly necessary to become more directly and explicitly involved in theory and theorizing. (Jorgensen, 1989: 110)

Often, but not always, the findings of a qualitative piece of work will aim at theory development. When this is the case, it is in the process of selective coding that certain categories are adopted as theoretical concepts, since they will most certainly become part of the theoretical model. For authors such as Strauss and Corbin the end result of qualitative research is a grounded theory, and for them selective coding equals 'selecting the core category, systematically relating it to other categories and filling in categories that need further refinement and development' (Strauss & Corbin, 2007: 116). Theoretical coding as an alternative term for selective coding is therefore a rather good choice for this sophisticated level of coding (Charmaz, 2006).

The term 'core category' originally stems from the grounded theory approach and refers to a category that is central to the integration of the theory (Strauss, 1987). In other traditions, the term may not be used at all. One of the most compelling examples of a core category is the awareness context of Glaser and Strauss (1965) (see Box 1.4). We have already come across some examples of core concepts in our illustrative studies, such as 'trust' in a research on residents in nursing homes (Bosch, 1996) and 're-connecting the person to humanity' in a study on care of psychiatric nurses for suicidal people (Cutcliffe et al., 2006).

The core category or core concept is a construction of the researcher which does not magically emerge from the data. The core category describes and explains

the researcher's observations. A number of characteristics of the core category are provided by Strauss (1987: 36). These characteristics, which may facilitate the identification of the core category, are shown in Box 6.4.

BOX 6.4 CHARACTERISTICS OF THE CORE CATEGORY

The core category:

- is central, meaning that a lot of other categories are linked to it
- is the heart of the analysis: it indicates that the category accounts for a large portion of the variation in the behaviour in the data
- appears frequently in the data; otherwise stated, the indicators pointing to the core phenomenon must appear frequently
- is not easily saturated because there is so much material related to the core category
- can be formulated in a more abstract way, which can then result in the possible application to other fields of research
- facilitates analysis, makes the pieces of the puzzle fit together

Since the phase of selective coding marks the end of the analysis phase, it is useful to answer the questions listed below. When an answer can be given, it shows that you understand how the pieces of data fit together:

- Which themes have turned up repeatedly in the observations?
- What is the main message that the participants have tried to bring across?
- How are the various relevant themes related?
- What is important for the description (What) and the understanding (Why) of the participant's perspective and behaviour?

Researchers looking for explanations for the observed phenomena will probably ask the following questions as well (Lofland & Lofland, 1995):

- Under which circumstances does phenomenon A emerge?
- What facilitates experience B?
- What influences phenomenon C?
- When is event D absent?

Although by some, selective coding is mainly considered as the hunt for the core concept, the synthesis of data by establishing relationships is in itself just as vital as the selection of the core category. Connecting is also necessary when one chooses not to work with a core category and to focus on the main concepts instead. Reassembling is initiated during axial coding, but gains importance as the research progresses towards the final stages. In deciding which categories will stand out in the findings, the researcher can use the following elements as guidelines:

1. Research question and purpose: the most influential factors in determining how the data will be integrated and what the findings will look like.
2. Literature: the results are contrasted with the relevant literature and demonstrate how the sensitizing concepts have functioned.
3. Data: the outcomes are guided by what stands out in the data in terms of richness and the insights they have yielded.
4. Fascination: the surprising, fascinating and original parts should be included in the findings.
5. Actuality: the results occasionally grow in value if they fit the actual context of societal and scientific debates or events.

So far, attention has mainly focused on discovery of relationships between concepts. A second activity in reassembling is checking the relationships between the concepts with new data. Data have to contain the evidence for the asserted claims. In order to substantiate the conjectures, the researcher may initiate a new round of data collection. In analytic induction, testing the propositions with new research material is important right from the start (see Chapter 5). Theoretical selection (see Chapter 2) and in particular the analysis of negative cases provides the researcher with sufficient evidence to verify whether conjectures are justified. The crucial question is whether all the data, including the negative cases, can be described and understood within the conceptual framework.

There are several pitfalls in establishing and verifying connections between phenomena. When the researcher is not alert to these errors, they threaten the validity of the results and the eventual conclusions. Diverse procedures to ascertain quality of the research, including the interpretation of the data, are addressed in Chapter 9. Box 6.5 depicts some areas of possible misinterpretations that researchers should be aware of (Miles & Huberman, 1994).

BOX 6.5 POSSIBLE CAUSES OF MISINTERPRETATIONS WHEN CONFIRMING FINDINGS

- An unbalanced selection of the data: important information is missed and other information is weighed too heavily.
- Overemphasis of the first data or factual and dramatic events that made a lasting impression on the researcher.
- Selectivity, especially when the researcher is trying to validate the most important finding (and would rather not see it disproven).
- Interpreting co-occurring events as connection or even causality.
- Having trouble with estimating the value of various sources of information.
- Drawing unwarranted conclusions based on the frequency with which events occur.
- Putting a hypothesis aside too easily based on certain information.

FAQ 17

Q: During selective coding I have the feeling I am not 'coding' anymore at all. I feel as if I am doing all kinds of different things and I keep thinking about the data. I have not touched my computer for days, other than for writing texts, and I have neglected to update my coding scheme. Do I need to do that?

A: When your thinking goes on independently of the software support, congratulate yourself! Do not feel compelled to change the contents of your computer project to follow what you have already done. Other tools (see FAQ 16) might enable you to think outside the coding scheme. Unless for obvious reasons you want the code tree to reflect your latest work; for instance, when you need to present your provisional knowledge to others as a code tree, a matrix, or a map, or when you want to continue using the computer from where you are now and cannot do it otherwise.

Arriving at the end of selective coding, we are far removed from 'putting words in the text margin'. In the final phase of the analysis, the definitive findings are shaped. Instead of the chronological order in which the data were originally collected, the data now stand in the order implied by the data and the research questions. This new order is the surplus-value of the analysis. When the analysis provides a fresh, theoretical look at the phenomenon under study, the findings might be a source of inspiration and an invitation to the reader to think about and pose questions about the meaning of the findings.

BOX 6.6 SUMMARY OF SELECTIVE CODING

Purpose:	Determining important categories and possibly a core category, formulating the theoretical model, reassembling of the data in order to answer the research question and realize the research aim
Phase:	End phase of the research
Activities:	Determining core concept(s), determining relationships between the concepts and verifying them, writing, interpreting and positioning findings in the existing literature, thinking about the answers to the research questions and drawing conclusions
Starting point:	All available research material
Result:	Description of the most important concepts, coherent story in which the relationships between concepts become apparent, answers to the research questions
Validation:	Saturation, meaning that new data collection provides data which are consistent with descriptions thus far and fit the theoretical model

Reflections on analysis

Let us take a step back and look at what we have been doing this chapter. Thus far, this chapter has discussed two activities which, when taken together, form the core of the analysis. One, the segmenting of the data. This is largely achieved by means of open coding. The researcher puts markers into place which signify interesting parts in the research material. Two, the reassembling of the data in light of the problem statement. This is mainly achieved through selective coding. Axial coding occupies a certain in-between position and bridges segmenting and reassembling. In all three types of coding, thinking and doing go hand in hand.

Coding basically fulfils two functions. First, coding is important for data management. Interviews, field notes and other materials are generated throughout the research, and can be stored chronologically. But the researcher needs to develop an indexing system, to store the data based on the contents, from which the data can easily be retrieved. Coding supports this by reducing the amount of data, since only relevant sections are selected for storage and are assigned a summarizing and meaningful code. Coding also ensures that the data are filed in a way that enables easy retrieval: data are archived by code. In this manner, a code is a reference to passages in a text which relate to a common theme.

Second, coding is also important as a way to explore and interpret the data. It is a tool for mining a new research area. After all, coding is a way of constantly interrogating the data and asking questions about the meaning of the data. Coding is also the driving force behind categorization and conceptualization. Coding forces the researcher to generate categories and to be clear about what distinguishes them from each other in relation to their properties, as well as clarifying which indicators there are for each category to occur. Finally, coding plays a role in the reassembly of the data, so that the data are looked at from a new perspective and the research questions can be answered.

In the spiral of analysis the three types of coding are depicted as linear stages, but in practice the three activities are not neatly distinguished. It is possible that some areas of the field still need further exploration, while other areas are already subject to axial coding. Open coding and axial coding often merge into one another, as do axial and selective coding. It is does not matter which activity you are engaged in at a given moment, as long as you work systematically from the data to the findings.

The qualitative research process is characterized by alternating between data collection, data analysis and sampling. These activities cannot be strictly separated, as the researcher jumps backwards and forwards between them. Simultaneous data collection and analysis is seen as the backbone of the spiral of analysis (Figure 5.4). Its procedure is referred to as cyclical, spiral-shaped, iterative or recursive. In contrast, quantitative research is characterized by separate phases of data collection and data analysis; the two phases take place one after the other, also known as a 'linear' research process. Towards the end of the qualitative research process, data collection decreases in favour of analysis. An abstract representation of the models of qualitative and quantitative research is given in Figure 6.7.

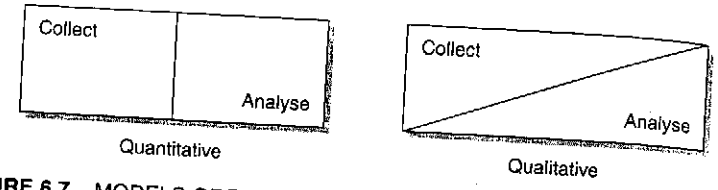

FIGURE 6.7 MODELS OF THE QUANTITATIVE AND THE QUALITATIVE RESEARCH PROCESS (MAYS & POPE, 1995: 184)

There are several reasons for adopting a zigzag style of data collection and analysis approach. Working this way contributes to the explorative capacity of the research. Each stage of data collection is informed by the analysis of the previous data collection stage. The topic list or any other measuring instrument can be revised for the next round of data collection to be more specific and focused. As the amount of data grows, new paths can be followed. This could even lead to a fine-tuning of the research questions (see Chapter 2) and ultimately lead to more relevant outcomes. The principle of purposive sampling and in particular theoretical selection (see Chapter 2) supports this way of working, since the interim results of the analysis determine where the researcher will collect new data.

The interchange between data collection and analysis not only benefits the explorative aim of qualitative research, but also its explanatory aim. Conjectures and propositions about the connections between phenomena can be tested for accuracy when the researcher collects new data from new cases. This testing of provisional findings against new information can lead to modification, rejection or elaboration of the first ideas. This is part of analytical induction (see Chapter 5). The newly gathered data is compared with everything that has been learnt from the data that was collected and analysed previously – concepts, conjectures or assertions. What was thought to be a known result or assumption is then adjusted accordingly.

Finally, there is a third much more practical reason for the data collection and analysis to proceed simultaneously. Imagine returning from field work carrying tapes (or digital recordings) of at least 20 interviews, plus some exercise books full of field notes. Where would you start? Would you listen to the first interviews, already knowing their content? Would you instantly start writing the results from memory without having the basic material in front of you? Either method would double the work and most likely lead to frustrations. Postponement of analysis only adds to the pressure. The yield of the research will most likely be better if analysis is started right at the beginning, because the exploration and verification opportunities of an emerging design can then be utilized to their fullest potential.

Readings I learnt much from

Charmaz, K. (2006). *Constructing grounded theory. A practical guide through qualitative analysis.* London: Sage.

Jorgensen, D.L. (1989). *Participant observation. A methodology for human studies.* Newbury Park, CA: Sage.

Lofland, J. & Lofland, L.H. (1995). *Analysing social settings. A guide to qualitative observation and analysis.* Belmont, CA: Wadsworth.

Ritchie, J. & Lewis, J. (2003). *Qualitative research practice. A guide for social science students and researchers.* London: Sage.

Doing your own qualitative research project

Step 6: Open, axial and selective coding

1. After collecting the first data, the analysis can be started. Start open coding after the first two or three sessions of data collection. Although the sequence of the activities involved in this series suggests a linear process, it is emphasized again that data collection and data analysis alternate.
2. Have you decided upon the use of software yet? If you have decided to use a package, try to choose one that suits your purposes and start using it right from the beginning.
3. Do not forget to write memos as you go along. Write memos on methodological issues as well as on the contents of your project, in particular the development of categories and concepts.
4. Stay close to the data when conducting open coding. When you have reached saturation after a few rounds of data collection (see Box 6.2), you will automatically progress to axial coding. It is highly likely that some areas within your project will still need open coding as well.
5. If you work in a research team, use each member's capabilities and knowledge in coding and memo-writing.
6. When you have defined and delineated different categories, you should begin selective coding. Start processing all the assumptions and theories that have accumulated during the research project up to this point. Do use all the work of earlier phases. This is by far the most difficult stage of the research but also the most rewarding. Most importantly, you should capitalize on your theoretical sensitivity that can be stimulated by the integrative procedures presented in the next chapter.

7
INTEGRATIVE PROCEDURES

The integration of the data into a coherent, analytical format is a difficult part of every qualitative research project, even with the use of methods and techniques such as coding and memo-writing. Ten integrative procedures and heuristic devices are presented in this chapter to help you to extract the interesting elements from your data. They should be seen as thinking aids, and examples are given here as inspiration. To illustrate the decision-making processes during the analysis, we present a 'think aloud' report of a real-life analytical procedure. Modern software packages are immensely useful for coding and performing the more mechanical tasks, and are a useful tool when interrogating the data and heuristically thinking about the material.

LEARNING AIMS

By the end of this chapter, you will be able to:

- Pursue integrative procedures that benefit your analysis
- Create diagrams and know how to present them
- Capture social processes in a mould encompassing chronological stages
- Work out whether and how searches and counts are useful in your own project
- Incorporate the influence of group interaction in the analysis of focus group data
- See how visuals are used as a stand-alone method or as an elicitation technique
- Value code-and-retrieve and theory-building software tools in your project
- Reflect on the pros and cons of computer-assisted data analysis

Heuristics for discovery

The reassembling phase in analysis is without doubt the most difficult phase in the entire research process. It is this phase that requires the researcher to be theoretically

sensitive, resourceful, open-minded, skilful and well-informed on the literature. No ready-made solutions are available for integrating the data. The literature on qualitative analysis does not tell us much about this analytic stage either. This is probably because this phase of the investigation is unique for all research projects and intrinsically bound to the contents of the specific research project. It is also generally assumed that the capabilities needed in this part of the project are the hardest to acquire. According to Cutcliffe and McKenna (2004), it is only the experienced qualitative researcher who can recognize the over-riding or underlying dynamic, principle or process in a data set. It is precisely this proficiency that is difficult to articulate or teach.

This so being, there are several methods and devices to help solve these problems. These 'heuristics' are rules of thumb, thinking models and creative instruments to turn your data around and look at them from different angles. The ten integrative procedures and heuristic devices presented in this chapter are:

1. Creating visual displays
2. Looking for the core category
3. Using moulds
4. Creating a matrix
5. Reading your memos
6. Searching and counting
7. Acknowledging the presence of feelings
8. Formulating a typology
9. Constructing arguments
10. Outlining the main message

No universal solutions exist for the reconstruction of data. There are, however, different kinds of thinking aids which may be applicable to your work. They are discussed in this section and in order to clarify them, examples are included where possible.

Creating visual displays

According to the motto 'You know what you display' (Miles & Huberman, 1994), drawing figures, flow charts, maps, decision trees and Venn diagrams (sets of interlocking circles) is an excellent way for researchers to test what they know and do not yet know regarding their topic. Creating visuals will raise awareness about the possible gaps in the available knowledge and stimulate thinking about how to fill these gaps. Maps and diagrams may help develop interpretations. In a project that is informed by social theory from the outset, a map may offer insight in all the adjustments and refinements made to the theory based on the current research.

TIP

Modern software can be used to visualize connections in maps and diagrams. The mapping tools enable a different and often more complex set of linkages that go beyond the hierarchical coding scheme (Lewins & Silver, 2007). Most packages allow an integration of the source data – documents, memos and codes – and the maps and support a visual way to rethink connections.

If a diagram is solely an aid for the researcher it is not necessary to include it in the presentation of the findings. However, a visual display may be added to a presentation and to the final report when it clearly illustrates the structure of the findings, most commonly a proposed theory. A diagram enables the researcher to present a lot of information in a concise manner. Figures must be kept as simple as possible, they should be analytically sound, and all elements should be clear, logically connected and simple to follow (Morse, 2006).

Graphical representations of both interim results and final results are common (Strauss, 1987). Examples of such depictions were given in Figures 5.1 to 5.3 about the quality of care in nursing homes. The diagram depicted in Figure 7.1 is taken from the study on alcohol use among adolescents during their social gatherings and night life (Engineer et al., 2003). It is an example of how a model can summarize a lot of information. The fishbone diagram depicts the most important risk factors for alcohol abuse and the resulting criminal and disorderly behaviour. While reading the text of the original report, the reader moves logically through the model.

Looking for the core category

Reassembling the data may be achieved by seriously looking for core conceptual categories. It is possible that one category is very central but remains unnoticed until one actually looks for it. It has to be a category which matches all of the characteristics described in Box 6.4. The benefit of a core category lies in the fact that the researcher may continue analysis from the new perspective, as if all pieces of the puzzle fall into place. Then the 'old' material and earlier steps are recycled in terms of the newly found theoretical angle. It is not possible or desirable to appoint a core category for every research project. In some studies, multiple categories may claim a leading role.

In the study on prolonged use of sedatives, Haafkens (1997) appointed 'rituals of silence' as the core concept. It became the most important social pattern of what led women to continue using tranquillizers over a prolonged period of time. As time went by, the study tells us, the doctor stopped asking whether the tranquillizers helped and the women stopped talking about or asking for feedback or opinions. The people in the women's social network who knew about their use would not offer their opinion. This allowed those involved to go on living under the assumption that the drugs were still working and not producing any negative effects. This mutual silence largely explained the prolonged use of tranquillizers and became the main theme in the study.

Using moulds

In Chapter 5 we touched upon the moulds or 'coding families' which were developed by Glaser (1978) in order to structure the data. The coding paradigm of Strauss and Corbin (2007) – a mould consisting of context, conditions, interactions/strategies

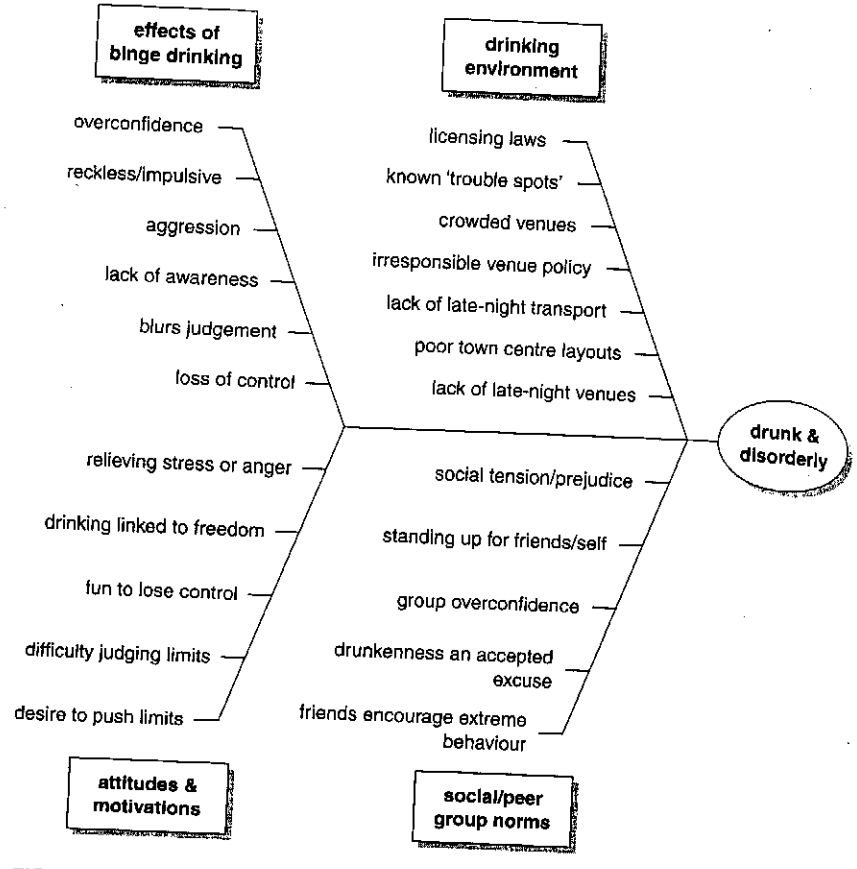

FIGURE 7.1 KEY RISK FACTORS OF ENDING A NIGHT OUT IN DISORDER (ENGINEER ET AL., 2003)

and consequences – was discussed in Chapter 6. A different mould may incorporate processes, stages or phases. Since qualitative researchers are often interested in processes and changes over time, a mould incorporating stages would be useful in dealing with the different materials.

A well-known example is Becker's phase model, in which he described the stages that a person should go through to start finding the use of marihuana pleasurable (Becker, 1982). He deduced that no one will become a user without:

- learning how to smoke the drugs in such a way that the effects are noticeable
- learning to recognize the effects and attributing them to the drugs (learning how to get high)
- appreciating the experiences.

Haafkens (1997) was inspired by this research, and distinguishes between five phases in the use of tranquillizers:

- The period which leads to the first prescription.
- The first six months of the use on prescription.
- The last six months of taking the tranquillizers.
- The decision to stop.
- The stopping in itself.

TIP

Some software packages enable the tracking of a sequence of events that is talked about throughout the text (Lewins & Silver, 2007). The tools allow hyperlinking between points in the text. You can either link several segments or pairs of segments. A sequential model (and other models as well) can also be visualized using the mapping tools of the software packages.

Creating a matrix

Miles and Huberman (1994) recommend the use of a matrix, as creating it facilitates thinking about the linkages between the various categories. A matrix displays whether and how phenomena appear when other phenomena are also present. A matrix may have a very simple shape. Whether a trait is present or absent or whether a phenomenon does or does not occur is noted in the rows and columns. In order to depict this, a category is divided into parts (see Chapter 6). Thus, connections become apparent. A matrix that looks simple can still be very informative and labour-intensive to create.

An example of a matrix about causes of aversive behaviour and feeding issues of nursing home residents is shown in Table 7.1 (Pasman, The, Onwuteaka-Philipsen, Van der Wal & Ribbe, 2003: 311). The rows contain the reasons why nursing home residents do not eat and the columns specify the domains to which the reasons apply. The cells contain words or short descriptions. This matrix is far from simple and summarizes an important part of the findings.

TABLE 7.1 FRAMEWORK OF CAUSES OF AVERSIVE BEHAVIOUR AND DOMAINS OF FUNCTIONING

	Domains		
Causes	**Physical**	**Psychological**	**Social**
Not able to eat/drink	Difficulty in swallowing Paralysis Acute illness	Grief Apathy	Distracting environment Insufficient time for feeding
Not understanding the need to eat/drink	Visual and audio problems	Apraxia Agnosia	Insufficient nursing skills
Not wanting to eat/drink	Declining/loss of appetite Afraid of choking	Not wanting to live any longer Disliking food	Unpleasant environment

Source: Pasman et al., 2003

A matrix may be used to document a profile of participants. In this manner, Haafkens (1997) summarized the life histories of the women in the tranquillizer study. The rows contained the interviewed women. The columns contained the following information: the year of the first use, the age of the woman at that moment, the most important biographical event, the reason for using the tranquillizer, the role of the doctor, the role of others, the nature of the medication and the pattern of use. The life history of a person is found by reading through the matrix from left to right. Reading through the matrix from top to bottom yields the variation in content of a phenomenon, for example the various reasons for using tranquillizers.

Reading your memos

Theoretical or analytical memos are created during the project to record any ideas that are related to the interpretation of the data (see Chapter 4). They reflect attempts to organize the data in a way to derive meaning from them and record the development of the interpretation. They allow researchers (and others) to track down the development of the concepts. In the memos is described, for example, how a particular observation has been interpreted, why codes are split or merged, how a sensitizing concept has slowly gained empirical weight, with which colleagues the interpretations are discussed and so on. A theoretical memo might contain just one statement or it might comprise several pages of relevant information.

The reading of memos helps to track the various categories and follow or understand the conceptual progression of the findings. Memos are particularly valuable in providing a chain of reasoning (Cutcliffe & McKenna, 2004): the 'evidence' found in the data, the inferences drawn from the data, the choices for the main categories and their properties and so on. Your chain of reasoning should be sound and convincing. Memos can then remind you of the development history of the various categories and the relationships between them, and alert you to missing or yet incomplete elements in your line of thinking.

TIP

When you use software in your project you can attach memos to different elements of the project, such as documents, codes and text fragments. Memos can be managed in a systematic and useful way, and their contents can be searched and retrieved. It might be useful to write memos as standalone documents in your project, especially for the overarching themes and patterns that are discerned at the end of the research (see Chapter 6). A memo not attached to other sources has the advantage that you can assign codes to them and use all the functions that are available to 'normal' documents (Lewins & Silver, 2007).

Searching and counting

Software packages offer the option of searching for words or codes either in a section of the data set or in the entire data set and determine their frequency. Two types are distinguished: searching and counting for specific words or expressions as they are used by the participants, and searching and counting how often a specific code has been used.

TIP

Most software packages offer tools to easily search for words and view them in context. They also have tools for counting words and codes. Most programs allow you to search whether several codes co-occur in the data and whether certain words are used in each other's vicinity.

First, there is searching for particular words as they occur in the texts as a way to interrogate the data. For instance, in the example study into the quality of care in nursing homes (Chapter 5), staff members often spoke about 'attention' paid to residents. To verify whether 'attention' was indeed often used in the interviews and in the field notes, the empirical material was searched for the word 'attention' and the hits were counted. When the occurrences were reviewed in context, it was confirmed that 'attention' was an important activity that occurred under several conditions and had several definitions attached to it.

Counting can also be used to explore large data sets and to compare the use of certain words and phrases in different groups. One example of counting words is keyword analysis, illustrated by Seale, Charteris-Black, Dumelow, Locock and Ziebland (2008). The authors compared joint interviews with individual interviews with men and women. One assumption was that in mixed-sex joint interviewing the perspective of one gender will be favoured, i.e. that of the overbearing male informants who see themselves speaking on behalf of the couple. To test this hypothesis the authors counted how often the singular pronouns ('I', 'me' and 'my') were used, as well as plural pronouns ('we', 'us' and 'our'). Analysis of the individual interviews showed that men were more likely than women to use plural pronouns that referred to both themselves and their partner. Women, on the other hand, used the individual interview to speak about their own personal experience, reflected in their more frequent use of 'I', 'me' and 'my'. In joint interviews 'we' became used more frequently by women to indicate inclusion of their partners, while no such change occurred for men. Keyword analysis is not a qualitative analysis in the strict sense, but it shows that counting should not be dismissed immediately, especially when large data sets need to be scanned for interesting subjects.

Second, there is searching and counting for the number of times a certain code has been assigned to the whole or parts of the data set. Due to the nature of the data collection, we have to use these frequencies with even more caution. The content and time spent on each topic varies according to the questions asked and the participant's own salient concerns. So when counting codes, you count your own concerns

as well as the participant's. Your concerns can change during the research process as you come to understand the topic being studied. As a consequence, the outcome of a count will not lead to a reliable estimate of the true prevalence of the various social phenomena, and not even of their weight.

Then again, during analysis it is the researcher who assigns the codes, so that information stored will be retrievable. Generally, the frequencies with which the codes are used will give additional evidence. Instead of relying on implicit notions of frequency only, counting how often a theme appears in the transcripts or field notes may demonstrate that a theme is weighty and be considered supportive evidence. Here is one example of my research on veterans who developed post-traumatic stress disorder (examined later in this chapter). I coded for positive and negative experiences during their deployment in Cambodia. Negative aspects for instance consisted of 'bribing', 'witnessing misery', 'poverty', 'cruelty' and 'disparaging human life'. Positive aspects were 'beauty of nature', 'resilience of local people', 'children', 'hospitality' and 'travelling'. When I counted the fragments that were assigned these codes, I noticed a ratio of 3:1 for the negative aspects in relation to the positive ones.

The outcome of this counting was used with caution. The fragments I counted were all of a different nature and had a different impact. Is 'travelling' of the same nature as 'bribing'? Is that a 1:1 ratio? Maybe only one negative, unpredictable life-threatening event could spoil the entire exercise. What I want to demonstrate with this example is that you cannot rely on counting only, you have to interpret what you are counting in the context of the particular source that your segments are taken from and in the context of all your sources, all your analytical efforts, and all your interpretations.

Acknowledging the presence of feelings

Researchers' feelings during field work, such as embarrassment, unease, fear, guilt, anger and sadness, can be used in analysis. Feelings were previously discussed in Chapter 3, when researcher's stress during data collection and data analysis were addressed. We have learnt to ignore and suppress feelings because they are thought to bias our data collection and analysis. However, feelings can serve as heuristic instruments when one tries to become aware of them and tries to find out what possibly could have caused them. Feelings can be used to your advantage, especially when trying to place yourself in the shoes of another person (Harris & Huntington, 2001; Kleinman & Copp, 1993). When emotions are acknowledged they can be used as analytic tools.

It is possible to feel upset, for example when veterans tell you about the injustice and misery they observed when deployed in a war-zone area. You will be affected each time you read the transcript about mine incidents with children. By using your emotional reaction as an analytic tool you can try to understand how others feel, and why (see the next section). Your own feelings then serve as a means to see the world through the eyes of other human beings, who try to share their experiences with you. Your own feelings might mirror those who hold a particular role in the setting you are studying (Kleinman & Copp, 1993). Feelings can alert you to ask many questions about what it is that upsets veterans, what strategies they use to deal with

their emotions, how they protect themselves against becoming traumatized, and how they deal with emotions in a culture perceived to be tough. Feelings should not be considered inappropriate in field work; rather they should be seen as advantageous in understanding the data.

There are times during analysis when it is appropriate to respect and acknowledge your own emotions, and these should be included in the field notes or memos. However, it is still comparatively rare for authors to include descriptions and analyses of their own intense emotions experienced during field work in the final report. This has to do with the common writing styles used in scientific reporting, although some writing formats allow you to write about your situation in the field and your personal background (see Chapter 10). The disclosure of feelings can be considered helpful when analysing qualitative data.

Formulating a typology

Qualitative research may result in the formulation of a typology. The four awareness contexts of Glaser and Strauss (1965) for instance, constitute a typology (see Box 1.4). The construction of a typology stimulates the analyst's imagination. In Chapter 6 we identified that two types of analyses are commonly distinguished: issue-focussed or cross-case analyses and case-focussed or within-case analyses. Until now we have mainly concentrated on issue-focussed analyses by using codes to find common themes through all collected materials. Classification into types helps identify connections between different phenomena within one case.

Types can be constructed on the level of an individual, for instance focussing on the beliefs, intentions, behaviour and the consequences of behaviour in women using sedatives. They can also be constructed on the level of groups or organizations, such as the relationships between culture, degree of organization, traditions and involvement in violence of football club supporters. This kind of connecting may illuminate how several elements of a case are linked and may show the flow of events over time.

TIP

To generate a typology it is useful to code relevant segments that pertain to a certain type. When using software, consider the use of hyperlinking tools to link distinct parts of the type within a text (the columns in Table 7.2) or to link the values on the criteria in all the available texts (the rows in Table 7.2). A useful software tool is the overview of codes that appear in one document. This index of coded segments per document provides a text profile that can help you to focus on the specifics of a type.

Types are useful tools for readers as well. First, because the names of the types usually reveal a lot about their contents. In most cases the types are easily recognizable and easy to work with. Second, a typology allows readers to compare the types

with each other as well as with empirical cases in real life – organizations, persons, activities, situations – in order to trace differences and similarities. These comparisons provide knowledge on these cases and may provide clues about the reasons of possible deviations from the types or about the consequences attached to different types.

TABLE 7.2 TYPOLOGY OF TWO PERSPECTIVES OF PEOPLE WITH MULTIPLE SCLEROSIS

Criteria	Perspective	
	MS will never have the upper hand	I have nothing to do with MS
Feelings about MS	Confident, realistic	Bitterness, powerless
Identity	Integration of illness in life	Loss of former identity
Coping strategies	Normalizing, managing	Hiding, banishing
Dealing with the family	Openness, support	Isolation, distraction
Dealing with help	Organizing, control	Rebellion, resistance

Source: Boeije, 2002

The typology in Table 7.2 stems from research on the perspectives of people who have the chronic degenerative illness multiple sclerosis (MS). Most interviewees were found to live by a certain motto. These mottos were used as titles for different types which were distinguished in the data and of which two are described as an example here (Boeije, 2002). Putting a typology together forces one to think about the criteria by which different cases can be distinguished, whether situations, individuals, organizations, groups or activities. These criteria are put into the rows of a matrix and the cases in the columns. The values for each of the examined cases are systematically filled out in the cells.

Constructing arguments

This section describes the use of logic as a heuristic device for analysing your data. Logic has to do with reasoning, with what is considered proof and with drawing inferences (Cutcliffe & McKenna, 2004). The reader should be able to follow the progression of the findings in the research report, i.e. to trace the path from the data to the inferences. Cutcliffe and McKenna remark that one person's chain of reasoning may well be different from another's, yet both can be logical. These different chains of reasoning may even produce different conclusions. To think about your chain of reasoning during the reporting phase is a good exercise in thinking about your data. In your report reasoning takes place on at least two levels.

The first level is the logic you apply to the participants' accounts. While doing your research you try to dissect how participants view their world and how this view relates to their actions. When you claim for instance to have 'found' a type – 'MS will never have the upper hand' (Table 7.2) – this claim should be supported by arguments, such as observations, interpretations, regularities and theories. To substantiate your logical reasoning you have to ask all kinds of questions: What are the participants' arguments

for behaving in a certain way? Are there any contradictions? What do they take into account? How do they weigh different aspects? Do they perceive these aspects as within their control? Do they have different experiences as well? In other words, your claim as a scientific writer must be based on plausible arguments (Smaling, 2002).

TIP

To support this heuristic aid, you can use the hyperlinking tool in the software to link contradictory statements, points in the data where strong arguments are demonstrated or where doubts arise.

The second level of reasoning concerns the report as a whole in which you are constructing an argument. A strong argument persuades the reader to accept the writer's viewpoint (Charmaz, 2006). The most important argument is the overall message of your research: the one that elaborates the salient contribution. Therefore you can go back to the elements of your research proposal, but you cannot copy and paste them, because now that you have carried out the research you can add the insights you have developed into the argument. We know that in a scientific piece of work we have to present a straightforward research design including the gathering of adequate data, the conducting of a sound analysis, and the construction of the overall argument and the logical reasoning it involves.

All chapters, all sections and all paragraphs consist of smaller arguments. Each paragraph has a core message comprising a main issue and some side issues. This core message is connected to your argument in the section. And the argument of the section builds up the argument of the entire report. All of them are connected. The point you want to make in each part should be clear. Charmaz is helpful in formulating four questions to help you find arguments (Charmaz, 2006: 157):

- What sense of this process or analysis do you want your reader to make?
- Why is it important?
- What did you tell your readers that you intended to do, and why did you tell them that?
- In which sentences or paragraphs do your major points coalesce?

Charmaz suggests that working on the data, taking notes about an emerging argument, talking to colleagues, talking to yourself or thinking aloud may help formulate your initial arguments. Or take a rest and sometimes the right reasoning comes to you ('That's it!').

Outlining the main message

Suppose that the deadline to hand in your final research report is imminent. How would the main argument of the story be formulated and what would be the

message to the reader? This question forces you to think about the preliminary table of contents and in particular the contents of the results section (see Chapter 10). When writing different drafts of your results your ideas and especially the order or logic of the presentation will develop (Jorgensen, 1989: 120). Jorgensen sees an outline as a heuristic device that helps you order your ideas and prepares you for writing the final report:

> Outlines may be very useful once you have notes and files in need of further organization. Outlines, however, are purely heuristic devices. You need an outline only to get a handle on the possible arrangement of ideas or an argument: The outline serves temporarily to organize your thinking and provide a definitive direction to the work. Be prepared to deviate from your outline.

Below is an example of the table of contents from the study by Cutcliffe et al. (2006) on the support of psychiatric nurses for suicidal people.

BOX 7.1 EXAMPLE OF A TABLE OF CONTENTS

1. Introduction
2. The study of suicide: a brief historical overview
3. Research method and design: modified Grounded Theory
   - 3.1 Ethical considerations
   - 3.2 Selection of the sample and data collection
4. Data analysis
5. Findings
   - 5.1 Core variable – 're-connecting the person with humanity'
6. Stage one: reflecting an image of humanity
   - 6.1 Experiencing intense, warm, care-based human to human contact
   - 6.2 Implicitly challenging suicidal constructs as a results of encountering contrary experiences
7. Stage two: guiding the individual back to humanity
   - 7.1 Nurturing insight and understanding
   - 7.2 Supporting and strengthening pre-suicidal beliefs
   - 7.3 Encountering a novel interpersonal, helping relationship
8. Stage three: 'learning to live'
   - 8.1 Accommodating an existential crisis, past, present and future
   - 8.2 Going on in the context set by the existential relationship with suicide
9. Discussion
10. Conclusions

There are sufficient aids to properly round off the final stages of the analysis. Of course, these aids do not need to be applied all at the same time. You should use them when they fit your data and only if you are convinced of their value. In choosing heuristic devices, a personal component is involved: when you are always sketching you will probably be drawn into visualization, when you are sloppy you will find it difficult to read and use the memos (if you can still find them), and when you enjoy seeing a bigger picture you will probably like to construct a matrix and an outline.

Think aloud analysis

Although you can learn a lot from books, some things are best learned by watching someone doing the job. Therefore I recorded my thoughts during an analysis to generate a 'think aloud' report. I would like to show the type of mental activities I was engaged in while analysing qualitative interviews with Dutch veterans who had been deployed to Cambodia on a peace mission. Half of the group had been diagnosed with post-traumatic stress disorder (PTSD), while the other half had not. I did not conduct the interviews myself. I started out with ten interviews of veterans who had been diagnosed with PTSD. The research question formulated for the entire project was 'How do Dutch veterans attribute meaning to their deployment on a peace mission in Cambodia and what are the similarities and differences between veterans diagnosed with post-traumatic stress disorder and the ones not diagnosed?'

Data management

The first thing I am going to do is rename the interviews of the group diagnosed with PTSD. They have these complicated numbers like 31–287 that even after some time do not ring a bell and I keep forgetting who is who. The texts are numbered as simply as possible with an abbreviation of an Interview Diagnosis Number (ID-01, ID-02 and so on). I generate a project in MAXQDA2007, the software I am working with, and import the texts. I suspect I will be using different codes for this group and for the group that is not diagnosed, and therefore decide to analyse the two groups one after the other so that I can concentrate on each group separately. Afterwards the two can be compared and combined.

Start open coding

The experience of control during deployment seems to have two dimensions: mastery on the one hand and powerlessness on the other hand. I decide to merge the codes 'comradeship' and 'collegiality', and decide to use the code 'comradeship' because this expresses the meaning of what the participants tell me. How they deal with each other while on a mission seems to be a mixture of working together

professionally and friendship. The relationship between soldiers and officers is sometimes described as the formal soldier/officer relationship in which the officer is clearly the superior of the soldier, but sometimes the relationship is described as equal.

While analysing ID-04 I have to generate new codes because he gives a lot of information about the nature of the mission and what he perceived to be their ambivalent task in Cambodia. Because they were on a peace mission they were not allowed to fight and shoot except when seriously threatened themselves. I discover that many men in this group have some sort of conflict with their (former) employer, the army. I decide to generate an attribute (a variable) to cover the current relationship with the army either as neutral/all right or as a cause for conflict. Later on when I add the interviews with the veterans without PTSD I can create a matrix and compare the groups.

Half-time score

After coding four interviews of around 80 minutes each, I have generated around 75 codes organized in nine themes. On average I have assigned around four fragments per code. I tidy up the code tree and order it in a logical sequence. It now runs from the motivation to join the army, to motivation for deployment, their experiences in Cambodia, their home front, the comrades, returning home, and current experiences.

Theoretical memo 1

I remember having glanced over an interview with a veteran who did not develop PTSD in which he was talking about 'just doing your work' in a kind of instrumental and pragmatic way. In this group there is an interviewee who also says 'you just do your job' but he seems to use it in a different meaning. It is a demoralized expression of duty, a way of expressing that you just do your duty without thinking about it in order to get the job done. The veteran without PTSD seemed to use the word 'job' as employment, like any other job.

Theoretical memo 2

In the meantime I have studied an article about self-enhancement in the context of trauma and recovery, since I felt troubled by my lack of knowledge on the subject. Victims of danger and fear often feel guilty or blame themselves for not having done enough. Their confidence has to be restored, and this is what Taylor (1983) refers to as 'self-enhancement'. The interviewed veterans seem to restore their confidence by making comparisons, for instance with colleagues who are going through a divorce or even some who committed suicide. At first I did not grasp the meaning of these comparisons, but now I see them as attempts to restore confidence. In comparison with these other veterans they are doing quite well.

Continuation: open coding and axial coding

While coding ID-06 I still need new codes, which means that the exploration is still going on. For example, about carelessness, awareness of danger and risk, about exhaustion, and about the macho culture of the Marines. I am a bit wary of creating new codes because I already have so many of them.

Now coding is speeding up. I feel the need to find out where I stored all the fragments and to check if I have worked systematically. I guess I have not. I refrain from checking because it takes a lot of time. I am confident that when retrieving segments and writing, I can filter out the useful fragments.

Theoretical memo

While coding ID-06 I notice that the veterans confide mainly in male persons (sometimes before their partners), like an uncle, a father-in-law, a grandfather or a male friend.

I discover that the veterans are impressed by the children they met in Cambodia. Are they a symbol of innocence and disarmament? When something happens to these children it is experienced as very painful. They attach positive and negative emotions to children in general. To check this conjecture I do a lexical search to find the word 'children', and indeed children are mentioned in all interviews except one. The exception or negative case is an interview with an interpreter who had a different task and role than the other interviewees, and during the interview mostly concentrated on the negotiations with Cambodian rulers.

Methodological memo

The idea of bracketing or the role of presuppositions in the analysis becomes meaningful. It would have been a nice experiment to see whether I could have detected the interviews with the veterans diagnosed with PTSD and the ones without. I think there is only one interview in this group that would have left me in doubt.

Finding a core concept

I have read Janoff-Bulman's (1992) book, in which she builds an argument that extreme situations and life-threatening events can lead to shattered assumptions. She presupposes that we hold basic assumptions, namely that the world is benevolent and meaningful and that the self is worthy. A trauma can suddenly and powerfully threaten these assumptions of a safe and protected life and a just world. One response to threatening events is that the fundamental beliefs are destroyed. Of course the veterans in our study do not tell me that their 'assumptions are shattered' but there are many indications that they found themselves in a frightening place and that they are still filled with images representing malevolence, meaninglessness and self-abasement. In Cambodia they found themselves in an inconceivable, unimaginable world.

And then the puzzle is suddenly solved. With hindsight the veterans state that you cannot prepare yourself for deployment. This makes sense because you cannot prepare for the unimaginable. They all say their families do not and cannot possibly understand them. This is self-evident, taking into account that their families' basic assumptions are still intact and that they see the world through different eyes than the veterans in this group. The concept of shattered assumptions allows me to understand their accounts, and I am connecting several themes with the core concept, such as the training and preparation, the experience of danger and the ignorance they perceive when coming home.

Methodological memo

I am analysing ID-08 and I become a bit bored. With ID-10 I seem to have reached saturation in my categories. Now that I have analysed all the interviews of the veterans with PTSD I feel the urge to let it out and write it all down. So I start writing. This is really the best part of qualitative research.

Writing

The structure of the written report seems to make sense. I start out describing the life-threatening events and danger. This is central for the experience of trauma. Then I go on with the shattered assumptions because this is, I think, what happens after having experienced danger and fear. Next, I contrast shattered assumptions with the taken-for-granted world they used to know at home. Subsequently, I describe the negative things and the positive ones they encountered, and notice that they often go hand-in-hand but that the negative ones are much stronger. And I finish with dealing with family and friends when returning home in what once was a taken-for-granted world.

Each time I start writing again, I reread what I have written up till then to get my mind back into the project and see whether it still makes sense. Each segment that I retrieve is read and reread to see what is in it. When retrieving segments that belong to a certain category, I sometimes come across the same segment again. This is obviously because I coded it twice. When I retrieve a segment that I do not understand, I store it in a file to be read later on. When I retrieve a segment that I probably need in a different section, I copy the entire segment into that other section. I am afraid that otherwise I will forget the segment altogether, but it makes the concept of the manuscript very unclear and unwieldy.

It is important to describe the danger and life-threatening events that the veterans met, like shootings, minefields and accidents that involve children. I make inventories of the less important themes that are expressed. All veterans talk about the backwardness of the country and I just list what they mention, like poverty and the lack of asphalt roads, sewerage and appropriate health care. In this way all their impressions are used to give one image for the reader to have when interviewees refer to being in a primitive country.

Now that I have described the danger and life-threatening events, I come back to these issues to demonstrate shattered assumptions. I desperately need the expressions

of the interviewees, since the subject is difficult and it is also hard to indicate data to be evidence. After having processed the fragments that were assigned the most important codes, such as shattered assumptions, negative and positive aspects of being there, taken-for-granted world and danger, I can fairly well assess which other codes I will be needing, like distrust, ignorance and misunderstanding of family, adventure, growth, returning home and the Marines' culture. I am trying to connect all themes, which means that the different parts of the manuscript will be interrelated and that there is consistency in the piece of work.

Half-time score

MAXQDA2007 showed me that I had coded 654 fragments and when I have processed all the categories that I need for this particular manuscript I think I might have read half of them. Because I did the coding myself I know exactly what is coded and what I can save for later publications. For instance, the conditions under which these veterans went on deployment, their education and training, their motivation and relationships with their families before departure. A comparison between the group with PTSD and without will be very interesting and I am burning with curiosity to find out how they compare.

Special devices: focus groups and visuals

Focus groups

Focus group data are special because there are group dynamics involved. The individual might be influenced by the group and the group as a whole might be influenced by the individual. Although the analysis of focus group data follows the same processes as the analysis of other types of qualitative data to a certain degree, we have to be aware of the specific group composition as well as the individuals finding their own meaning within the group (Duggleby, 2005). How to incorporate the impact of the group context in the analysis of focus group data is a matter not yet resolved (Vicsek, 2007).

Ultimately, there seem to be three levels of analysis, namely the intragroup level, the intergroup level and the individual level. First, there is the group itself. For intragroup analysis it is useful to have a combination of two analyses. One is of the group discussion on a particular subject, and the other is a more fine-grained analysis of the subject itself (Kidd & Parshall, 2000).

When analysing the larger units of texts, attention can be paid to group interactions, such as levels of agreement, consensus, conflict, censoring and characteristics of the discussion. Attention should go to tracking individuals' contributions as well, to see if or how they are influenced by group interaction. In an intragroup analysis the effects of group composition and characteristics are identified, such as a very talkative member taking part or the use of a small group due to absenteeism (Morrison-Beedy, Côté-Arsenault & Fischbeck Feinstein, 2001).

When analysing for content, the focus can be directed at different points of view when addressing the subject. Counting how often a subject is addressed during a focus group meeting needs to be approached carefully because of the effect it may have on the group. For instance, one group member may put a topic that is prominent to him or her at the top of the agenda, or during the discussion participants may select items that they consider of interest or importance to the other group members, such as themes they think would be good to discuss together (Vicsek, 2007). So when a theme appears to dominate the discussion, it could be a sign of its significance but it could also be an artefact of the method used. It is up to the researcher to decide how the outcomes of counting have to be interpreted.

The intergroup level refers to the comparison of the groups with an eye to group differences, the shared experiences and topics in each group, and the reasons for this. Situational factors that can be taken into account are the characteristics of the participants, the moderator, the environment, the time factors and the content (Vicsek, 2007). By comparing the groups it becomes possible to say something about how these factors might have influenced the results of a group discussion. An intergroup analysis can determine whether there is consensus on the range of issues deemed relevant to the participants. A form of consensus across groups can be identified in terms of the range of issues concerned.

And finally there is the individual level of the participants of the focus groups. The responses and behaviours of separate members can be examined at this level. It is possible to link the contribution of each specific group member by marking all statements of a member in the transcript. In this way the use of certain terms and the contribution of particular themes can be linked to certain roles, identities and personal attributes (Sim, 1998). If people were invited for a particular reason, such as their vision, opinion, or expertise, it is possible to address whether this is evident in the results. A more thorough investigation of group influence on an individual member taking part can be made. A sequential analysis can be conducted for any changes of proposed views or re-emphasizing certain experiences. For instance, social pressure can bring a group member to soften his view on a particular topic and conform to the group norm.

Not every study demands an analysis on all three levels. When analysing with respect to the influences of the group, group interaction data is needed (see Chapter 4). They will help the researcher to understand how contents are collectively created and how interactions build on one another. Group interaction data can be generated in different ways (see also Chapter 4). In reporting focus group results, excerpts of group interactions can be presented instead of quotes of individuals.

Visuals

Visual data are valuable because they communicate through images instead of language, showing aspects of the lives of the participants involved. The conceptualization of images – putting the meaning of the picture into words – is different from conceptualizing verbal data, which expresses thoughts, emotions, ideas and arguments. In Chapter 4 it became clear that neither verbal nor visual data can be considered exact representations of reality. When using visual images, whether media images

(e.g. magazine covers), photographs or drawings, the researcher should be aware that people viewing the images may interpret them differently from those who created them.

Collier (2001) differentiates between indirect and direct analysis of visuals. Indirect analysis means that the images are used to trigger others to tell and explain what can be noticed in them, for instance in photo elicitation (see Chapter 4). Direct analysis means examination of the content and the character of the images as data by the researcher. Direct analysis can be employed as open viewing or open immersion and is based on initial grouping in a natural way. Different criteria can be used to organize and compare the images in order to detect themes and patterns. Direct analysis can also be employed in a more closed and structured way. Much of what then takes place is based on the assumption that visuals are often part of a collection or series.

Collier (2001) distinguishes two types of collections of images. A collection may consist of many different images of the same subject, for instance panoramic views of a park, details, different angles and aerial photographs. In the analysis they are put together to allow descriptions and interpretations to emerge, for instance how the various corners of a park are used differently by visitors or how the park gradually changed over time. Another type of collection consists of different images that help to identify what is depicted in a certain image. For example, an image of a child engaged in playing with a toy that is compared with photos of the child in class, the child playing in other locations or with other toys and photos of family members.

As the literature does not offer much in the way of guidelines for incorporating visual images into an interpretive analysis, I draw on my own experience with student projects in which we practised with the use of photographs and drawings. It was clear that pictures reduced the amount of description needed to understand what a situation looked like, whether a religious lesson in Koran texts, skiing lessons for small children, or women at home and at work. Photographs produced elements of what these situations were supposed to be; images as writing, in fact (O'Reilly, 2005). In this respect they added to the verbal descriptions and could be coded according to certain categories.

The photos also showed the environment, such as a classroom with rows of chairs and tables, a mountain slope, and artefacts like books, skis or computers. They also showed the number of people present as well as something about the relationship between the different photographed individuals, such as teachers and children. In picturing all this, the photographs give us a glimpse of how these situations were experienced, whether seriously or jokingly, weary, bored or eagerly, or intimately or casually. These elements could all be coded, although the meaning and 'proof' was debatable since the photographs could have been interpreted in many different ways.

It seems that today visuals are often used alongside other kinds of methods, maybe to overcome the difficulty in using and interpreting them. The use of different methods to study a research subject is referred to as 'methods triangulation' (see Chapter 9). Radley and Taylor (2003) are among the few who offer some valuable points for analysing visual material in combination with interviews:

- Set out all photographs to examine the kinds of images selected: what scenes, events, places, spaces, and objects have been visualized?
- Listen to the interviews with the photos at hand to find out more about the reasons for taking these particular photographs.

- Analyse what the pictures denote in terms of times, places and experiences. Sometimes this becomes clear if a participant has given the photograph a caption.
- In what way do the images represent a story that the participant wants to tell?
- What are the limits to which photographs can show what they are intended to exemplify?

In sum, to study human experiences and culture visual materials are of value. Open viewing can help to explore the field of study. Once identified, themes and patterns can be tested via structured analysis and complemented with indirect analysis as in photo elicitation.

Using the computer in analysis

Today many of the above-mentioned activities can be conducted with the use of software packages for qualitative data analysis. Most of these programs have been specifically developed for this purpose. In the literature the term CAQDAS, or computer-assisted qualitative data analysis, is used to indicate analysis supported by software. Some researchers still conduct their analysis manually. They copy documents, cut fragments, archive them in folders, create cards for each concept, stick Post-it notes to the wall to find relationships, draw maps and so on. Computers can make these activities much easier, although they have their limitations as well.

Software developers would like you to believe that the computer will allow you to search the data and present the findings with lightning speed. I call this the 'Furby-syndrome': Furby is a furry toy animal, and children love it when Furby talks back and even changes its facial expressions. They also enjoy having conversations with the toy because it is programmed to give adequate responses. Software for qualitative data analysis also gives you appropriate responses when you search for words, codes and frequencies and willingly shows text portraits, matrices, tables and colourful charts. But software cannot help you to judge the worth of a document, it does not break up the text into meaningful fragments, it does not decide what codes to assign, it cannot interpret any relationships between categories, and it does not come up with great ideas to capture in memos. The artificial intelligence part is still not that well-developed.

I am a big supporter of computer use in analysis, but it is the researcher who has to know what it is the computer needs to do. I expect that with digital technology still developing fast, the debate whether to use a computer or not will lose its currency. Vast amounts of digital documents, digital visual material, voice recognition software and simultaneous analysis of the transcript, as well as the audio-recording, are all within easy reach via the computer.

If the majority of the effort remains the work of the researcher, what then are the functions of computers in qualitative analysis? In this section the most used functions of qualitative analysis software packages will be described. In the following section the description will continue with a reflection on computer-assisted data analysis.

Tools

In writing about the tools of software programs I refer to the code-based theory-building software (Lewins & Silver, 2007). These packages assist the researcher in coding texts and retrieval along thematic lines as well as in memoing. They have incorporated search tools allowing the researcher to test relationships between categories and concepts. Within the search tools they have sometimes taken on language-based searching tools, like word frequencies, word indexing, the creation of keyword co-occurrence matrices, and proximity plots. Some programs enable the graphic visualization of connections and processes using mapping tools (Lewins & Silver, 2007).

In the description of program tools I draw heavily on my experiences with MAXQDA2007. I will focus on the support of software with the analytical activities that were addressed in this and the previous chapter.

Filing

All materials that you use, whether field notes, interview transcripts or existing documents, can be imported into the database of your project. In MAXQDA this is referred to as a 'document system'. Just as in a real filing cabinet, you can systematically file your documents in 'drawers' of your choice (text groups), whether ifi type of data (for instance, interviews and observations) or chronologically (first-wave interviews and second-wave interviews) or participant groups (teachers and pupils, organization X and organization Y). The document system is very flexible and makes it easy to combine documents from different 'drawers'.

Editing

Most programs allow you to change your documents once imported. Not only can you remove, for instance, spelling errors, but you can also change the fonts, use underlining, italics or bold and colour words or text fragments.

Coding

One of the basic functions of software for qualitative data analysis is the coding function. In MAXQDA you work in the text browser to select an excerpt and award it a code. The codes are hierarchically ordered in a coding scheme (code system). New codes are easily generated. Needless to say, it is the researcher who develops the codes. Modifying coding, like merging, splitting, renaming, deleting and moving excerpts from one code to another is all very easy and can be carried out systematically. For example, if you change the code 'pet' into 'dog' in the code tree, the program will automatically change the code in all documents in which this code was used.

The codes are displayed in the margins of the working documents. You can assign more than one code to a fragment, overlap fragments, nest them and so on. If you want to stress the interconnectedness between the sub-codes and the main code, you can use colours. You can choose not to display all codes and make some of them invisible, for instance those of a particular colour or awarded by a particular user. This is a useful function if you work in teams and want to compare codes without being influenced by what the others have done.

Retrieving

In software for qualitative data analysis it is easy to retrieve your coded segments. It is also very flexible. You can retrieve from all your documents or from a selection. You can retrieve one code at a time or several codes at once. More advanced retrieval is possible with the use of 'boolean' searches, for instance fragments that have awarded code A and code B or fragments with code A and code B within only five sentences.

In MAXQDA you can award attributes to your documents; for instance, sex and age of an interviewee, or the number of members of a group when your document is a focus group transcript. You can also award attributes that relate to your analysis and codes. For instance, you assign the code 'happiness' and then decide to turn this code into an attribute. The number of times this code is used in a particular document is considered as the value on this attribute for this document. Later on you can retrieve fragments from documents in which, for example, the code 'happiness' was used more than five times ('happy people'). Then you can make all kinds of combinations, for instance, you can search in the group of female happy elderly and male unhappy youth. The attributes can then be used in software for quantitative data analysis.

It is very easy to get an overview of all retrieved segments in all documents. The program counts the frequency of a certain code and it shows you the line numbers and position in the documents. Per text this can give you a kind of profile in terms of codes that you can compare with other texts.

One of the criticisms about coding and in particular coding with the computer was that retrieved fragments were decontextualized, that is, taken out of the situation or conversation they had taken place in. However, in modern software one click on a retrieved segment opens the original text and shows you where the fragment was originally located.

Searching

Most programs have a function to search for words or strings of words in your texts. You can also search for words that are often found in each others' vicinity. The results of these lexical searches show you the context of sentences in which certain words were used. Once again, frequencies can be obtained.

Memos

Memos tend to get lost, since they are like the yellow Post-it notes we use on our books, desks and walls. Modern software incorporates the Post-its in our projects. In MAXQDA the memo icon is in the form of a yellow Post-it. This function enables you to write project memos, document memos, code memos and text memos. Memos can contain all kinds of remarks, and are easily adjusted and completed. You can give them colours or other symbols to remind you of their content and the type of annotation. Using the software, it is easy to generate an overview of memos.

Visualizations

Visualizations are novel. MAXQDA started with cross-case visualizations. It used a matrix to combine documents and codes and another matrix to combine only codes to search for co-occurrences. Recently per case visualizations were added. They

make use of the colours that have been attached to certain codes and show which colour (=code) often pops up in a certain text (textportrait) or in what sequence the colours (=codes) appear in a text (codeline).

In addition to these features, there is a special graphical tool available within MAXQDA called MAXMaps. This allows you to import elements of your MAXQDA project and draw diagrams and maps showing the relationships between the elements. Of course, you have to link the elements; the computer does not do that for you.

Writing

Memos and retrieved segments serve as the basis for writing the chapters in the research report. Systematically all fragments that you awarded a certain code are retrieved and can either be printed or transported to your text processor. You can use the fragments as a prompt to refresh your memory of what to write about or directly as quotes in your report. Again, the computer does not write for you, but it provides the basic material that you need to write about.

Methodological memos can be easily retrieved for writing the methods section of your report. Memos are also used to have all your ideas and brain waves at hand while writing. Other types of output are tables of, for instance, search commands, or visuals that can be used in reports and presentations.

Of course there is more. As we have seen already, text parts can be linked to other parts in the same text or in a different text or even to other sources (hyperlinking). Sometimes visual materials such as photographs can be imported and coded as for any other document. Team members can work on the project at the same time and merge projects later on. There are options for quick coding, in vivo coding and so on. However, the core function of software for qualitative data analysis is to facilitate the data management and analytic tasks.

TIP

Lewins and Silver (2007) ascertain that most of the students they train in software find innovative ways of using the packages. They seem to work independently, using the tools in a way that benefits their work and can be legitimated by their methodological approach. Consider a cupboard that you can use for shoes, CDs, jewellery, clothes or books. It can be used standing up or lying down, underneath your bed, in the living room or in the hallway. With the use of software it is much the same: tools can be used in any way you consider appropriate for your project. I think this opinion from such experienced teachers in software as Lewins and Silver should boost our confidence!

Reflections on computer-assisted data analysis

There are a lot of programs on the market that assist in qualitative research. Examples are QSR Nvivo, ATLAS.ti, HyperRESEARCH, MAXQDA2007, QDA Miner,

The Ethnograph, Transana, Qualrus and Framework. Information on specific software packages, such as tutorials and manuals, as well as demo versions, are available on the software developer websites. Programs are not static products; they are continually revised and developed. Some authors who have written books about qualitative analysis recommend the software they are familiar with. Examples are Nvivo (Richards, 2005; Bryman, 2008) and HyperRESEARCH (Hesse-Biber & Leavy, 2006).

A few authors have made attempts to compare the programs and help researchers to select the best one for their use (Lewins & Silver, 2007; Fielding & Lee, 1991; Tesch, 1991). There are also useful software sites that provide information on computer applications, education and conferences, and advice about the purchase of software such as the Computer-Aided Qualitative Data Analysis Networking Project (http://caqdas.soc.surrey.ac.uk). In my opinion, except for the software enthusiasts, most researchers do not have the time nor inclination for long deliberations about their choice of software.

You can make a choice by looking at a few criteria. The first one is to clarify your needs and subsequently your expectations of the program. Ask yourself what part of data analysis worries you and what would help you to do this task. St John and Johnson (2000: 397) suggest some helpful questions:

- What are the advantages and disadvantages of this package for my research?
- What purpose will this package serve for this research project?
- Will this package handle the kind and amount of data I intend to collect?
- Will this package enable flexible handling of data for this project?
- Will this package enable me to interact with and reduce data in a way that is consistent with my methods?

Another important criterion to consider is user-friendliness. How easy and comfortable it is to work with the package is one aspect. User interface and interactivity are key words here (Hesse-Biber & Leavy, 2006). However, this is partly a matter of taste. 'Interface' refers to the organization of the screen, whether it is appealing to you and whether it matches your way of thinking and working. For instance, do you like a structured interface that always looks the same, or is a flexible interface more convenient for you? Do you remember things by their colours or do colours confuse you? Interactivity means that there is a good, instant hyperlink (one click or double click) between an object in one pane and an object in another, for instance a code can easily be linked with its context and retrieved segment and there is instantaneous access to source data files once introduced into the project (Lewins & Silver, 2007).

Another aspect of user-friendliness is how fast you can learn to work with the program. Again, except for the enthusiasts, most researchers do not want to spend hours and hours before they can start using the program. So it is worthwhile to know the capabilities of the packages and to estimate the time and effort required to use the package. Programs that offer many capabilities are generally speaking harder to learn. On the other hand, they might have all the tools that you need.

A last criterion to consider when purchasing a program is its price. Websites are a useful guide in this matter. Many programs offer free demos that can be used for a

certain period of time, offer upgrades, additional modules and special deals for students. Search for the 'best purchase' for your purpose, which means that for you the optimal balance is found between capabilities of the program and its price.

There are current debates about the strengths and possible concerns of data analysis software (Lewins & Silver, 2007; St John & Johnson, 2000; Fielding & Lee, 1991). Here are a few remarks from a user's perspective.

A particularly attractive advantage is that software packages free the researcher of many manual and mechanical tasks which in turn saves time. The examination of the data can be more complete and rigorous too. Once you have gone over the documents and coded them and wrote the memos, you can retrieve them at any time, and a complete and systematic overview will be given. According to Lewins and Silver (2007), one of the main aspects of computer use which has most changed the process of interpretive analysis is the ability to gain direct access to all relevant areas in the data in immediate response to the researcher's train of thought.

The ease by which data are handled might seduce researchers into using larger amounts of data (Seidel, 1991). This is not a problem for computers providing the memory capacity is large enough for the job. Data overload is actually a researcher's problem. Qualitative researchers often have a vivid memory of data collection. With too much data they run the risk of not remembering anything of the data themselves and of losing the overview. The consequence might be a more superficial analysis rather than the more complex analysis aimed at originally.

Software opens up opportunities to look at the data from different angles. The tools provided might reveal patterns otherwise not noticed. At the same time, by providing these tools the software in some way predetermines what will be drawn to your attention (St John & Johnson, 2000). This may be an extra time investment that does not need further consideration, or it could take your attention away from other more relevant parts of analysis. I will give one example drawn from my own experience. In the beginning, when MAXQDA was still called Winmax, there was only one colour involved in coding: green. In the new product, several colours can be picked to distinguish different codes. At first it never even crossed my mind to use the colours because I considered thinking about using the colour and clicking on it as extra mental and manual activities. But now I find myself using colours more frequently. First, simply because the option is available and this might even open up new opportunities. Second, other newly added visual tools, like textportrait and codeline, build on the code colours.

Some software developers argue that their packages support researchers to examine relationships between categories and engage in theory building. They point at tools like retrieval with boolean search operators, visual tools that show co-occurrence of different codes in a matrix or show code-profiles of texts, linguistic search operations and the use of attributes. All these functions seem to be 'under construction'. Lewins and Silver (2007) advise stepping away from the computer now and then to use paper and pencil. There is nothing wrong with scribbling diagrams or provisional frameworks on paper, and there is also nothing wrong with boolean searches and graphical representations of texts. In some way it is all a matter of taste, as long as it helps you to notice new things in your data.

Readings I learnt much from

Hesse-Biber, S.N. & Leavy, P. (2006). *The practice of qualitative research.* London: Sage.
Lewins, A. & Silver, C. (2007). *Using software in qualitative research: a step-by-step guide.* London: Sage.
Seidel, J.V. (1991). Method and madness in the application of computer technology to qualitative data analysis. In: N.G. Fielding & R.M. Lee (Eds.), *Using computers in qualitative research* (pp. 107–116). Newbury Park, CA: Sage.
Strauss, A.L. (1987). *Qualitative analysis for social scientists.* Cambridge: Cambridge University Press.

Doing your own qualitative research project

Step 7: Using heuristics and software

1. When coming to a closure of the data transformation that leads to the findings, you may benefit from several thinking aids. Have a thorough look at the heuristics described at the beginning of this chapter and apply the ones you think will be useful in your project.
2. Analysing focus group data and visuals is somewhat more pioneering than the analysis of the commonly used individual interviews or observations. You might consider searching for interesting publications of these sorts of data and examine how they have been analysed and how the findings were reported.
3. It is quite obvious when using qualitative software that you will need to use the code-and-retrieve tools. Take a look at the less evident tools that might be helpful for your project, such as searching, linking, visualizing and creating subsets of texts and codes. Make sure that you use the biggest surplus value of the software, which is that your texts, code system, memos and interim products are all linked to shape new constellations that might offer new insights into your research subject.

8
FINDINGS

We have addressed many topics related to qualitative research and to qualitative analysis in particular. Having worked with the data during analysis, one could assume that you know how the findings of a qualitative research are conceived and what they will look like. In practice this is not so easy, and social scientists are often puzzled as to what can be considered to be the outcomes of a qualitative research project. This chapter will focus on the different types of findings that a qualitative effort can produce. Findings can exist in their own right to add to theoretical knowledge or to professional practice, but they can also be used for re-analysis in a meta-synthesis. Mixed methods research is also addressed, in which qualitative research has a specific role and is chosen for its distinctive findings in comparison with the statistical findings of quantitative research.

LEARNING AIMS

By the end of this chapter, you will be able to:

- Grasp the relationship between data, interpretations and findings
- Identify and prevent common mistakes in reporting qualitative findings
- Adjust the degree of data transformation and the findings to the research purpose
- Define qualitative meta-synthesis and indicate different elements of it
- Elaborate on the advocates' and adversaries' points of view with regard to combining different qualitative traditions or approaches
- Reflect on the debate about the paradigm–method fit
- Address a mixed methods design and explain the role of the qualitative part in it
- Direct instrumental and conceptual use of findings

Results of qualitative research

Findings in qualitative research are conceptualized differently from findings in quantitative research. This relates to the methods and the particular role that theory plays