University of Huddersfield Repository

Gibbs, Graham R.

Using software in qualitative analysis

Original Citation

Gibbs, Graham R. (2013) Using software in qualitative analysis. In: SAGE Handbook of Qualitative Data Analysis. Sage, London, UK, pp. 277-295. ISBN 9781446208984

This version is available at http://eprints.hud.ac.uk/id/eprint/14873/

Using software in qualitative analysis[1]

Graham R Gibbs
University of Huddersfield

Computer software to assist with qualitative data analysis (QDA) has become established as an essential tool for many researchers in the last 20 years. One of the most commonly used acronyms for this software, CAQDAS, introduced by Fielding and Lee following a 1989 conference on the programs (Fielding & Lee, 1991), emphasises that the software assists: Computer Assisted Qualitative Data AnalysiS. However, the assistance given can be seen in different ways. For some, the use of software constitutes a separate kind of analysis to be considered alongside conversation analysis or grounded theory. *The SAGE Handbook of Qualitative Research* (Denzin & Lincoln, 2011b) lists "computer assisted analysis" as a "method of analysis" (table 1.1: 12) and comments that "faced with large amounts of qualitative materials, the investigator seeks ways of managing and interpreting these documents, and here ... computer-assisted models of analysis may be of use" (Denzin & Lincoln, 2011a). Fielding (2000) finds this unsatisfactory because it suggests a coherence in the approaches taken by the software that underplays their differences and because it "confuses a technical resource with an analytic approach." (Fielding, 2000: para. 6). One consequence is a kind of side-lining of such software as a special interest which, as Fielding notes, promotes a pattern of adoption where novices such as graduate students are more likely to use the software than established researchers. A similar point is made by Marshall in her literature survey of the impact of CAQDAS on qualitative sociology. She finds that novice users are apparently less concerned with the object (analysis) than with the tools they are using (Marshall, 2002). All too often this results in research proposals (at both student and research council levels) which imply that the analysis will be 'undertaken' using CAQDAS software. This is what MacMillan and Koenig (2004) refer to as the 'wow' factor where the use of CAQDAS is seen as the method rather than treating the software as only an organisational support for a method. This chapter will argue that CAQDAS is not a distinct method or approach to analysis, that the software does not 'do' the analysis. On the contrary, a major function of the software is to help organise the analysis. In particular it is a way of managing the data and the analytic thoughts that are created in the analysis. The software no more 'does' the analysis than the word processor I am using now writes this chapter for me. Nevertheless, the use of technology is not neutral. Just as word processing has changed the way some people write, so CAQDAS has changed the way analysis is done and there is considerable debate about the extent to which the software has affected practice.

Development of CAQDAS

There is no doubt that the idea of code and retrieve as a form of analysis had a central influence on the development of CAQDAS. It was the development of the personal computer in the early 1980s that made clear the possibilities for supporting qualitative analysis. Breakthroughs came

---

with the development of the Macintosh computer and then the publication of Windows for the PC which presented much more accessible interfaces to users. The first programs focussed on code and retrieve functions by implementing simple ways the researcher could attach codes or tags to sections of text and then undertake retrievals of all the text similarly coded, or in more complex cases text that had been coded that matched a, usually, Boolean combination of codes. (Such combinations are discussed later in the section 'Development of the analysis'). Some software addressed additional issues such as the incorporation of memo writing, supporting the analysis of audio and video, the analysis of narratives or the undertaking of case based analysis and hypothesis testing. Even though many programs have 'fallen by the wayside' over the last few decades, many survive: at least 25 are still on sale or freely available and their persistence suggests they are serving a variety of analytic niches.

Many of those producing software were themselves academics undertaking their own data analysis and their close ties with and understanding of the research needs of the social science community from which they came meant that early adopters felt confident that, once bugs were ironed out, the programs would have the capability to assist in the analysis of qualitative data (Mangabeira, Lee, & Fielding, 2004: 176). The first developments were in the USA but were followed very rapidly by programs written in Germany, Holland, Australia and the UK. In the UK at least, this period coincided with a growth in the use of qualitative methods (Fielding & Lee, 1996). But the most significant factor promoting dissemination in the UK must be the activities of the CAQDAS Networking Project at the University of Surrey, funded for over 15 years by the UK Economic and Social Research Council to promote software use and train researchers to use it. Over 6,000 postgraduates and researchers have been trained on courses run at the Project.

Despite the early beginnings, the adoption of CAQDAS in academia has only taken off since 2000. A bibliographic search of all social science journal papers using any of a range of qualitative methods and using CAQDAS (Figure 1) shows the increasing number of papers mentioning CAQDAS from 1983 to 2011.
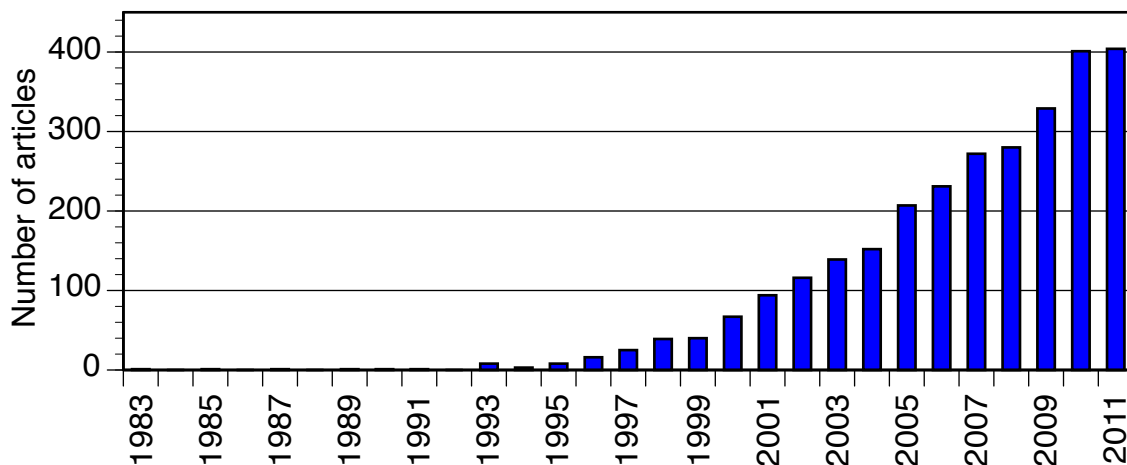

Figure 1. The number of refereed papers published using qualitative methods that used CAQDAS, 1983-2011. (Original to the author.)

2

Set against this upwards trend is the fact that of the recent papers using thematic approaches such as grounded theory, qualitative content analysis and framework analysis only around 2.5%, mentioned using CAQDAS. This low level of use is found outside academia too. In a recent survey of UK companies that undertook market research only 9% said they usually used special computer software for the analysis of qualitative data and only two out of 153 which replied used it regularly (Rettie, Robinson, Radke, & Ye, 2008). Ereaut suggests that one reason for this low take up is that the use of software is seen as undermining the essentially humanistic approach of qualitative research. In market research qualitative research is seen as the opposite of quantitative; an approach that emphasises the insight of human understanding in the analysis and especially the interpretation of highly experienced human experts. In contrast, in academic social research the use of CAQDAS is seen in a positive light, showing one is using the best possible approach to ensure the work is high quality, reliable and exhaustive (Ereaut, 2002).

Of course, the low percentage of academics apparently using CAQDAS may simply reflect a decision not to mention the software they used (just as they didn't mention the word processor they used). If, as I am suggesting, CAQDAS is not a method but rather a way of managing the analytic process then its use may not be of key relevance to the reporting of the data analysis methods. On the other hand, where its use, for example to manage large-scale teamwork, has affected the analysis then readers need to know.

Selecting the program

Most of the software now available provides a common core of functions to support the thematic coding of data and for the comparison of themes across cases. So these programs provide good support for analytic approaches that use these ideas, such as grounded theory (Charmaz, 2006; Glaser & Strauss, 1967), interpretive phenomenological analysis (IPA) (Smith, Flowers, & Larkin, 2009), template analysis (King, 2012), framework analysis (Ritchie & Lewis, 2003) qualitative content analysis (Mayring, 2000) and general thematic approaches. This core of software functions includes:
- The construction, modification and maintenance of code lists,
- The use of these to code documents (usually by selecting data with the mouse and then applying a code to it and usually with some visual way of showing what has been coded),
- Retrievals (at its simplest the retrieval of all materials coded the same way, but also including more complex retrievals of material coded by several codes – in Boolean combinations),
- Ways of dealing with case-based data (often quantitative variables or attributes),
- Writing memos and linking them with other elements in the project,
- Sophisticated text searches (which may also include concordance and word list features),
- A range of diagrams and charts (in many cases with their elements linked back to items in the project) and
- The ability to deal with a range of documents including digitised media documents such as images, audio and video. Several programs now have much improved interfaces for applying coding to these media using rectangular subsets of the image or sections of the timeline (in the case of video and audio).

One of the questions most often asked of the CAQDAS Networking Project at Surrey is 'what software should I use for my project'. In truth, for any project that is undertaking straightforward coding and thematic analysis, just about any of the popular CAQDAS programs will do the job: Atlas.ti. MAXQDA, NVivo, HyperRESEARCH, QDA Miner, Qualrus. Of course, if the research requires additional features then an examination of the software must be made. Fortunately, most software companies have informative websites and they allow the download of a trial version of the software so these features can be checked. Nevertheless, frequently the choice comes down to other factors: what is available, what support and expertise is there, what can be afforded. For many researchers undertaking doctorates or working in collaborative projects, the decision about software will be made for them. It is what the university or project is already using. But this is not necessarily a bad thing, as it may mean that help and expertise are available. One point to bear in mind is that learning how to use the software is not the same as learning about how to undertake analysis using a specific method such as grounded theory (Charmaz, 2006; Glaser & Strauss, 1967), Interpretive Phenomenological Analysis (IPA) (Smith, et al., 2009) or qualitative content analysis (Mayring, 2000). It is best to have a clear idea of how the analytic approach should be undertaken before learning the technical skills needed to use a new piece of software.

A recent attempt to generate more information about the strengths and weaknesses of the software was the Kwalon Experiment, named after The Netherlands Association for Qualitative Research, the body from which the exercise originated, which was run in collaboration with The CAQDAS Networking Project, based in the UK; *FQS*, an international journal, based in Germany, and Kwalitatief Sterk, based in Belgium (Evers, Silver, Mruck, & Peeters, 2010). Researchers from five different software teams (Atlas.ti, Cassandre (Lejeune, 2010), MAXQDA, NVivo, Transana) were given a common set of data to analyse. The data consisted of freely available data, collected from the Internet; of newspaper articles, websites and weblogs, video and audio files. There was a broad variety of geographical origin, actors involved and file formats. The comparison was not in terms of how good an analysis could be done with the software as it was recognised that the quality of analysis still depends largely on the researcher. Each software team reported on their experience, and this highlighted the strengths and occasionally the weaknesses of each program in dealing with particular kinds of data. All dealt well with the core kinds of documents, word-processed transcripts, audio recordings, images. But there were differences in dealing with the range of video formats, with documents in pdf format and material taken from the Internet. In the end, for most users the choice should be based on whether the software has the tools or functions needed, or which is found the most convivial/easy to use.

Some programs offer things that lie outside the mainstream. One issue is price. Most programs are expensive, ten times the price or more than the cost of a mobile app. But Weft, Open Code, AnSWR and TAMS Analyzer (and some others) are free, though this often comes at the price of limited support/updating and programs that may not run in new operating systems. Transana has specialized in dealing with video and has some of the best facilities for this (and is cheap). But other programs (such as Atlas.ti, NVivo and MAXQDA) also now have ways of dealing with video. HyperRESEARCH is the only software that comes in a Macintosh version. All the others run under some version of Windows, though they can be run on Macintosh computers using Boot Camp or on Mac OS X using virtualisation software. Dedoose and Saturate are new approaches to analysis in that they work through the web browser and data are stored 'in the cloud' – that is

on rented Internet servers – rather than on the local PC. This means that the software will run on any operating system, even some tablet computers, but some feel this may compromise the security of the data. Access to data is encrypted to a high degree (in a similar way to secure transactions using credit cards) and the software companies claim this is as secure, if not more so, than data stored on local university PCs or institutional storage areas. It remains to be seen whether research ethic panels will be convinced that data stored this way are properly curated.

The use of CAQDAS makes most sense when most, if not all, of the source files being used are in digital form. Most commonly this means having transcriptions of interviews, focus groups and field notes as word-processed files. Most of the popular programs can also import images in a variety of common formats as well as audio and video files, and they provide ways of viewing and coding such media.

Analysis using CAQDAS

CAQDAS software is essentially a database that holds the source data such as transcripts (including ethnographic notes), video, audio, memos and any other documents that are available in electronic form and then supports the annotation, coding, sorting and other manipulations of them and keeps a record of all this activity. The one key advantage that most researchers using the software claim is that the programs help them keep everything neat and tidy and make it easy to find the material they need later in the analysis. In order to keep a clear mind and not become overwhelmed by the sheer amount of data and analytic writings, the analyst needs to be organised. The bigger the project and the more researchers who are involved, the more sensible it is to use software to support the analysis.

What follows is an attempt to highlight the main issues and functions that need to be thought about as a project is analysed with the help of CAQDAS. Read the instructions and help files and watch the online video lessons that many of the program's publishers have now produced. But at least knowing what the software can do means it is possible to use the help files to see exactly how to do it.

Setting up a new project

The best advice when starting with new software is to spend a bit of time 'playing' with it using some real data; and be prepared to throw it all away and start over again. Prepare all the sources, documents, video etc. so that they can be imported into the program. For information about transcription, see the chapter in this volume. The programs can accept transcripts in a variety of file formats. Atlas.ti, MAXQDA, NVivo and QDA Miner can import pdf too and these programs now do a pretty good job of keeping all the layout information intact so that both text and images can be coded. Word-processed files may use formatting such as colour, fonts, bold, italic, underline, indents etc. and all are preserved on import. (Although, check this, as some of the more basic programs can only accept plain text.) There are, however, some aspects of formatting that are not preserved on import, such as footnotes, table of contents, bibliographies. But usually the textual content is imported in some way even if the formatting is not preserved.

Dealing with video, audio and images can be more complicated. Video camera manufacturers use different codecs in their cameras; that is, different ways of storing and compressing the digital signal. Often the recording will need to be processed through an editing suite to get a file that is usable, but this is pretty normal. Audio is relatively simpler. Most people use digital recorders so simply choose standard formats like WAV for uncompressed, high quality but rather large files, or MP3 for compressed files. Most software can import both these formats. In the case of images there is a range of formats, but most CAQDAS programs can handle the common ones. Just make sure files are saved in the appropriate format from the image editing or scanning software.

At this stage it is important to think about where the project files and all the associated source files will be stored and how to ensure they are safe and secure. Some programs (such as NVivo) will store all the materials in a single file that includes all the source files along with the coding attribute data etc. Others (such as Atlas.ti) store the source material separately from the project file. Because video files can be very large (hundreds of megabytes) even NVivo needs to store them separately from the project file. Once a project is started it is not an easy matter to move externally stored files from one folder to another. So keep all this material together in one folder (which can have subfolders) on the hard disk. Make sure the work is safe. Do not keep just one, single copy on the hard disk of one computer (or even worse on one memory stick). If this breaks down or gets stolen, all the work will be lost. So make back-ups and keep copies on other machines in other places. Most universities now have a form of secure storage on campus that can be accessed from machines with the right password access and which is backed up on a regular basis. So use this if possible as where one of the copies is kept. It may also be necessary to keep the data private, especially if it is of a sensitive or confidential nature. Password protected computers and even password access to the project can help here. Keeping data secure and confidential are key ethical issues. See the chapter in this volume on ethics for the other issues that all qualitative researchers need to consider.

A less technical issue to think about when setting up a project is what the cases are. This is particularly important if some quantitative, perhaps biographical data will be used in the analysis. For example, if the cases are people the data might be their age, gender, education or work experience, or if the cases are places the data might be population, crime rate, state, or if events they might be date, duration, size, type etc. Which will depend on the research question and in complex studies there may be several different case types. Several CAQDAS programs can import quantitative data in spreadsheets (or sometimes in SPSS format). These allow the use of variables or attributes alongside coding in the analysis. Thus it is possible, for example, to narrow down a retrieval to just those cases that match a certain variable value – such as, they are female. Such variable data has to be attached to cases, which in the simple case might just be represented by a document (the interview with that person, for instance). But some software allows a range of sources or even parts of sources to represent a case (all the interviews, sections of group discussions, and video that represent one person, for example) that can be combined into one case. If biographical or other quantitative case data are needed in the project then it is best to decide on the cases when the project is first set up. Things to resolve are the names of the cases (these are used in the variable data files) and whether such variable data will be imported (for example, from a related quantitative survey). A good idea here is to use informative case name abbreviations that include some of the key descriptions relevant to the study such as

F23MC8 to mean family 23, male child aged 8, or Josie_FPhysio to mean Josie (a pseudonym), female physiotherapist. Normally it is best to introduce case data early in the analysis, but attributes or variables can be added after the project has been set up, either from data collected in the field or on the basis of the analysis of the qualitative data as a way of classifying the cases (See, Miles & Huberman, 1994: 102-9).

It makes sense to import into the project a substantial number of the sources before starting analysis. However, depending on the analytical approach, they may not all be needed at the start. For example, when following an IPA or template analysis approach then it is common to start the analysis and the coding on the basis of just a small number of documents or cases. And of course, if the design uses theoretical sampling (see the chapter on grounded theory) the analysis must be started on the early cases in order to identify which further ones are needed for the sample.

Coding

Once some sources (documents, video etc.) have been entered into the project most researchers want to start analysis and the core analytic procedure that the software supports is coding. Usually this is done as a way of indicating all the content (text, images, video etc.) that is relevant to some identified theme. At its simplest, coding enables researchers quickly to retrieve and collect together all the text and other data that they have associated with some thematic idea so that they can be examined together and different cases can be compared.

There has been a long lasting debate about the role of coding in QDA and *a fortiori* in CAQDAS. Some writers, such as Coffey and Atkinson (1996) suggest that coding is just a matter of data management. It may be part of the analysis process but should not be thought of as a substitute for analysis, which they see as an interpretative process. Others say it is more than that. Tesch (1990) argues that coding is not merely the random division of text into smaller units, but requires skilled perception and artful transformation, and for Richards and Richards (1994: 148) coding is "a theorising process" It involves the expedient retrieval of categories, theory building and the pragmatics of breaking down or dissecting one's data into manageable and meaningful analytical units (Bong, 2002: 31). King, in describing his template analysis goes even further and considers the development of the coding template as a key part of the analysis (King, 2012).

However, whichever view the analyst may take on this issue, the software remains agnostic. For the programs, coding is simply a process of attaching a name or tag to a passage of text or an area of an image or a section of a video or audio recording. The software does not care about the analyst's motivation for this act of tagging and it certainly does not understand any interpretation given to it. The intention may be to use the code to represent a deeply thought through, cross-case, thematic idea, or may just be used to highlight some text as a reminder to come back to it later. A similar distinction is made by Seidel and Kelle (1995) for whom codes are differentiated in two basic ways; they can act as "objective, transparent representations of facts" or they are heuristic tools to enable further investigation and discovery. In the latter case, codes are used simply as a way of marking the source without the implications that thematic coding brings with it. Thus the coding of some passage or timeslice can be used simply as a way of highlighting the

text or video as an *aide mémoire* (Weaver & Atkinson, 1994). Of course the text can be highlighted in other ways using italics, bold, colour etc. but using coding to do this is advantageous because it is easy to find all the text highlighted that way by a simple retrieval. Moreover, because it is done electronically it is very easy to delete such coding when there is no further need for it. The point to recognise is that coding in a CAQDAS program can be used in a variety of ways and for a variety of purposes and researchers are free to use it any way they want, to use several different ways in the same project and to change and delete coding as they wish.

All the CAQDAS programs make a fundamental distinction between the data sources (text, audio, video, image) that can be coded, and the list of codes, with which they are coded. Usually the data will be shown in a different window or pane on the screen from the pane or window showing the list of codes. This division goes deeper in the software design though; as there is a range of different things the programs can do with data sources, especially textual data. Aside from coding, one of the most important of these is searching for words or phrases (similar to that found in word processors). Alongside this there is a range of things the programs support that are to do with codes. This includes at the most basic, retrieval (showing all the text that has been coded at that code) but also includes more complex retrievals or queries that allows the user to retrieve source material that meets a combination of criteria that includes being coded by a combination of codes and being in documents or cases that have certain attributes.

The discussion of coding in qualitative analysis has been dominated by the grounded theory approach. This is an essentially inductive approach that starts from a *tabula rasa*. The idea is to put aside any pre-existing theory or understanding of the data and use coding to allow new ideas, themes and theories to emerge. Following this advice, the researcher using CAQDAS would start reading the text (or listening to the video or audio) and code the data as they come to it, interpreting what they see or hear and developing new codes with which to code it. In this process, which is often referred to as open coding, the coding frame, or list of codes, is built up, inductively, by interpreting the data. The software makes it easy to do this both by having simple ways of creating new codes once the text to be coded has been selected and by having clear ways of re-using these codes when further instances are found later in the data. In the software the actual process of coding usually consists of selecting the text to be coded with the mouse cursor and then assigning it to a code. This latter step can be achieved in a number of different ways. Often it is a matter of dragging the selected data (text, image etc.) to the code name (or *vice versa* in some software), or choosing a code name and clicking on a code button, or using the right mouse button, pop-up menu to select a dialog box that enables the user to choose the code(s) to be used. In the past, when using pens and paper transcripts, analysts would indicate the passage they were coding by drawing a line or a bracket in the left or right margin of the page. Several of the programs mimic this by showing a coding bracket or coding stripe with a code name to the left or right of the text it is coding. For most users, this gives a very strong visual cue as to what they have coded and how they have coded it, which is both re-assuring (see how much I have done!) and analytically useful (e.g. when looking for co-occurring codes or consecutive coding). See Figure 2.
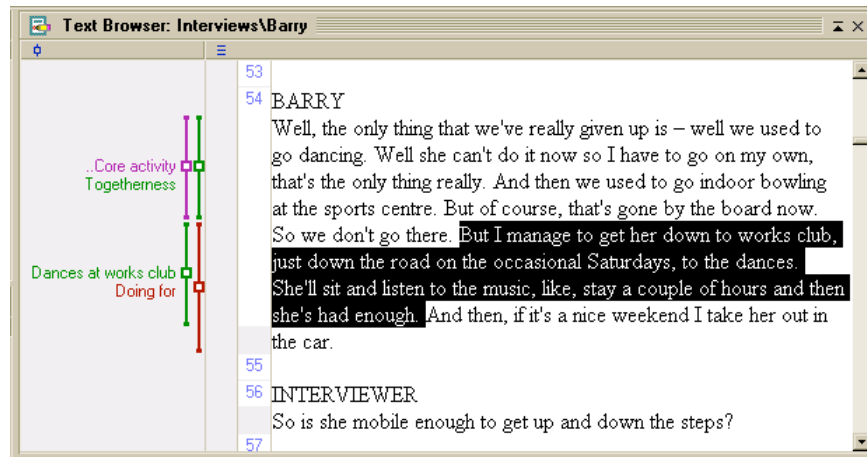
Figure 2. Document showing coding brackets in MAXQDA

But such inductive coding is not the only way to proceed. For a variety of good reasons many research projects already have a set of codes before any analysis is started. Sometimes the analytic approach actually recommends this, as is the case with template analysis. But it is also the case that many research projects are addressing a set of predetermined issues that can constitute an initial, *a priori* coding scheme. Typically this is because the bidding process has obliged them to do so, and perhaps even specify this in the grant application documents, or because the project is of an applied or policy based nature where the issues and themes are well established. Another source of *a priori* coding schemes is what a review of the literature in the field indicates are key issues, themes and interpretations. Such ideas are often represented in the interview schedule used in interview-based research, so these schedules are another source of initial coding ideas. More grounded sources of codes are both the hunches that researchers have because of their familiarity with the field they are researching, or ideas that the researchers have generated or thought of during the field research or the interviews.

Working with the codes

Although for some writers coding is simply a form of data management and has no analytic implications on its own, many researchers have found that thinking about the codes, writing about them (for example, writing memos about them) and manipulating them is a central part of the analytic process they go through in order to extract a coherent and novel understanding from their data. The software includes a variety of tools for the manipulation of codes that supports this kind of thinking.

Most programs have ways of attaching definitions, and in many cases, memos to codes. The latter thus supports the grounded theory practice of writing 'analytic memos'. At its simplest this keeps a record of the definition of the codes and any deliberations about their implications and importance. Such analytic thinking alongside inspection of the coded data following a retrieval may indicate that there is a need to make some changes to the codes themselves or to the data they code. So the programs have tools to support such changes. Retrieval, in order to inspect all the data coded the same way, is usually just a matter of selecting a code, double clicking it or activating it using the pop-up menu. This supports two analytic activities. First, it enables a

9

check that the theme identified makes sense, is well evidenced by the sources that have been coded and that the coding has been done consistently. It is thus a way of checking the quality of the analysis. Second, the researcher can begin to look for patterns within the data sources coded to the same theme. For example, the results across different cases can be compared to see whether there are differences in what has been coded between groups of cases – perhaps all the older cases talk about this thematic issue in a different way from the younger ones. In this way it is possible to build up a more sophisticated account of what is happening amongst the cases in the study.

In addition, the programs have ways to rename codes and to combine them or even divide them into two or more new codes (and to separate the data coded at each new code). In the software this is achieved by cutting and pasting the codes or by dragging and dropping them. The data they code moves with them so that if one code is merged with another, the data each codes is merged together so that the merged code codes all the data that both original codes coded.

Coding crisis

Researchers using an inductive, grounded approach often find they have created a very large number of rather descriptive codes. Gilbert refers to this as "a coding trap" where endless and often needless codes are assigned to the data (Gilbert, 2002: 220). Others refer to the dangers of "data fragmentation" and "decontextualisation" that may follow such over-coding (Ereaut, 2002; Welsh, 2002). Having so many codes is not necessarily a bad thing. It may simply reflect the heterogeneity of the data and the complexity of the analysis. However, it can be a barrier to further analytic work and especially to developing a clear understanding and explanation of the data.

To deal with this researchers often find that they move away from the CAQDAS program for a while. One reason some researchers like to do this is that they are attempting to rethink the analysis. Moving away from the existing project stops it interfering with the rethink. One way to do this is to print out all the codes (possibly with definitions and even short samples of the coded text), cut them up, one code to a piece of paper, and then try rearranging them. This can help to see connections and arrangements that are currently hidden by the arrangement of the codes. Alternatively such details might be cut and pasted into a spreadsheet with rows for each code and columns for things like definitions and examples and for other thoughts about the code. In both cases, look to re-arrange and sort the codes or possibly re-express them; look for patterns and categories amongst the codes.

However, CAQDAS programs do also offer a range of functions to help with the problem. One common approach is to create a hierarchy or tree of codes. The logic of this is that lower level codes are in some way types of, cases of, example of etc. the codes under which they are arranged. Lewins and Silver suggest there are three ways this is done in the software: functional, organisational and cosmetic (Lewins & Silver, 2007: 93). MAXQDA, for instance, has a functional hierarchy so that codes can be arranged into top-level codes and subcodes and sub-subcodes. Activating a top-level code (e.g. for a retrieval) automatically activates all its subcodes. NVivo has organisational hierarchies of codes (which the program calls 'nodes'), they can be arranged into a tree but there is no automatic inheritance of the data coded by its subcodes

when doing a retrieval. Atlas.ti has no hierarchy arrangement of codes but they can be given names that cosmetically mimic a hierarchical arrangement or they can be arranged into families, that is, collections of codes. Many programs also have a way of representing the coding in diagrammatic or network forms. This allows not only hierarchical arrangements to be portrayed but also other kinds of relationships to be represented. Of course this can be done in other software or even with pen and paper, but the great advantage of doing it in the CAQDAS program is that the connection between each code and the data it codes is preserved and usually the researcher can undertake quick retrievals from items in the charts and diagrams. Rearranging the coding scheme and combining codes can be assisted and directed by one very simple activity which is to use the information about how many times and in how many documents a code has been used. Codes used only once or in one case are candidates for combination with other codes or for rearranging into subcodes in a hierarchy.

Such rearrangements allow a variety of analytic ideas to be captured. For example, Corbin and Strauss in discussing grounded theory suggest that codes can have dimensions (Corbin & Strauss, 2008). That is, when participants do something, they may do this in a variety of ways, or when they adopt a particular strategy this may be one of a range of strategies for achieving the same outcomes. Reorganising codes this way (along with renaming, combining and dividing codes) is also a way in which the analyst may begin to identify key or core themes or categories in the data. These may be higher-level codes in the hierarchy or even entirely new codes that combine coding in new ways.

Development of the analysis: retrieval and searching

CAQDAS programs support two kinds of searching, searching for text, also called lexical searching, and searching for codes. Both can be used as ways of advancing the analytic process. In some programs a query tool is used to do one or both of these searches.

Lexical searching is rather like the word searching facility in a word processor, only more powerful. The CAQDAS program can search for a number of different terms (perhaps synonyms) at the same time, search for word roots (and find all the words with different endings) and some programs can even do 'fuzzy searches', that is, search for words spelt like the terms being used or with similar meanings. Lexical searching can help the analysis in a number of ways. First, it can be used as a way of getting familiar with the text, for example, search for terms that are connected with the theoretical hunches in the research and then inspect the passages where the terms are found in the original documents. This might produce new ideas or candidates for new codes. Second, such searching can be used as a way of looking for passages similar to those already coded. These will contain terms, words or phrases that might occur elsewhere and indicate similar topic matter. Put these terms and others related terms into the text search tool to find all the further occurrences. Of course, this will not necessarily find all relevant passages but it can complement a reading of the documents. Third, the approach can be used as a way of checking the validity of the analysis and in particular to check for the occurrence of negative cases – that is, instances that are inconsistent with current explanations in the analysis. These cases may have been missed because they weren't expected in the context they appear. But if they use the same terms or words as other instances then lexical searching will find them. Of course, the approach is not infallible. Relevant passages of text might just not use the terms

being searched for and so will not be found. In the end it is still necessary to read the text and inspect the other sources in a comprehensive way.

Searching for codes, or rather the data they code, is another key way of extending the analysis. At its simplest this amounts to a retrieval on one code. It is quite common, even in published work, for researchers just to summarise the major thematic codes in the report on their study having done such retrievals. This expresses what they have found and, naturally tends to be quite descriptive. Sometimes that is interesting, but as Seale notes, all too often what is produced is "an impressionistic and anecdotal reporting of data" that, at its worst, highlights data that is dramatic but unrepresentative (Seale, 2001: 657). Qualitative studies can go a lot further and offer accounts of the patterns of the occurrence of such themes and, perhaps, suggest causes or explanations for those patterns. In CAQDAS programs it is searching for coded data and searching in cases having certain attributes (often combined with a search for codes) that supports this type of investigation. In this case what is compared in the search is the actual data coded at or linked to the code or attribute. Thus in the simplest case, in a search for one code *or* another the search will find all the text coded at either code, if any (including that coded at both codes, if any).

Searching (or running queries as many programs refer to the activity) with such combinations of codes is divided into two kinds, Boolean and proximity. Boolean searches combine codes using the logical terms like 'and', 'or' and 'not'. Proximity searches rely on the coded data being near, after or perhaps overlapping some other coded data. Commonly used proximity searches are 'followed by' (also referred to as 'sequence' or 'preceding') and 'near' (also referred to as 'co-occurrence'). Boolean searches are most useful in examining hypotheses or ideas about the data and rely on consistent and accurate coding, whereas proximity searches can be used more speculatively and to explore the data, often at an early stage of coding.

For example, in a study of the rehabilitation services available to men and women in prison to help them after their release it would be interesting to compare what respondents said about different kinds of courses; courses focused on family life such as home finances and parenting, courses focussed on work skills such as job application or basic literacy and courses focused on academic areas such as distance learning qualifications. Assuming there were codes in the project for each course type and the data had been comprehensively coded then retrievals on each code in turn would be a first step to analysing what prisoners thought about them. However, other hypotheses and ideas might be investigated; for example that there were differences in men's and women's attitudes to the courses or that reactions to the courses varied with the educational background of the prisoners. Assuming the project has variable data on the gender and educational level of prisoners (the cases in the study), then a search for, say, the code 'work skills course' along with the attribute 'case is female' would retrieve what the women said about work skills courses and this could then be compared with what the men said. This could be repeated for other course types and for cases with different educational levels.

Such thinking about the data, its constant comparison and the close inspection of passages, language, events and actions is a key part of moving the analysis forward. In an influential paper Agar suggested that qualitative researchers using CAQDAS should use the left brain more, by which he meant being able to synthesise and explore the patterns in the data using diagrams,

charts, colour and visual representations of their data and their analytic thinking (Agar, 1991). These suggestions were taken up first by Atlas.ti which now has a particularly well integrated modelling function where diagrams can be created with elements representing codes, quotations (passages of text or parts of images) and memos in the project linked with lines indicating a range of different relationships. Several other programs such as NVivo and MAXQDA now offer similar charting features and have extended them to include representations such as word clouds, cluster analyses and frequency tables. Many researchers find this ability to think visually a liberating way of developing models about their results.

An example of CAQDAS use

The Coping study is a large scale, cross-country study of the effect of the imprisonment of parents on their children and in particular is concerned with what affects the resilience of the children to deal with the situation (http://coping-project.eu/). The study of children, prisoners and carers in four countries: Germany, Romania, Sweden and the UK, had several stages the first of which was a large quantitative survey of the families involved. The second stage, on which I worked, involved in-depth interviews with a subsample of the families in the survey – a total of 343 interviews. NVivo was used to support the qualitative analysis of this interview data, all of which was transcribed. As the researchers undertaking the coding were not necessarily those who undertook the specific interviews, interviewers also wrote short reports on each interview that were stored in the NVivo project as memos attached to the interview documents. These reports contained summaries of the key points the interviewers thought had come out of the interviews, their impressions of the interview (particularly important in the case of some of the younger children) and suggestions for coding and/or analysis.

Initially, an *a priori* coding framework was used. This was developed in large part from the themes addressed in the interview schedule that had been used and also reflected the researchers' prior familiarity with the literature. Major codes used were: Resilience, Stability, Honesty/disclosure, Family relationships, School, Friendships, Interests/sports/hobbies, Bullying, Achievements, Prisoner-child contact, Significant past events, Services and support, Health and well-being, and Experience of the criminal justice system. Many of these also had sub-codes. The researchers ensured that they coded exhaustively, taking particular care to code passages to as many different codes as applied. So typically passages were coded to two, three or more codes.

Teams in the four countries started coding with this framework but at regular team meetings to discuss the coding, several changes were suggested. Some of these were new inductive coding ideas, often suggested by one researcher. The ideas were discussed and either accepted as new codes to be used by all (which meant that other researchers had to re-read their data to incorporate the new code and its coding) or were incorporated into existing codes (in which case the data already coded was merged with the respective existing code). Revised coding frames in the NVivo project were shared across all four countries. In some cases the new codes were added as top-level codes in the hierarchy. One particular example of this was the introduction of two new codes to indicate positive experiences, outcomes, feelings etc. and negative ones. This meant that the researchers could use a Boolean combination of two codes in the query tool to retrieve, say, all the positive outcomes to do with Honesty/disclosure (and contrast it with the

negative outcomes of Honesty/disclosure). The first stage of the analysis consisted in taking all the text about key major themes, for example, School, and identifying the issues and patterns of experience faced by the children by using a combination of Boolean queries along with simple inspection of the overlapping coding using the coding stripes. It became clear that a positive outcome in school was associated with the school having in place systems for dealing with the bullying and stigma which children of prisoners experienced and that the ability of children, carers and professionals to express their emotional experiences was a key aspect of this.

A subset of data from the quantitative survey (selected variables and selected cases) was imported into the project and used to create attributes (such as age and gender) for all the cases in the study. Amongst the attributes introduced was a scale, the Strengths and Difficulties Questionnaire (SDQ) score, that was a measure of the individual's mental health (Goodman, 1997). Honesty and complete disclosure (or at least disclosure appropriate to the age of the child) was identified as a major issue in promoting the resilience of the children. Using these SDQ attributes in Boolean queries it was possible to determine in more detail the kind of disclosure (and what other intervening experiences) promoted the development of good resilience in the children.

Analytic approaches supported

Thematic analysis
There are significant differences in the degree of use of the software by those adopting different approaches to qualitative analysis. Without question, given the central role of coding in most programs, analytic approaches using this, and that includes grounded theory, framework analysis, thematic analysis, IPA, template analysis and qualitative content analysis, are served best by the software. In general, approaches to analysis that are concerned with the development of themes and with analysing data across cases are well supported by the software's functions. Many programs now support a very wide range of types of data to which this coding can be applied. So projects using video, audio, images and documents in the form of pdfs (e.g. projects concerned with analysing the printed media) should all consider using CAQDAS.

Discourse analysis
Approaches that are more ideographic or are more concerned with how language is used and do not use cross-case, thematic analysis are less well served by the software and its functions. This includes discourse analysis, conversation analysis and narrative analysis. Some have argued that the software is of little use at all in such analysis. For example, after a detailed investigation, MacMillan concludes that as there are no universal procedures in discourse analysis, the software cannot offer functions to support any such procedures and is therefore of very limited use in discourse analysis (MacMillan, 2005). On the other hand Silver and Fielding, whilst admitting the reluctance of discourse analysts to use CAQDAS, suggest that the increased power of search and retrieve tools in the programs means they should look again (Silver & Fielding, 2008). They point to the data management functions of the software, "the storage and accessibility of data and interpretation" (2008: 343) and suggest that the ability of the software to isolate passages of text (for example, the procedure of specifying quotations in Atlas.ti) and annotate them will be useful for discourse analysts. In addition, they propose that the text searching and data mining tools that are now built into many programs will be of use in an exploratory way. For example, the

Keyword in Context (KWIC) search in programs like MAXQDA enables the researcher to search for terms quickly and see their use in context; just the kind of recontextualisation that MacMillan complains CAQDAS programs prohibit.

As I suggested earlier, the software is indifferent as to how it is used. It is the human analyst who has to work out their method of analysis and use the software as a tool for their needs. This is shown convincingly by Ryan in her description of how she used NVivo to assist with her discourse analysis (M. Ryan, 2009). She used coding and linking to identify and retrieve examples of various kinds of discursive activity such as a range of types of positioning. Having identified appropriate passages she then used the linking of documents, discussions and drawings to produce what she calls 'multimodal pastiches' in a critical discourse analysis of the data. As she concludes, "the benefits of quick retrieval, efficient linking of data and creating illustrative models, do not preclude contextualised and rigorous qualitative inquiry" (2009: 158).

Mixed methods
One of the biggest recent developments in CAQDAS programs has been their ability to hold quantitative data about cases and to use it in retrievals and queries alongside the coding. This clearly makes the software of interest to researchers engaged in mixed methods research – a growing approach to research (Tashakkori & Teddlie, 2010). The most common situation here is a project design in which there is a large-scale quantitative survey from which a smaller subsample is selected for further qualitative investigation, for example, by interview. In this case the appropriate subset of the data from the survey (selecting only the relevant cases and only the variables required) can be imported into the CAQDAS project and linked with the appropriate case files. Another mixed methods design is where some of the qualitative data is used to create descriptive variables (Bazeley, 2006; Miles & Huberman, 1994: 102-9). In this case the qualitative data are interpreted and applied to cases to classify or even rank them. Such information can then be incorporated into the project as variable or attribute data for further investigations.

Team working
Another situation where CAQDAS programs are proving particularly useful is research involving teams of researchers. These may involve teams of qualitative researchers, perhaps covering different geographical areas or, as just discussed, mixed methods designs with qualitative and quantitative research being used.

Several programs now support team working by allowing, for example, the merging of projects so that the work done by several different researchers can be combined and/or compared. Some programs have server versions that allow multiple users to access the project data at the same time. When there are several researchers, then it is useful to set up different users' accounts and user details for each so that it is possible to see what work in the project each researcher has done. The work of different researchers can be identified by reference to their registration in the project if the software supports that or by using specifically titled codes, memos, annotations etc. This means it is possible to check on the coding work done by different researchers on the same data as a check on intercoder reliability (Armstrong, Gosling, Weinman, & Martaeu, 1997; Bourdon, 2000; Hesse-Biber & Dupuis, 2000; Morse, 1997; G. W. Ryan & Bernard, 2003; Welsh, 2002). In some research, often that using broadly defined, *a priori* coding schemes that

are not altered during the analysis, this can be done on all the coding to ensure that it is done in a consistent way throughout the project. In other cases, the comparisons of coding and code construction can be used as a discussion point in teams. By discussing differences in coding and code construction between researchers the team will be able to come to a more consistent use of the thematic ideas and produce more reliable research, but also, crucially, they will be able to identify deeper and less obvious difference in their interpretations that may point to richer interpretations of the data.

Debates about CAQDAS

For a long time now there have been researchers who have remained suspicious of CAQDAS. Some feel it alienates them from their data by creating an apparent distance between them and their participants. However, whilst early versions of the software made it relatively difficult to see the wider context from which retrieved data came, current programs either show retrievals directly in context or make it easy and rapid to display. Others have suggested that the software's design owes too much to grounded theory and that its use will thus enforce certain analytic strategies (Coffey, Holbrook, & Atkinson, 1996; Lonkila, 1995). But as Kelle (2004) has noted, the coding, indexing, cross-referencing and comparison techniques of the CAQDAS programs are simply different version of the 'age-old' techniques of data management used implicitly in social science research. The same might be said for the often-expressed criticism that software use is associated with a superficial and descriptive analysis of the data. There is nothing in the software that either demands or encourages this. In fact, if anything, the software makes it transparent or more visible when analysis is so limited (Johnston, 2006).

In fact, transparency is one of the advantages of using CAQDAS. Software assisted analysis means that all the analytic work can be kept organised and together and with some planning it is possible to keep a good record of how the analysis has developed and (if the methods of identifying researchers discussed above are used) who has done it. Analytic thinking is thus rendered transparent, although, it has to be said, very few research reports or journal papers take advantage of this to discuss the development of their theoretical conclusions and outcomes. Moreover, far from imposing an analytic method the use of software actually makes the analysis more flexible. Researchers using just paper and pen are reluctant to modify their coding schemes and find it very time-consuming to undertake anything other than very simple retrievals based on their coding. Software makes the restructuring and reorganisation of the coding scheme possible without starting from scratch (although it may still involve a lot of work) and especially if variable or attribute data is included in the data base, all kinds of complex retrievals are relatively quick to perform. The software does not require certain kinds of analysis or particular analytic activities; it is simply a tool that can be used any way the researcher wants.

But, if that is the case, does that mean that CAQDAS use has had no impact on actual analytic methods? Whilst, as I suggested at the start, CAQDAS is not itself an analytic method, I think its use has, and will continue to have an impact on how analysis is undertaken. First, it has made it easier to deal with large and often mixed (qualitative and quantitative) data sets. For example, market researchers surveyed by Rettie et al. said of CAQDAS "that it helps when handling a large volume of data, that it helps when handling complex data, that it makes the analysis more

scientific and that it makes the analysis more systematic." (Rettie, et al., 2008: 81). Some commentators had worried that in-depth, careful analysis may be lost to a shallower type of exploration of larger and larger datasets, just because it had become more possible. However, as Fielding and Lee (1998) have argued, it is not the availability of CAQDAS packages that has led to large projects but rather researchers who are undertaking large projects seek out the software. The size of the project, they suggested, is more likely to be determined by methodological stance or by the sponsor. In fact, Fielding and Lee provide some evidence that on average, datasets and samples in qualitative projects have not been getting larger.

Second, many programs now include a range of functions that offer various quantitative measures and statistics. These include a range of word search (and phrase search) functions and statistics (such as cluster analysis) and tables of coding frequency and amount. Whilst I suspect these are not much used at the moment, they do offer ways of undertaking mixed methods research that goes beyond the simple combination of qualitative and quantitative analysis. Given the growing popularity of mixed methods, the use of these analytic methods seems set to grow.

Third, and I believe this is a pointer to the future, CAQDAS programs are now offering mechanisms for dealing with a wide range of digital data beyond the images, audio and video already mentioned. This includes, for example, geo-tagging data in the project (as done by Atlas.ti), linking the data with an external geographic information system (Fielding, 2012; Fielding & Cisneros-Puebla, 2009), importation of pdfs (long waited for by media researchers) and making accessible a range of web sources such as blogs, discussion groups, Facebook and Twitter. As more of our social life happens on the Internet or at least is mediated by it, so there will be a growing need to use appropriate software to investigate it.

Resources

The CAQDAS Networking Project provides practical support, training and information in the use of a range of software programs designed to assist qualitative data analysis and has links to all the manufacturers' websites. (caqdas.soc.surrey.ac.uk)
Manufacturers videos are mainly on YouTube. There are links to these on OnlineQDA (onlineqda.hud.ac.uk).

Suggested Further reading

N.B. Some of the books mentioned here cover older software version. However, their advice and much of the detailed instructions will still apply to the most recent versions.

Lewins, Ann and Silver, Christina (2007) *Using Software in Qualitative Research: A Step-by-Step Guide*. London: Sage.

Written by two experts from the CAQDAS Networking Project, this covers in detail the three most popular programs; NVivo, MAXQDA and Atlas.ti, as well as some discussion of other programs. There is good advice on how to choose the software and how to set up and use it in a project. A new edition, covering the latest versions of the software is due out soon.

di Gregorio, Silvana and Davidson, Judith (2008) *Qualitative Research Design for Software Users*. Maidenhead: Open University Press, McGraw-Hill.

Focuses on the issues around the ways that a research design will influence and be influenced by the use of software. In particular it contains sage advice about things to consider when first setting up data in a new computer project.

Bazeley, Pat (2007) *Qualitative Data Analysis with NVivo*. London: Sage.

This works through all the stages of undertaking an analysis using one program, NVivo. There are detailed instructions on how to use the software at each step of the research.

Friese, Susanne (2012) *Qualitative Data Analysis with ATLAS.ti*. London: Sage.
Does the same for Atlas.ti and is particularly strong on the use of the network explorer in the program.

Refs

Agar, Michael. (1991) 'The Right Brain Strikes Back'. In N. Fielding & R. M. Lee (eds.), *Using Computers in Qualitative Research*. London: Sage. pp. 181-194.

Armstrong, David, Gosling, Ann, Weinman, Josh, & Martaeu, Theresa. (1997) 'The place of inter-rater reliability in qualitative research: An empirical study', *Sociology,* 31(3): 597-606.

Bazeley, Pat. (2006) 'The Contribution of Computer Software to Integrating Qualitative and Quantitative Data and Analyses', *Research in the Schools,* 13(1): 64-74.

Bong, Sharon A. (2002) 'Debunking Myths in Qualitative Data Analysis', *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research,* 3(2): [np].

Bourdon, Sylvain. (2000). 'Inter-Coder Reliability Verification Using QSR NUD*IST', paper presented at the Strategies in Qualitative Research Conference: Issues and Results from Analysis Using QSR NVivo and QSR NUD*IST, The Institute of Education, University of London, London, UK.

Charmaz, Kathy. (2006) *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London, Thousand Oaks, New Delhi: Sage.

Coffey, Amanda, & Atkinson, Paul. (1996) *Making Sense of Qualitative Data Analysis: Complementary Research Strategies*. London, Thousand Oaks, CA: Sage.

Coffey, Amanda, Holbrook, Beverley, & Atkinson, Paul. (1996) 'Qualitative Data Analysis: Technologies and Representations', *Sociological Research Online,* 1(1).

Corbin, Juliet M., & Strauss, Anselm L. (2008) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (3rd edn.). Thousand Oaks: Sage.

Denzin, Norman K, & Lincoln, Yvonna S. (2011a) 'Introduction: The discipline and practice of qualitative research.'. In N. K. Denzin & Y. S. Lincoln (eds.), *The SAGE handbook of qualitative research*. 4th edn. Los Angeles, London, New Delhi, Singapore, Washington DC: Sage. pp. 1-19.

Denzin, Norman K, & Lincoln, Yvonna S (eds.). (2011b) *The SAGE handbook of qualitative research*. 4th edn. Los Angeles, London, New Delhi, Singapore, Washington DC: Sage.

Ereaut, Gill. (2002) *Analysis and interpretation in qualitative market research* (Vol. 4). London: Sage.

Evers, Jeanine C., Silver, Christina, Mruck, Katja, & Peeters, Bart. (2010) 'Introduction to the KWALON Experiment: Discussions on Qualitative Data Analysis Software by

Developers and Users', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 12(1).

Fielding, Nigel G. (2000) 'The Shared Fate of Two Innovations in Qualitative Methodology: The Relationship of Qualitative Software and Secondary Analysis of Archived Qualitative Data', *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research,* 1(3).

Fielding, Nigel G. (2012) 'Triangulation and Mixed Methods Designs: Data Integration With New Research Technologies', *Journal of Mixed Methods Research,* 6(2): 124–136.

Fielding, Nigel G, & Cisneros-Puebla, César A. (2009) 'CAQDAS-GIS convergence: Toward a new integrated mixed method research practice?', *Journal of Mixed Methods Research,* 3(4): 349-370.

Fielding, Nigel G, & Lee, Raymond M. (1996) 'Diffusion of a methodological innovation: CAQDAS in the UK', *Current Sociology,* 44: 242-258.

Fielding, Nigel G, & Lee, Raymond M. (eds.). (1991) *Using computers in qualitative research.* 2 edn. Newbury Park, London: Sage.

Gilbert, L S. (2002) 'Going the distance: 'closeness' in qualitative data analysis software', *International Journal of Social Research Methodology,* 3(5): 215-228.

Glaser, Barney G, & Strauss, Anselm L. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Chicago: Aldine.

Goodman, R. (1997) 'The Strengths and Difficulties Questionnaire: A Research Note', *Journal of Child Psychology and Psychiatry,* 38: 581-586.

Hesse-Biber, Sharlene Nagy, & Dupuis, Paul. (2000) 'Testing Hypotheses on Qualitative Data: The Use of HyperRESEARCH Computer-Assisted Software', *Social Science Computer Review,* 18(3): 320-328.

Johnston, Lynne H. (2006) 'Software and Method: Reflections on Teaching and Using QSR NVivo in Doctoral Research', *International Journal of Social Research Methodology,* 9(5): 379-391.

Kelle, Udo. (2004) 'Computer-assisted qualitative data analysis'. In C. F. Seale, G. Gobo, J. F. Gubrium & D. Silverman (eds.), *Qualitative Research Practice.* London: Sage. pp. 473-489.

King, Nigel. (2012) 'Doing Template Analysis'. In G. Symon & C. Cassell (eds.), *Qualitative Organizational Research: Core Methods and Current Challenges.* London: Sage. pp. 426-450.

Lejeune, Christophe. (2010) 'From Normal Business to Financial Crisis ... and Back Again. An Illustration of the Benefits of Cassandre for Qualitative Analysis', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 12(1).

Lewins, Ann, & Silver, Christina. (2007) *Using Software in Qualitative Research: A Step-by-Step Guide.* London: Sage.

Lonkila, Marrku. (1995) 'Grounded theory as an emerging paradigm for computer-assisted qualitative data analysis'. In U. Kelle (ed.), *Computer-aided qualitative data analysis: theory, methods and practice.* London; Thousand Oaks, Calif.: Sage.

MacMillan, Katie. (2005) 'More Than Just Coding: Evaluating CAQDAS in a Discourse Analysis of News Texts', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 6(3 Art. 25).

MacMillan, Katie, & Koenig, Thomas. (2004) 'The Wow Factor: Preconceptions and Expectations for Data Analysis Software in Qualitative Research', *Social Science Computer Review,* 22(2): 179-186.

Mangabeira, Wilma C, Lee, Raymond M, & Fielding, Nigel G. (2004) 'Computers and Qualitative Research: Adoption, Use and Representation', *Social Science Computer Review,* 22(2): 167-178.

Marshall, Helen. (2002). 'Alchemists, Housekeepers or Artisans? Approaches to Computer Assisted Qualitative Data Analysis Systems', paper presented at the International Sociological Association, Brisbane, Australia (ISA)

Mayring, Philipp. (2000) 'Qualitative Content Analysis [28 paragraphs]', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 1(2).

Miles, Matthew B., & Huberman, A. Michael. (1994) *Qualitative data analysis: a sourcebook of new methods* (2 edn.). Beverly Hills, CA: Sage.

Morse, Janice M. (1997) '"Perfectly healthy, but dead"- The myth of inter-rater reliability', *Qualitative Health Research,* 7(4): 445-447.

Rettie, Ruth, Robinson, Helen, Radke, Anja, & Ye, Xiajiao. (2008) 'CAQDAS: A supplementary tool for qualitative market research', *Qualitative Market Research: An International Journal,* 11(1): 76-88.

Richards, Lyn, & Richards, Thomas J. (1994) 'From Filing Cabinet to Computer'. In A. Bryman & R. G. Burgess (eds.), *Analyzing Qualitative Data*. London: Routledge. pp. 146-172.

Ritchie, Jane, & Lewis, Jane (eds.). (2003) *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: Sage.

Ryan, Gery W, & Bernard, H Russell. (2003) 'Techniques to Identify Themes', *Field Methods,* 15(1): 85-109.

Ryan, Mary. (2009) 'Making visible the coding process: Using qualitative data software in a post-structural study', *Issues in Educational Research,* 19(2): 142-161.

Seale, Clive F. (2001) 'Computer-Assisted Analysis of Qualitative Interview Data'. In J. F. Gubrium & J. A. Holstein (eds.), *Handbook of Interview Research: Context and Method*. Thousand Oaks, Ca: Sage. pp. 651-670.

Seidel, John V., & Kelle, Udo. (1995) 'Different Functions of Coding in the Analysis of Textual Data'. In U. Kelle (ed.), *Computer-aided qualitative data analysis: theory, methods and practice*. London; Thousand Oaks, Calif.: Sage. pp. 52-61.

Silver, Christina, & Fielding, Nigel. (2008) 'Using Computer Packages in Qualitative Research'. In C. Willig & W. Stainton-Rogers (eds.), *The SAGE Handbook of Qualitative Research in Psychology*. Los Angeles: Sage. pp. 334-351.

Smith, Jonathan A, Flowers, Paul, & Larkin, Michael. (2009) *Interpretative Phenomenological Analysis: Theory, Method and Research* London: SAGE.

Tashakkori, Abbas, & Teddlie, Charles (eds.). (2010) *SAGE Handbook of Mixed Methods in Social & Behavioral Research*. 2 edn. Los Angeles, London, New Delhi, Singapore, Washington DC: Sage.

Tesch, Renata. (1990) *Qualitative research - analysis types and software tools*. London: Falmer.

Weaver, Anna, & Atkinson, Paul. (1994) *Microcomputing and qualitative data analysis*. Aldershot: Avebury.

Welsh, Elaine. (2002) 'Dealing with Data: Using NVivo in the Qualitative Data Analysis Process', *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 3(2).