Thematic analysis

Thematic Analysis is an approach to dealing with data that involves the creation and application of 'codes' to data. The 'data' being analysed might take any number of forms – an interview transcript, field notes, policy documents, photographs, video footage. As I said before, there is a clear link between this type of analysis and Grounded Theory, as the latter clearly lays out a framework for carrying out this type of code-related analysis. Similarly, many CAQDAS (Computer Assisted Qualitative Data Analysis) packages are designed to facilitate thematic coding.

'Coding' refers to the creation of categories in relation to data; the grouping together of different instances of datum under an umbrella term that can enable them to be regarded as 'of the same type'. To reiterate a point that I made in our first meeting (and in spite of what some radical grounded theorists might say), decisions about what counts as a category come from all kinds of 'places' – theory, literature, research experience, the data itself.

There is a distinction often made between 'open coding' and 'relational coding'. The former refers simply to the creation of categories and the latter to the *relating* of those categories to each other. Some ways in which categories can be related include:

Cause – Code A *causes* Code B

Property – Code A is a *Property* of Code B

Aspect – Code A is an *Aspect* of Code B

Associate – Code A is *Associated* with Code B

Result – Code A *Results* from Code B

Contrast – Code A *contrasts* with Code B

There are of course many more possible relations (just think of a relational word of the above type and it could be a code relationship). The point here is not to provide a complete list, but to suggest some of the ways in which you might think about your codes while you are developing them. Here are some other things that researchers often do with codes:

Combining or 'merging' Codes - everything that had been coded as 'A' or 'B' now just become 'A'
Splitting a Code into different parts - code 'A' becomes 'A1' and 'A2'

Creating a new code (sometimes called a 'Super code' from two existing codes -anything that has been coded as 'A' AND 'B' is coded as 'C' – Codes 'A' and 'B' are still maintained as distinctive codes though.

Creating a Code Family – placing a group of codes into a code grouping (without actually using that grouping as a new code, as with 'super codes').

Understanding the general principles of coding is pretty straightforward; the idea is to develop themes and to work out how they relate to each other within your data. Key texts that may help you work through these issues are listed at the end of these notes. The reading I have provided for this week comes from the Flick et al collection and is just a brief and comparatively recent discussion of some of the issues which includes some useful references.

While the 'concept' of thematic coding may be straightforward it is a lot less easy to do in practice. It is important to write down the definitions of codes when you create them and to be sure that you are applying the codes in the same way every time you use them. Quite often it is extremely hard to say that a particular instance of code 'x' is the same as instance 'y' as subtle differences in

terminology/speech/approach to issue formulation among interviewees (say) may make it quite hard to decide that 'interviewee A is saying *the same thing* as interviewee B'.

As you work through the application of a code it may be necessary to modify the definition of the code you are working with. This can mean that the more times you apply a code, the broader the definition of it becomes. If a code definition does change, it may be necessary to go and check that the new definition still fits the previous uses of the code (depending on how the definition has changed though). I have found that I tend to use a small number of very general codes in the first wave of coding, and then subsequently develop more specialized codes. Many researchers advocate keeping a 'code log' to note when you change the definition of a code (i.e. the date/time), where you have applied it, and how it relates to other codes you are doing. While most CAQDAS programs are designed to help you do this in an efficient way that makes use of the power of computers (and they are very good for these purposes I must add), it is certainly possible to do all of this stuff without a computer. However, CAQDAS is brilliant at pulling out the sections of data that you have coded. With the click of a mouse you can have all the things you have coded as, for example, 'attitudes to health' uploaded into a single document. When you are working with paper on the other hand, you may well need to (as I have done in the past) spread all your interviews out on the floor, and flick through each one to find the relevant bits. This is a much slower way to work.

'Theoretical' Issues in Thematic Analysis

There are some big issues in undertaking thematic analysis. One of the central positions associated with qualitative research pertains to the idea of interpretivism: that we are interpretive in our actions and in our understanding of the actions of others; that we impose meaning on the world; that we inhabit cultural worlds and engage in cultural practices that are defined by shared interpretations (i.e. that we do not operate as 'isolated individuals' in our interpretive actions, but share with groups of people, certain interpretations).

Language is a prominent aspect of the ways in which we make sense of and order our experiences of the world. Wittgenstein's famous phrase 'the limits of my language are the limits of my world' draws attention to the idea that language forms, in some quite profound way, a tangible context for our actions. Exploring these contexts – their specific features and interrelationships – is, in essence, the central project that drives much of anthropology, sociology, social linguistics, and many other 'human sciences'.

We might say that the human sciences (e.g. the ones listed above) are driven by an attempt to deal with a riddle; a riddle that has been stated by anthropologist Clifford Geertz in the following terms:

…the consoling piety that we are all like to one another and to the worrying suspicion that we are not". (Geertz, C. (1983) *Local Knowledge: further essays in interpretive anthropology*. p. 42. New York, Basic Books. *(I include this quote largely because I am a huge fan of Geertz.)*

In other words, there is an attempt to examine the ways in which language is used in our practices (and 'practices' can mean really anything that we do) in an attempt to understand not only the ways that individuals do things, but the ways in which group practices are structured. Many of the key methodological issues dealt with in the human sciences (e.g. the extent to which individuals 'worlds' are accessible by others; Relativism vs Realism; Positivism vs 'Humanism'; Nature vs Nurture) hinge around debate over the relationship between the individual and the social.

Now, the observations about the importance of language as framing our practices and debates about the extent to which these practices may be intelligible to others also applies to social researchers. The issue here is not so much whether our activities as social researchers make sense to other people (although this is of course pretty fundamental), but rather 'is what we do with other people's practices appropriate?' Who are we to say that 'x' is like 'y' when we are outsiders to the communities being studied? Are we really in a position to say that what we are calling 'John's negative attitude to health care' is the same as 'Sam's negative attitude' if the two people live in entirely different contexts, and have had utterly opposing experiences of the healthcare system.

There is, to put it bluntly, an implicit ‘quantification’ involved in some forms of thematic analysis; an interest in creating ‘bundles’ of instances of behaviour that can be described as ‘alike’ in some way or another. But this quantification does not sit easily with the kinds of assumptions that underlie many of the approaches to qualitative research which are often galvanized around the assumption that the differences between people and the contexts which make those differences are profoundly important.

This is not to suggest that thematic analysis is destined to failure, but merely to point to one of the dilemmas that researchers inevitably face while dealing with data using this approach. There can be value in creating links between, say, people’s attitudes or beliefs, and comparing opposing cases can be a very effective way of throwing things into relief. Thematic Analysis is not then the wrong road, but it is a tricky one to follow without getting lost in (or at least distracted by) some profound theoretical issues…

Suggested Readings

Miles, M. Huberman, M. (1994) Qualitative Data Analysis: an expanded sourcebook. London, Beverley Hills.

This is a very thorough text that is particularly good on Thematic Analysis and is, in many ways, a kind of ’industry standard’ text for this type of research.

Flick, U., E. Vvon Kardorff et al. (2004). A Companion to Qualitative Research. London, Sage.

This is a general text that covers all kinds of issues in qualitative research, and has some good and short chapters on thematic analysis.

Dey, L. (1993). Qualitative Data Analysis: a user-friendly guide, Routledge.

Again, this book offers a particularly thorough discussion of qualitative data analysis.