

Analysing data

There are many ways to do qualitative analysis. Tesch (1990) identified over twenty separate approaches that between them encompass a broad range of methodologies and epistemological perspectives. There are also many texts providing detailed guides to the practicalities of doing qualitative data analysis (see, for example, Glaser & Strauss, 1967; Strauss, 1987; Dey, 1993; Miles & Huberman, 1994; Lofland & Lofland, 1995; Strauss & Corbin, 1998; Taylor & Bogdan, 1998). In this chapter we have opted to avoid 'potted versions' of research methodologies and restating analysis techniques that are readily available elsewhere.

While there are many different approaches to qualitative data analysis, in practice most approaches involve similar stages. Whatever the research purpose and question, certain analytic choices have to be made—what data to collect, from whom, how to focus the analysis and how to structure the research report.

Data for analysis may come from many sources and be in many forms. They may be obtained through interviews, observation, or content analysis of existing materials, and may include interview tapes or transcripts, observation field notes, notes on interview context and process, analytic notes and memos, or journal entries. The level and extent of analysis of data from the various sources will depend on the purpose for which the data was collected and involves choices that need to be made for each project.

While we do not hold firmly to any one way of doing data analysis, we are absolutely clear of the need to be as explicit as

possible about what is being done and why. Taylor and Bogdan (1998) suggest that all qualitative research reports should 'provide enough information about how your research was conducted to enable readers to discount your account or to understand it in the context of how it was produced' (1988, p. 168).

The following section focuses on some key stages of qualitative data analysis. Our comments are particularly relevant to approaches that seek to identify themes or patterns in the data. These include studies that are conducted inductively, with a view to generating new theory, as well as those to test an existing theory which use a more deductive approach. This is followed by the presentation of some perspectives on data analysis from four of our researchers: Catherine McDonald uses a deductive approach to analysis, testing the applicability of neo-institutional theory to the non-profit organisations she studied, while the approaches used by Cheryl Tilse, Anne Coleman and Caroline Thomas are essentially inductive.

Transcribing

Most researchers would agree that any audio-recorded data that are to be systematically analysed will need to be transcribed. Transcribing is a time-consuming process, especially for researchers who lack the resources to employ a transcriber. Estimates of the time it takes to transcribe an hour of tape vary according to the level of detail required for the transcripts, the quality of the tape, and the number of voices on the tape. A reasonably good quality recording, with an interviewer and one respondent, transcribed at a general level of detail by a proficient transcriber with a transcribing machine, will take around three hours per hour of tape.

For most qualitative research, transcription at a general level of detail would include, at a minimum, identification of long pauses (with the number of seconds or minutes typed in) and bracketed indications of obvious emotional content such as laughing, crying or sighing. There will inevitably be a great deal of contextual material that does not get onto the tape, including non-verbals and aspects of the interview setting that may impact on what is said and how. The researcher's field notes can provide at least one perspective on these issues and thus should be sufficiently comprehensive to include any contextual factors that may have a bearing on the research.

Some specialised forms of analysis, such as conversation analysis (Psathas, 1995), narrative analysis (Riessman, 1993), or discourse

analysis (Gee, 1999), may require more detailed transcription. Riessman recommends a more general transcription of the entire interview and then retranscription for detailed analysis of sections of particular interest. Gee on the other hand suggests 'transcribing for more detail than may in the end be relevant' (1999, p. 88).

Wherever possible, be prepared to pay for an experienced transcriber, who should be well briefed about the level of detail required. This will vary from study to study and from researcher to researcher. It is also important to decide a process for dealing with words or phrases the transcriber cannot hear or is unfamiliar with. We suggest identifying such places in the transcript with an agreed symbol such as a series of crosses or question marks, and having an understanding that the transcriber will not try to guess. Experienced transcribers will understand the importance of confidentiality, but it is essential to address this issue specifically with each job.

Transcribing is not only a specialised technical skill, it can also be hard work emotionally. Where sensitive topics are raised in interviews, be prepared to spend time to allow the transcriber to talk about their emotional response to the transcripts (see Matocha, 1992). For the transcriber, the powerful emotions experienced can be particularly intense, undistilled by other aspects of the interview and the research context.

Becoming familiar with the data

Once the transcript has been checked for accuracy, listening to the tape again, with the transcript in hand, can be an invaluable way of getting a fuller sense of what the text is about. It is difficult to obtain a good sense of familiarity with the data during in-depth interviews or while observing, as there are so many other things to attend to during data collection. While the researcher undoubtedly forms overall impressions, there is no substitute for this next stage of immersing oneself in the data. For Riessman (1993), 'A focus for analysis often emerges, or becomes clearer, as I see what respondents say' at this stage (1993, p. 57).

Coding

Coding is the process of creating categories and assigning them to selected data (Dey, 1993). In qualitative research this process is

sometimes referred to as indexing (Mason, 1996). While coding is a term used in both quantitative and qualitative research, very different processes are involved. In quantitative research, coding is part of data management and involves numerically transforming the data in preparation for analysis. In qualitative research, coding is an integral part of the analysis, involving sifting through the data, making sense of it and categorising it in various ways. The analytic choices made here about what to code and how will influence every stage of the research from here on.

Qualitative analysis is generally concerned with identifying patterns in the data—different ways in which the data relate to each other. The kinds of patterns identified depend very much on the focus of the study. Lofland and Lofland (1995), for example, have identified a number of levels at which analysis can be focused. These may range from the microscopic to the macroscopic and may be social practices, episodes, encounters, roles, relationships, groups, settlements, social worlds, lifestyles or subcultures. Within each, the specific aspect of focus may be cognitive meanings, feelings or inequalities, and any given study may be focused at one or more levels and one or more aspects. They stress that this approach is meant to provide a mindset for coding that 'should provide you a general orientation to the *kinds of things* for which to look in coding data, *not* a preformed schemata of things for which to code' (1995, p. 122). Bogdan and Biklen (1992) also present a useful list of possible kinds of codes to develop and likewise stress that they are meant to assist in categorising data and not to be rigidly adhered to. Their list includes setting/context codes, definition of the situation codes, perspectives held by participants, participants' ways of thinking about people and objects, process codes, activity codes, event codes, strategy codes, relationship and social structure codes, and methods codes. Again, a study may focus on just one or a number of these types of codes.

Lists such as these are helpful in so far as they draw attention to the vast array of possible ways to focus analysis. They open up the possibilities but also make the need to focus abundantly clear—no one study can hope to look at everything. The danger is that they may provide a false sense of security, a belief that a focus for analysis will emerge at some stage. In practice, decisions about the levels of analysis and types of codes should flow from the research purpose and question, and so be made early on—data collection, too, ought to have been focused in such a way as to obtain the kind

of data that will enable the research question to be answered. This approach does not predetermine what themes will emerge but certainly shapes the kinds of themes. For example, Yvonne Darlington's (1996) study of sexually abused women was primarily concerned with women's subjective experience of the impact of childhood sexual abuse in their lives. In keeping with this phenomenological focus, the interview transcripts were coded with a particular view to relationships and experienced emotions.

Even so, too tight a focus on particular types of data at an early stage of analysis carries the risk that unexpected and unanticipated relationships between the data will be missed. We suggest trying a number of different ways of looking at the data, including looking for differences as well as similarities.

Coding in a team

While coding in a team is essentially no different to coding alone, the processes have to be very explicit and consistently applied. While team coding can be a powerful motivator for rigour, if done poorly the potential for comprising both reliability and validity is multiplied.

Team coding worked well on a study on understanding hope in mental illness (Darlington & Bland, 1999), where both authors coded. We both coded the first few interviews in each of three data sets (consumers, family members and workers), and through this developed our code book. Once we were confident of our own and each other's coding, we divided the remaining interviews between us but still cross-coded a selection of each other's work. Discussions about difficult or unclear coding decisions were invaluable, stopping to ask questions like: Does this bit fit in a category we already have or is it really a new category, something we haven't come across before? Does it require a new code? Having to argue for any new code to a fellow researcher proved to be a good test of whether it was really needed.

Fielding and Lee found numerous instances of differences among team members getting in the way of completing qualitative analysis, but also note the potential of teamwork to enhance qualitative research:

[Team research] obliges researchers to be more explicit about their assumptions and particular understandings of qualitative research.

Team research makes the research process more transparent. Working in teams might help to counteract other pathologies, such as the unchecked proliferation of codes and the contrasting problem of coding which is too 'thin' or superficial (Fielding & Lee, 1998, p. 119).

Do I need a computer program?

The answer to the question of whether a particular piece of research requires a computer program will always be in the first instance, 'maybe'. It will depend on the amount of data and what is to be done with it. There is no doubt that computer programs to assist qualitative data analysis can be of enormous benefit, particularly for studies where there are large amounts of data. A computer program can manage amounts of data that it would be impossible for any researcher to keep in their head and, at the same time, retrieve, or bring to the fore, selected parts of the data for detailed analysis. There is a potential downside to this, in that the capability of the program may encourage sloppy data management early on—coding anything and everything 'just in case', or justifying very broad codes on the basis that one can come back later and refine the coding system if need be. While this can be an advantage in that closure is not reached too early on, the risks are that this can be mistaken for the analytical work integral to coding and that, after a lot of work, one is no closer to understanding what the data mean.

There are an ever-increasing number and variety of programs available, thus knowing what you want a program to do is very important. The two types currently most commonly used in qualitative analysis are code and retrieve programs and theory building programs (Fielding & Lee, 1998; Grbich, 1999). It is important to choose a program that will support rather than constrain your analysis. The availability of training and/or ongoing technical assistance is also an important consideration.

Stories from the field

In this chapter, our stories are relatively brief excerpts from the interviews with some of the researchers introduced earlier. These are not intended to provide comprehensive coverage of the possible

approaches to qualitative data analysis or of all the stages in the analysis process—that would require a book in itself. The array of approaches presented here is indicative of the complexity of the field and of the need for researchers to make choices about data analysis (as with every stage of the research process) in the context of the particular project. We have included excerpts from our interviews with Cheryl Tilse, Anne Coleman, Catherine McDonald and Caroline Thomas.

Cheryl Tilse—*The long goodbye*

Here Cheryl talks about transcribing in-depth interviews and analysing the data from the interviews, the observations and the nursing home documentation.

Transcribing

Cheryl: I had one person who is an experienced transcriber transcribe the interviews in *The Ethnograph* [computer program] format. So the prime tool to help the analysis was to have the interviews transcribed in absolute detail. I asked the transcriber to put in brackets, pauses, crying, that sort of thing but it wasn't the very fine detailed transcribing of conversational analysis. It was very much 'type down what they say' . . .

Yvonne: How important is it to have a good transcriber and a good relationship with your transcriber?

Cheryl: I think it's really important. I thought it was also really important that the same person did them all because she knew what I was talking about and she knew the study and she knew what I was trying to do. It was really important to say things like, 'Don't guess. It's not about that. It's about you typing them up for me and just tell me what's not clear and then it's my job to make the thing perfect. I will go back and check it.' It was also good that I had a good relationship with her in that she got very distressed—I mean they're really sad tapes, they're all about grief and loss and people

crying—and I could say to her, 'How are you doing with this stuff?' because I found when I re-listened to the tapes it was actually more upsetting than when I was doing the interview. When you heard just how profoundly unhappy some people were, it was very moving and I thought, 'She sits at home typing this' . . . So it was really important to give her a chance to debrief or say, 'Well, how do you go when it's so sad? I find them really sad.' We did a lot of this by telephone and she'd say, 'I get up and walk around the kitchen, have a bit of a cry and go back to it.' So, I think it was important that somebody was really listening to what they were typing and really trying to get it down. It wasn't a simple typing job.

Analysing the in-depth interviews

Cheryl: I had mounds of transcripts—very long interviews. I suppose I was always very clear that *The Ethnograph* was just a cut-and-paste tool and that I was doing the analysis. I just did the usual process of coding line by line into categories . . . trying to code as I went along so it was a developing process, rather than say, 'Here's all the interviews, now I'll sit down and code them.' I did a couple of interviews and got them transcribed and coded them and kept working on the general themes. I kept a code book and so if in Interview 3 I introduced a new code, I would go back to Interview 1 and say, 'Is it there called something else?' so that I didn't actually introduce a code in Interview 10 that I hadn't checked all the other interviews for. That was really important—because you suddenly got a sense of 'I'm calling it something else in the early stages but the more I understand what it is I can code it slightly differently, so I need to find those codes'. *The Ethnograph* was handy because you could add things together and do things like that . . . I took four interviews and then I wrote them up as if it was going to be a chapter on four interviews—here are the common themes; here are quotes that

support these themes; here are the differences. And then when I'd done eight, I wrote the chapter again on eight interviews. I was trying to see what was staying in as common themes or what was dropping out, so I really kept on top of the analysis . . . I did four big versions . . . It was a really good way for me to stay on top of such a mound of material and feel that I was writing about it all the time. The other thing I did was use the face sheet option [in *The Ethnograph*] for separating out the themes. The things I looked for were gender, dementia and non-dementia, because the literature says that the spouse with dementia is a much more traumatic care-giving situation than a spouse who has a physical disability. I also looked at age groups . . . to say, 'Well, do they differ?' If I re-analyse it and say, 'Well, I'm looking at it this way, how do these themes come out?', *The Ethnograph* was very useful for doing that—for grouping and seeing the differences . . .

When I saw a common theme I'd take that theme and then say, 'If it's a common theme then I have to find it in every interview'. Or I'd identify which interviews did not reflect the theme. So I was constantly trying to keep in this view, that there might be some people who didn't fit and I'm not just selecting the same four or five interviews that are very rich [data]. I wanted to have an example from each interview and then I could write it down . . . but what I did in the final draft was actually select out the examples that encapsulated the theme best. But the process of analysis was making sure that I said it was common because it was common . . . And when I knew who it wasn't common with, I addressed that in the next chapter . . . So I ended up with another chapter on difference and divergence and that's where I look at men and women and [other differences]. But I also then talked about three people who really didn't fit, which was almost like case study reporting and saying 'this one man is hardly quoted . . . I was determined not to leave him out but it was very clear he didn't fit in'.

Analysing the observational data

Cheryl: What I did here was try and sort through the observations. I was trying to link them to themes of inclusion and exclusion and so I really looked at them in those terms rather than just describing this happened to this visitor, this happened to that visitor—I looked at them all and said, 'Well, what practices and provisions included people and what excluded family visitors?' So I tried to structure it that way so it allowed me to contain heaps of information—notebooks of descriptions of what happened to people. I had some basic research questions, for example, 'How were they provided for?' So I did a description of what the provision was—were there family visiting rooms? Were there kettles? Was there equipment for families, notebooks for families to write things in? So I looked at provision and then I looked at the treatment of visitors and when I talked about treatment it was just this theme of inclusion/exclusion.

The document analysis

Cheryl: The content analysis basically just picked up what guides [the nursing homes] had—they'd have some guides to family visiting or notes for families. Then I did a content analysis and identified what roles they constructed for families—are they resources? Like, 'We appreciate your help with feeding and bringing clothes', or are they seen as, 'Come join us. You're part of the family. We'd like to get to know you.' You know, the underlying themes. So it was really interesting, seeing how particular facilities constructed very clear roles [for families].

Anne Coleman—Five star motels

Anne talks here about her thematic analysis of the data from her observations and informal interviews for her research with homeless

people (Coleman, 2001), and about her use of *The Ethnograph* program for managing text-based data (Seidel 1998).

Analysing the observational data

Anne: In Phase One [observation] I colour-coded everything. I was really only interested in three things—people, spaces and what they were doing in the spaces. So I took the computer notes that I'd typed from the original field notes and I went through every page of those notes with three coloured highlighters—green for spaces, orange for people and purple for acts—and every space named, every person, group, member of a group, whatever, I colour-highlighted the whole lot. I'm fairly visual so once I got it to this stage it was easy. It made sense to me because I'd never keep all that together in my head. And the journal—I didn't do much analysis on the journal because in a sense it was an analysis of what I was feeling that day and what I was thinking and it was also a record of why I'd gone one way methodologically and not the other. It was used all the way through in the final write-up as a way of illustrating, or making a point, or putting a date on something, but it wasn't analysed and used in that sense . . .

In Phase Two [informal interviews] I repeated the way I'd recorded the observations in Phase One so that any time I wasn't directly engaging with people or following that up I'd be having a bit of a sit and writing down what I'd seen. So that was an identical process. I didn't want to record those conversations because I knew they were brief and I was going to have most of them on the run, for example, 'I'm from Queensland University, I'm interested in this, do you come to the Valley often? Have you got five minutes to chat?' And that's all it would be. I'd have the conversation and then I'd run away and write some memory-dot points, this is this, this and this. The only exception was if somebody had said something in a way that was just so brilliant or so clever or just 'I wish

I'd said that', then I would write that down. I repeated the process of typing the [handwritten] field notes into the field notes computer format. I'd go home and from those dot point notes write up the conversation as a separate conversation . . . and again the analysis of that was basically thematic. What I was looking for was the breadth of the range of opinions, so I would have been really disappointed if everybody had said exactly the same thing—they didn't of course.

Using *The Ethnograph* to order the data from the in-depth interviews

Anne: I like to hold something in my hand while I'm reading it and I was extremely concerned about whether using *The Ethnograph* would somehow take the life out of the conversation or transform the information in some way, but I used it partly because I wanted to learn it and this was the perfect opportunity, and the other thing that it was really useful for was as a management tool. Although it was really useful, I've got to say I kept going back to the transcripts, to the hand-held ones. And I did a lot of stuff by hand . . . I did the initial couple of rounds of coding by hand. When I thought I had a broad enough idea about the whole length and breadth of the content and I wasn't coding myself in too tightly, I then entered them in *The Ethnograph* and got all the numbered lines and started to pull the themes out there. But I never totally relied on *The Ethnograph*. To me, it was a management tool.

Catherine McDonald—Institutionalised organisations?

Here Catherine talks about how she analysed the in-depth interviews for her study on the application of neo-institutional theory to the non-profit sector in Queensland. Catherine's work is an example of a very structured approach to qualitative data analysis—a deductive approach, essentially testing an existing theory.

Analysing in-depth interviews using a deductive approach

Catherine: I transcribed all my interviews and I had pre-set categories drawn out of the theory. From the theory I drew out what you would call third order concepts, which were theories. First order concepts were just descriptors, which I'd cut up into second order ones—another level of analysis, then into the third order. So I went through all the interviews once, did the first order analysis, then went through them again and looked at the [second order concepts]. I used *The Ethnograph*.

Yvonne: What happened in the first order analysis?

Catherine: I read the interviews through without my schema [of concepts in neo-institutional theory]. Everything that I thought was interesting I coded and gave a code word. It was a purely descriptive process and I ended up with a coding scheme of about 50 or 60 descriptors. I then pulled out each of those descriptors, had a look at them and did searches on them to see what they were saying. They would either split into higher order concepts or sometimes two descriptors would come into one to be a higher order concept. I then went back and re-coded all the interviews and found second order concepts. Sat back, felt sick. Went back to the schema that I'd developed from the theory, extracted all those concepts and printed them out.

Yvonne: These are your second order concepts?

Catherine: The second order ones. And then I looked at them according to the theoretical schema and I re-coded the whole thing again. So after three lots of coding I ended up with conceptual indicators and there were beautiful examples of all of this stuff coming out. It worked, but it was very tiresome . . .

Yvonne: Could you have gone straight to the higher order concepts—used your schema right from the start—or did you need to go through those stages?

Catherine: In hindsight, I reckon I could have. I mean I didn't because I was learning as I went along, but I've noticed since then when I've been doing this sort of thing that I could just pull [the schema] out and go

straight to the interviews. But the interesting thing was that I got a research assistant and tried to get her to do the same thing. She couldn't do it, just could not do it. She could do the descriptive stuff but she couldn't do what I was doing, which was going, 'Oh yes, I know that, that's an indicator', but it's just confidence and I think it is two levels of confidence. One is I know the theory so well, and because I've applied it once I know now what it is when I see it. I can see it and I go, 'Oh yes, that's so and so.' So, I can see how researchers get quicker and quicker . . . I'm now a more experienced researcher, more familiar with the theory, and I trust my own instincts whereas I didn't do that the first time around—I thought I was stupid and couldn't possibly be right. I felt very unsure of myself about the qualitative analysis in that field. I still feel a little bit uncomfortable with it but I'm getting surer now.

Caroline Thomas—Adopted children speaking

Here Caroline talks about data analysis and, in particular her use of the *NUD*IST* software program, for her interviews with children about their experiences of adoption.

Dorothy: Can you say a little about the data collection and analysis process? For instance, you did a content analysis and chose to use *NUD*IST*, is that right?

Caroline: We were discouraged from doing it because of the time needed to transcribe, code and then to use the volumes of material that *NUD*IST* sorts and reproduces for you. But I wanted to be sure that I had actually surveyed the breadth of the material and I wanted to try to guard against being too selective on the basis of what had stayed with me . . . Although this was an exploratory study, I still wanted to be able to convey whether the views expressed [by a child] were unusual or whether they were common. I wanted to try to see the patterns, to try to get some sense of the proportions of children who said similar things. I knew

that the findings were not statistically significant but I thought it was important to know whether what was being quoted was a one-off example of an extreme case or whether there was some commonality.

Dorothy: Did you look at, say, whether you were getting certain types of responses in relation to the gender of the child or the age at which the child was placed for adoption?

Caroline: We did think about doing that but given that it was such a small sample and that not all the children had been asked the same questions . . . we decided not to in the end.

Dorothy: So in this exploratory study you were interested in the overall themes that are coming through about the adoption process—the chronology of those points of the process you've identified. And were you happy with what *NUD*IST* did for you?

Caroline: Yes. I loved it. I know there are people who don't like using it and who find it cumbersome and time consuming but if I did the exercise again I would want to use a similar tool. I had too elaborate a coding system and consequently the coding of the transcripts took longer than it should have. But given it was the first time I'd used that sort of tool for looking at interview material, it wasn't a surprise that I'd over-catered rather than under-catered with the coding. I would use fewer sub-codes and more global codes, I think, in future.

Comments

At the beginning of this chapter we said we did not intend to provide a comprehensive coverage of qualitative data analysis—the diversity of approaches is far too great and there are many excellent sources for this information. Instead, we provided more general comments about some of the processes common to a range of qualitative data analysis approaches.

These four stories continue that theme. They don't provide a definitive recipe for qualitative data analysis, either individually or in composite. They do provide some glimpses of what happens in practice, how some researchers have gone about their analysis.

They also highlight the crucial role of the researcher who, whatever approach is being used, has to make many decisions about what actually to do in practice.

Qualitative data analysis is a dynamic process and no method can stand in isolation from the world of research practice. Any approach is mediated by the researcher and is only as good as its capacity to assist the researcher to make sense of the data collected. This is not meant to imply an 'anything goes' approach to data analysis. It is always incumbent on the researcher to be rigorous and purposeful, to be clear about the steps taken in data analysis and to be able to defend those steps. This is arguably even more so in qualitative research where there are many possible approaches and where the researcher is so integral an instrument at every stage of the process. The stories presented here, part of the accumulated experience of qualitative researchers, have a part in the continuing evolution of the complex world of qualitative data analysis.