Why use computers in qualitative research?
3

Why Use Computers in Qualitative Research?

Justifications for using the computer as an analytic tool in qualitative research are inextricably tied to the character of qualitative data. Qualitative research often produces an 'assemblage' of data (Lee, 1993). Typically, this assemblage is multi-stranded, derives from multiple sources, and frequently has multiple forms: transcripts, fieldnotes, documents and so on. Researchers might often also work with a range of data types: jotted notes of various kinds, fieldnotes organized according to topic, diaries, chronologies, card indexes, questionnaires, maps and diagrams, not to mention photographs, video or audio recordings and even artefacts (Ellen, 1984; Levine, 1985; Fischer, 1994). The emergent character of qualitative research also tends to encourage 'data promiscuity'. Because one cannot specify in advance what might eventually be significant, data of different kinds are collected 'just in case'. As a result of all this, qualitative researchers tend to accumulate large volumes of data. It might be helpful to visualize. Suppose a researcher carried out a fairly small-scale qualitative study consisting of 20 depth interviews each lasting about an hour. Assume each interview produces 25 pages of transcript. Data from the interviews will amount to 500 pages or a stack of paper 10 centimetres (4 inches) thick; approximately two volumes of the current London telephone directory. (Of course the density of data in the transcripts is much less than in the directory. On the other hand, the complexity of the data in the transcripts is much greater.)

One consequence of volume is that there is a good deal of redundancy in the data. Alongside material relevant to the task in hand, there is usually much non-relevant material that has to be sorted through and set to one side. Probably every qualitative researcher has had the experience of knowing that there is a vital piece of information buried somewhere in a great mass of material, without it being immediately apparent where it is to be found. The sheer volume of material produced by qualitative research also makes data vulnerable to disruption. An obvious solution to having lots of material is to spread it out on the floor. Doing so, however, renders it vulnerable to draughts, children, pets and over-zealous cleaners.

As well as being voluminous, qualitative data are also typically
WHY USE COMPUTERS IN QUALITATIVE RESEARCH? 57

unstructured, context-specific and recalcitrant. In survey research, responses to questions are pre-structured by the response categories the researcher builds into the questions. This is not the case with qualitative research. Indeed one of the justifications for using qualitative data is precisely that one is not imposing an a priori analytic structure on the data. This is not to say, of course that qualitative data are devoid of structure. Rather that structure is usually implicit. The researcher attempts to discern or elucidate a structure that is regarded as emerging out of the data as the analytic task proceeds. This 'post-structuration' of the data is much more difficult than the pre-structuring of responses that takes place in a survey. Meaning in qualitative data is contextually dependent. In other words, how one interprets a particular utterance in an interview, say, depends on the context within which that utterance is made. An implication of this is that it is important to be able to retain the original context in the analytic process and to be able to switch back and forth between one's analytic material and the original data. Finally, qualitative data are 'recalcitrant' in the sense that they do not yield up their meaning easily. One constantly needs to refer back to the data. Moreover, the process of returning to the data time and time again in itself transforms the analysis. This is so, because, as one's familiarity with the data increases, one's analytic understanding of it is transformed. It becomes important therefore to be able to impose and discard successive analytic schemata on the same basic data.

Not surprisingly in the light of this, one justification for computer use in qualitative data analysis is that the machine can facilitate the task of data management. Tesch (1990), an early CAQDAS enthusiast, argued that the mechanical difficulties inherent in working with qualitative data inhibited the analytic process. Mechanizing manual procedures in qualitative research, she argued, potentially offered considerable benefits in terms of time, efficiency and more thorough analysis. Computers are fast and flexible. One only has to think, for example, about how easy it is using a word processor to reorganize a piece of written text or to find a word or phrase in a lengthy document. The computer, Tesch argued, could bring some of this capability to the analysis of field materials. Tesch also pointed out that storing data on a computer means having less paper to shuffle around. This in turn makes it easier to keep track of material and to find specific items when they are needed. Decreasing the amount of time devoted to managing data makes the analysis process less tiresome. The analyst becomes less fatigued and in consequence can devote more time and mental energy to the analysis itself. A related and important aspect of this, according to Tesch, is that using a computer for qualitative analysis allows one to 'play' with the data. It becomes possible to look at data in different ways, and to try out new analytic approaches even though there is no guarantee that they will work. This kind of activity is important because it potentially increases creativity in dealing with one's data. It can be difficult to do with manual methods.
A second justification for harnessing the computer to the needs of qualitative researchers is that in so doing one is potentially extending the capabilities of qualitative research. The computer in other words provides analytic possibilities difficult to accomplish by 'traditional' methods. For example, replication becomes a possibility (Conrad and Reinharz, 1984). Qualitative analysis can seem a mysterious and non-reproducible process. There is sometimes a suspicion that the analyst has done little more than string together a series of 'juicy' quotes extracted from transcripts. By contrast, the computer encourages users to be clear about what they are doing. It can also be used to keep a log or trail of analytic procedures. Both these features make it easier for a second person to replicate an existing analysis. Computer assistance might also conceivably make team research more feasible in qualitative research (Conrad and Reinharz, 1984). Lone-hand research is common among qualitative researchers. Indeed, it is the norm for graduate-level research. But restricting research to settings which can be handled by only one researcher produces a bias in the total universe of studies towards small-scale social entities; the factory rather than the business, the school rather than the education system. The bias in survey research is in the opposite direction, towards large-scale social aggregates. It might be – and this is just speculation – that social science leaves out of the picture those intermediate structures of social life which arguably form a crucial link between the micro and macro levels of social organization (Fielding, 1988). If there is such a bias, team-based studies might help to overcome it. In a similar vein, Ragin and Becker (1989) argue that the advent of desktop rather than mainframe computing might help to bridge the traditional gulf between 'variable-oriented' (quantitative) and 'case-oriented' (qualitative) researchers. For both groups the microcomputer encourages closeness to data and an intensive, interactive analytic style. This in turn might encourage a certain degree of methodological convergence as quantitative researchers find the detailed analysis of subpopulations easier and qualitative researchers are able to examine comparative contrasts within their case materials more fully.

A third justification for software use is that it can enhance the acceptability and credibility of qualitative research. Some of the early enthusiasm for CAQDAS in the United States seems to have been driven by a concern to make qualitative research appear more 'scientific', and in this way to destigmatize it within a not entirely sympathetic academic environment. While we would judge this to be less true of the UK, qualitative research is not always well regarded by policy makers. Social researchers of all kinds in the UK have for some years faced sustained pressure to deliver research findings in a comprehensive and timely manner. For example, the principal funding body for graduate study in the United Kingdom, the Economic and Social Research Council, has consistently pressed the academic social science community to secure better completion rates for doctoral studies, where 'better' is construed primarily as adhering to a
WHY USE COMPUTERS IN QUALITATIVE RESEARCH? 59

three- or, at most four-year timetable. One justification for software use is that by offering efficiencies and the possibility of more sophisticated analysis, it gives the qualitative researcher some leverage in dealing with such pressures.

Evaluating the Advantages of CAQDAS

Claims that CAQDAS can bring about an advantageous transformation of analytic method in qualitative research have of course generated counter-claims which stress the possible disadvantages of software use. Some of these counter-claims will be discussed in detail later on. The general point we would make at this juncture is that the claims advanced for and against CAQDAS use need to be evaluated in a processual manner. Different advantages and disadvantages are manifest at different stages of the researcher's exposure to CAQDAS. Nor is it simply a question of there being clear and unambiguous advantages or disadvantages; different benefits and disbenefits might be experienced, and these might register differently with different users. It is also sensible to differentiate what we might call 'anticipated' and 'experientially validated' advantages and disadvantages. As we will see and as we have implied above, it is not simply a question of the perceptions users have of what the software does or does not do for them; the perceptions of others are also important. A processual perspective accommodates the dynamic character of the research environment and of the negotiation of the place of CAQDAS within it. This is particularly important when we examine criticisms arising from early experience. In a number of cases technical solutions to problems experienced by the users who participated in our focus groups have emerged which were not available at the time of our fieldwork. Users themselves also become more sophisticated.

Analytic Uses of CAQDAS

Taking a processual view, we begin by considering what factors persuaded users to take up CAQDAS in the first place. Often it was the prospect of working with a lot of qualitative data: 'I started working with [package] to help sort out that amount of data' (FG2). Variety of data sources is an associated consideration. 'I had transcripts, law reports, courtroom scripts, solicitors’ case files, fieldnotes, and to use such a diverse body of data... I felt putting it on computer would be much easier' (FG2). Experienced researchers drew on their knowledge of the time-consuming aspects of qualitative analysis. 'It was expediency really, I knew that I would have
about 50 interviews, some going on to two hours perhaps, and I knew that I had a very limited amount of time. Having done a previous project very similar [where] I ended up with 100 interviews to sort and sift and it was cut-and-paste and photocopy ... I just knew how much space you need for a start. Piles of files and papers everywhere, and marking different categories in different coloured pens, just seemed to me that this would be a mechanical way of doing all that, to give me more time to look at the actual interview data' (FG2). Another respondent remarked that she had often done cut-and-paste during her long daily commute but had lost data as she sifted and sorted it on the train. Another user felt she could not 'keep it all in my head, so I wanted something that was a filing system. I'd been to a seminar in France suggesting ways of doing content analysis by hand and I wasn't at all convinced, just seemed a complete mess and I just couldn't imagine ever being able to do it myself with that amount of data' (FG5). A pragmatic view that CAQDAS would make cut-and-paste easier was also expressed. 'I saw people doing cut-and-paste, and chucking out bits and piles of comments there, and I thought there must be an easier way of doing this' (FG1).

In such cases, the prospective user approaches CAQDAS less as a means of conducting a formal, systematic analysis than as a means of organizing voluminous data. Initial expectations might presage patterns of use, in particular a failure to fully exploit the conceptualizing features of the package. In other cases, systematic analysis was the goal: 'I thought it would help me to be more organized about how I went about my analysis' (FG2). Also oriented to a particular feature was this account: 'I knew looking at my data and the ideas I had before that I did have hierarchical codes, but that I might not be quite sure what they were going to be at the initial stages and I might want that flexibility of being able to move things around' (FG5). An applied researcher had 'worked on fairly small-scale projects and my data sets were quite small so I had always used a cut-and-paste method', but then 'there was a project to look after a national-scale data set and that leapt from fifteen to hundreds' (FG3). This researcher made heavy use of display matrices (Miles and Huberman, 1994) in her doctoral research, 'and as a result of my PhD I had been thinking of what ways I could use to analyse this data' (FG3), so that CAQDAS offered a new means to make analysis systematic.

While convenient, the distinction between 'clerical' data management and analysis is somewhat artificial. After all, the task of 'organizing' involves looking ahead to the likely contours of the analysis. Also oriented to analytic considerations are users who have from the first been dissatisfied with the established approach to qualitative data analysis. 'I started off when I was an undergraduate being not very satisfied with the teaching I got on how to analyse the qualitative material. We were given lectures on how to do interviews and how to think about what codes might be, but it never got further than that. I was getting very interested
WHY USE COMPUTERS IN QUALITATIVE RESEARCH? 61

at that time in how to do statistics and SPSS, and just playing around on the computer, I thought there must be a better way' (FG6). Another user initially obtained CAQDAS for teaching purposes because I wanted to show students there was a systematic way of analysing qualitative data'. He felt that students did not respect qualitative analysis relative to statistical analysis because it seemed *ad hoc*: 'the qualitative teaching was unconvincing. The analysis side was underdeveloped, you had an interview and trying to transcribe it all, but then what do you do with it?' (FG6). In another case a sophisticated user's enthusiasm rubbed off: 'not many people in the university were using [package] but this person had been working with it for two years and was impressed' (FG2). These routes to CAQDAS involve an unfavourable comparison between the systematic procedures of statistically based analysis and the less clearly articulated, intuitive and less-reviewable character of qualitative methods.

Users typically anticipate both practical and analytic benefits. The idea that CAQDAS might help make analysis more systematic was frequently expressed. None of the users were dissuaded of this likely advantage as a result of experience with CAQDAS, yet the time-saving expectation was one that few were able to confirm. For people with a keen interest in qualitative analysis it can, as users often said, be a boost to creative work to have features available that prompt new approaches to the data. 'One of the things that fascinates me is the different ways in which it is possible to analyse text, and ... text of all sorts, and fine grains of analysis, exciting looking at other ways of analysis' (FG1). To another it was a question of discouraging 'this view that qualitative analysis is something magical... By having to think about the categories, codes ... it forces you to think about the systematic process' (FG3). However, it was a question of CAQDAS being an aid or encouragement rather than the *sine qua non* of good practice. Manual qualitative analysis could be rigorous too, it was simply that mechanical cumbersomeness might interfere. One respondent made the point that 'You have to think very clearly about what you are doing', going on to comment: 'You should have been doing that anyway' (FG3). Another remarked that using a computer was no guarantee of rigour: 'if you are unsystematic without computers I am sure you can be unsystematic with them' (FG3). This was one reason that respondents involved in methods teaching were committed to the idea that qualitative analysis principles should be grasped before introducing the software. It was also the case that, without exception, they now included the software in their qualitative methods teaching. Some felt that it helped to convey the principles of qualitative analysis, dividing the analytic process into a series of stages, each of which was accessible for discussion.1 'The initial getting-hold of [package] was for teaching purposes, because I wanted to show students that there was a systematic way of analysing qualitative data' (FG6). Another element of the encouragement to be systematic was the idea that qualitative analysis was iterative and therefore subject to
much reformulation, requiring reconfiguration of the data, elaboration of category systems and so on. CAQDAS could not only track the development of one's thinking but provide facilities to automate recoding and repeat patterned retrievals as new data became available. 'I knew we would want to go back into it over and over again because there would be different things that we would want to do with it' (FG6).

These advantages were not immediately apparent, however. It was often a case of anticipating help with managing and organizing data and then coming to appreciate the facilities for various retrieval strategies and working with the category system. 'I assumed people must be using computers to take a lot of the physical grind out of the analyses. I was amazed when I found out that some of them not exactly put together theories for you but connected things up in search ways that you could thread' (FG4). One user offered an apt process description of how the speed of the computer aided code development. 'In the process of coding a transcript for an interview and then being able to very quickly locate that code at each point through the transcript, [it] enhances my abilities to really think through the material, because there ... almost isn't a time gap between locating one thing and the next. And therefore you are on that roll mentally and then you see the next chunk of text that's about that area and you are still with it ... It's in those very practical ways that it's an enhancer' (FG4). Similarly, 'it allows you to be flexible and try things out and scrap them or move them around or change them without committing yourself' (FG5). Another boon was that it allowed careful checks of the proportion of a sample expressing a given view, which tightened practice and controlled glib summaries from the data. 'It's made us more thorough. When you are interviewing you say "everybody I've interviewed seems to say this, this is what is happening" and [when] you actually go through much more thoroughly in checking [you find] "actually that was only two very articulate people who said that, all the other people [are] saying X", it's made us do that' (FG4). A doctoral student felt that 'it has made me much better organized than I was when I started. It is very easy to shut the door on pieces of paper and scissors and go away, whereas when you produce tons of files with all the different extension names and millions of printouts, you have to have some sort of organization scheme just to know which the latest one is and how you've revised it' (FG4). While non-users sometimes express the reservation that CAQDAS takes researchers away from the data, users often referred to an enhanced grasp of the data as they repeated searches and so on. It made it harder to 'gloss' the data with abstractions. Looking back on projects whose analysis was done 'manually', a qualitative computing advisor remarked that she was now 'rather unsatisfied both with the depth and quality of handling of the unstructured elements of people's data sets and the tendency to produce abstraction in order to legitimate what you are doing' (FG5). Thus do data management issues shade into new, more demanding criteria for data analysis.
Transparency

An effective data management system potentially encourages researchers to produce analysis which is explicit, systematic and documented. Doing so enhances the transparency of the research process. In other words, the researcher and/or others can trace the paths by which a particular analysis emerged. Transparency affects both the process and the product of the research. Some indication has already been given of how transparency makes the analysis process more fruitful by avoiding underinterpretation. If data are manageable, the potential for confusion and muddle is reduced. So too is the temptation to over-rely on the exotic, the near at hand or the voluble. However, the explicit, systematic and documented management of data also helps to avoid overinterpretation. It becomes easier to detect instances where the data have been forced or subjected to premature analytic closure. Transparency also aids analysis by indicating when and how further iteration through the data might be appropriate, and might promote creativity by making anomalous findings more visible. (Miles and Huberman, 1994, detail a variety of innovative strategies for making analytic procedures in qualitative research explicit, systematic and fully documented.)

A variety of audiences have an interest in the published products of research: other researchers, research participants, funders, stakeholders and the public at large (Miles and Huberman, 1994:280). For each of these audiences, concerns might arise about how conclusions have been reached. Some concerns revolve around the veracity of the research. In other words, the question asked is 'Have research data been misrepresented?' Concerns of this kind are most likely to arise where research participants feel they have been subjected to false, malicious or unflattering portrayals of their lives, or where stakeholders encounter unwelcome findings from evaluation studies. Veracity might also be an issue where scientific fraud is suspected although, as Miles and Huberman (1994:287) comment, 'we rarely consider this eventuality in qualitative studies'.

A second set of concerns about the transparency of data arises from attempts at secondary use. It is typically easier to reuse date which are well managed than those which are not. One indication of this is given by Lutkehaus (1990). Before embarking on her ethnographic study in Manam, an island close to Papua New Guinea, Lutkehaus read the fieldnotes Camilla Wedgewood had made in the same location 50 years earlier. Lutkehaus observes that Wedgewood carefully recorded, annotated and indexed her fieldnotes. In this, Wedgewood was partly following the promptings of her mentor, Malinowski. Lutkehaus also notes, however, that Wedgewood might have been influenced in her practice by the difficulties she had encountered using the fieldnotes of a deceased anthropologist, Bernard Deacon, who had left behind a set of field materials that were incomplete, unindexed and fragmentary.
Although secondary analysis is often couched in terms of sharing data (Sieber, 1991), 'sharing is useful if the user is provided with meaningful information rather than data' (White, 1991). To put this another way, data sets which are well documented facilitate secondary analysis better than those which are not. Miles and Huberman (1994: 281) point out that qualitative researchers 'don’t report clearly on their methodology because there are no shared conventions for doing so'. Miles and Huberman’s assessment is an understatement when it comes to fieldnotes. As Atkinson maintains, 'their construction and interpretation has been part of the tacit craft knowledge that is handed on from generation to generation as part of the oral culture of various disciplines' (1992: 19). The privacy of fieldnotes is a recurrent theme in major anthropological treatments of fieldnotes (Sanjek, 1990b). Indeed, some anthropologists seem to treat fieldnotes as 'sacred texts' with powerful and dangerous properties.

A third set of concerns relate to the analytic (re)assessment of the research. Here are issues to do with audit, meta-evaluation (the evaluation of evaluation studies) and replication. Schwandt and Halpern (1988) have advocated the use of audits to determine how far the findings produced by a qualitative study are dependable, confirmable and credible. The analogy here is with financial auditing where a third party attests to the financial health or otherwise of a company by examining its records and procedures. Auditing depends on the existence of an audit trail; a well-organized system of record keeping through which is documented the planning and progress of the research together with a contemporaneous running record of the investigator’s actions, thoughts and feelings. Miles and Huberman note that auditing of the kind Schwandt and Halpern endorse is rare. They go on to observe that the audit metaphor with its connotations of 'an external, stern, obsessive expert' (1994:282) is a forbidding one. However, they concur with Schwandt and Halpern in asserting the benefits which flow from audit-like procedures. From this point of view, operating as if research procedures are open to independent, external scrutiny improves quality by encouraging systematization, openness and reflection. Because establishing an audit trail depends so heavily on systematic and highly structured record keeping, it is costly in terms of time and effort. (Indeed, audits of this kind, if they are to be undertaken, probably need to be incorporated into the research process from its inception.) Computer-based methods in themselves encourage systematization and, to a degree, are self-documenting. They might, therefore, alleviate to some degree the burden involved in maintaining an audit trail. Moreover, qualitative analysis packages increasingly incorporate facilities for producing analytic memos and for allowing codebooks to be maintained. So far, however, the utility of computer-based methods to the auditing process has been asserted rather than demonstrated.
WHY USE COMPUTERS IN QUALITATIVE RESEARCH?  65

To some of those we interviewed considerations of demonstrable rigour were clearly important. 'In certain places it [package use] is much more systematic [than manual methods], and you can actually say "here is something rational", you have done this' (FG1). This theme has several aspects. One is that computer use involves an element of laying bare the analytic process to critical inspection. It also, however, bears the implication that procedures similar to those always available are now seen in a new, more flattering light precisely because a computer is being used. Perceptions of this kind are important in fields or disciplines where qualitative research is looked on with relative disfavour. The credibility of what is seen as a 'soft' methodology producing interesting anecdotes but unable to establish their generalizability is boosted by association with that 'hard' research tool, the computer. Two researchers, one a psychologist, the other a geographer, independently noted that qualitative methods were increasingly visible in their respective disciplines. Each made the additional point, however, that such methods faced criticism for their supposed lack of rigour. As one of these respondents put it in terms that were closely echoed by the other, 'People are desperate to find more rigorous ways of dealing with data, they are fed up with being accused of being soft and woolly. They are looking for ways to demonstrate to their colleagues that they have gone through a rational process and with [CAQDAS] they can hold it up and say "this can be done on the machine"' (FG1). In some disciplines, and in some countries, this view is established in academic as well as applied research. Moreover, there is a perception that using a package gives one not only respectability but also leverage. Thus a respondent from Latin America studying in the UK said of her country, 'there is a strong tradition in quantitative research because that gives more power to negotiate, to say you are doing hard research, and we can get more funding' (FG2). An educational researcher who worked in a department that was 'very statistically based, [with] obviously a lot of software around' found that using software made a difference in securing resources. 'My department was very willing to stand $100 so we bought three copies [of package], one for each of us. They were quite sympathetic' (FG1). Of course, there are 'credibility' issues within disciplines, too. Even though many researchers might be uncomfortable with the quantitative/qualitative distinction, it remains an abiding division. As one user suggested, 'people ... using qualitative data analysis packages said that they have always thought it gives them more ammunition to quantitative people' (FG3).

But disciplinary change and an interest in pushing forward the tenets of particular methodologies are not the only foci of the credibility issue. We encountered a number of cases where supervisors, despite their own lack of knowledge or expertise, had suggested to graduate students that
they might look into using computer software to analyse qualitative data. Other supervisors, though, could be hard to convince of the merits of software use. An element here seemed to be what one respondent described as the 'we've done it our way, what's wrong with you doing it our way?' argument' (FG6). Supervisors sometimes also harboured the kind of suspicions that people less comfortable with computers often bring to information technology, especially when confronted by a younger and more knowledgeable enthusiast. For example, the respondent just quoted encountered problems getting through her upgrade review because 'they wouldn't believe that qualitative methodology could be combined with quantitative methodology and both could use computing. So "you can't be using [quantitative package] and [qualitative package]!"' (FG6). In the event, this student got through the review, even if she felt that she had not entirely convinced her committee. In this, in fact, the computer aided her. In her qualitative work she had used the computer to document how her analytic thinking had evolved. 'I felt threatened that someone was going to take my thesis away from me [if] I didn't do it in the same way I did my quantitative stuff as well, tracking how I created variables' (FG6). In a similar vein, a different respondent commented, "That's what convinced my supervisor. Once I came out with a list of code words, "this is what I had a week ago and this is what I've got now" and "look, this is how it works" ... He said, "Gosh! Quite neat, isn't it"' (FG6).

Some research sponsors encourage use for quite other reasons. 'The first piece of [CAQDAS] software I heard about ... was quite a few years ago, when I was working in a research unit at [university] and my contact ... wasn't from a research point of view but from an administrative and cost point of view' (FG3). The research sponsor commended the package as saving analysis time. In a similar case, the department held a licence for a given package and standardly did all its applied research using it, so the researcher was 'stuck with it to some extent because the department has a case for using [package] ... and there was a resource element to it as well' (FG3). In other cases it is not a question of a sponsor imposing a requirement but of researchers feeling they are more likely to be seen as credible by sponsors if they use CAQDAS. 'One of the other attractions was the rigour that it imposed, not necessarily imposed because ... we were used to doing quite rigorous analysis with pencil and paper, but it was convincing outside people, government departments that we were doing research for. They're difficult to convince that qualitative research has anything to offer policy-relevant research for government. Because civil servants have a great turnover you are continually having to educate them. Someone else comes up who has never heard of qualitative methodology and you have to go through the whole process again, so we found it quite attractive that we could say "look, we use this but we are doing it this way" and that impressed them in their naive, ignorant way, which is
WHY USE COMPUTERS IN QUALITATIVE RESEARCH?  

useful for us. Wasn't all bluff from us, but it was a way, not of imposing rigour but emphasizing the rigour of it' (FG6).

While CAQDAS had been useful in justifying the credibility of qualitative research, it was suggested that at later stages the 'concreteness' of qualitatively based reports led to few queries from applied sponsors about what lay behind the analysis. 'No one has ever challenged us much about the analysis of qualitative data, from government departments... It seems paradoxical because they seem to be unwilling to accept it and we have to educate them to see it as a valid method but then when we come to present results it is, you are using quotations or whatever, concrete, they can understand it. And they go "oh yes", big pennies dropping all over the place but when you are presenting statistical [data] to them, my God, they'll go through it with a fine tooth comb!... We have never had to defend qualitative output in the sort of depth we have to on numbers ... It's much easier to present qualitative data because it's grounded in everyday language and if you write about it clearly it's very accessible' (FG6). Respondents were clear that the business of CAQDAS lending credibility was a phenomenon associated with non-academic sponsors. With applied research sponsors there seemed to be greater anxiety about legitimacy in general. 'It's respectability they are looking for, they ask "can I put your name on [the report]" or "can I say that you will do some consultancy on this" and "can you write which guides you have written about this and give a more scientific look to the whole thing" so I am finding increasingly my time is being taken up with making projects look more respectable so if you use a computer it must be better than if you don't' (FG4).

However, some respondents felt that there might in future be similar pressures on graduate students. 'It is also going to be a pressure for students because, as part of research training maybe, [it] makes it look more respectable if it's on a computer' (FG4). Against this was the peer review mechanism, which acted as a brake on innovation. 'I have put up three projects to [funding body] where I have had bum reader's comments based around the software, based on their absolute total and profound ignorance' (FG4). It is also worth recalling the apocryphal tale about the package that was allegedly banned in its country of origin because it made gaining a PhD too easy. The alternative version was that 'people were warned about saying "I will do qualitative analysis and I will use [package]." [It] would not earn many marks in their application for research grants. It was not considered a particular advantage any more because everyone was doing it' (FG5). Thus, CAQDAS might lend credibility at an early stage for reasons of innovation, but once the innovation is established the issue is one of how well the research is being done.

Boosting credibility is perhaps a narrow way of looking at the matter. Anything that heightens the sponsor's interest in the workings of the research and the basis of the analysis is to the good. Two users offered examples of CAQDAS having such an effect. 'I work with [police force] at
the moment. They are actually coming in to query the database themselves and I am finding that works very well, I am not driving the analysis in any way. They are getting what they want out of it. It's been quite easy to show them how to use the package. I was somewhat concerned that their interest in the research project, which after all they are funding, might have got completed. By the person driving the computer, by making it easy for them to come in and either follow what I'm doing or actually do it themselves has had a great payoff in the satisfaction they are getting out of the research' (FG4). Since the police were also the subject of the study it is not just a heightened involvement of sponsors but of subjects which was facilitated. In like vein, 'when I have got people I am working with to the point where they can articulate what they want then I can hand it over to them, but it is just the simple process which every researcher knows, you do need to be in almost continuous contact with them' (FG4). Thus, there is more to the matter than a 'philistine' applied scene and a 'pure' academic one. CAQDAS users appear every bit as concerned about preserving the canons of qualitative enquiry as are non-users. Here is a comment from a user in an academic setting: 'once there is a link to real, theoretically informed research that is going on it doesn't really matter what program. It stops being an issue of advocating [package], it's about enhancing methodology, enhancing analysis, allowing a researcher to maximize their time' (FG5).

As we indicated, several respondents felt that CAQDAS stimulated thinking about criteria of proof, adequacy and so on. 'It actually inspires and motivates debate, it makes you revisit issues around qualitative and quantitative philosophy, and puts these issues back on the table. And thinking about technology in the social sciences' (FG4). The phrase 'puts these issues back on the table' suggests that CAQDAS puts familiar epistemological debates in a new light. We must be careful to use verbs like 'encourage', 'stimulate' and 'prompt' when we are considering the impact of software on qualitative analysis. Agency remains with the user, not the package. There is testimony from our respondents to the practical and analytic advantages of use. There is further a recognition that good practice principles are achievable manually. These principles were laid down before computers. But there is also a recognition that there is more to be gained from CAQDAS than the practical benefits of data management. It is perhaps like the introduction of any significant innovation: it helps us question and explore our assumptions anew. We have also seen that, here again, the research environment mediates the benefits. It will be a similar story as we turn to disadvantages.

Disadvantages

We have made the point elsewhere (Lee and Fielding, 1991) that, by and large, proponents of CAQDAS have been relatively modest and non-utopian in their claims, at least when compared to 'computerization
WHY USE COMPUTERS IN QUALITATIVE RESEARCH? 69

movement' enthusiasts in other fields (Kling and Iacono, 1988). Nevertheless a variety of cautions, objections and criticisms have been advanced against the use of computers in qualitative research. We examine here some suggested disadvantages of computer-based methods in qualitative research. These are (a) that problems exist in relation to the accessibility and availability of CAQDAS programs, (b) that the practical benefits of computer-based methods have been exaggerated, (c) that computer use distances qualitative researchers from their data, and (d) that the introduction of computer-based analytic methods in qualitative research might encourage users to emulate some of the more problematic aspects of survey research.

Accessibility and Availability

Russell and Gregory (1993) have drawn attention to factors that in their view restrict the accessibility and availability of CAQDAS packages. They point out, for example, that the move to computer methods has financial implications. While costly of time, manual methods of qualitative data analysis impose little monetary burden. As Russell and Gregory indicate, the hardware requirements – and therefore the cost – of machines capable of running CAQDAS programs are not always trivial. Especially where the software used is complex, the user will typically need a machine that has a hard disk, a fast processor and a substantial amount of RAM. Moreover, there seems to be a trend for requirements of this kind to creep upwards over time. This is not necessarily undesirable since increasing hardware needs often reflect the growing sophistication of the software. Nevertheless, developers and software companies are frequently over-optimistic about hardware requirements. The question is not how much RAM is needed simply to run the program, but how much does the program need to be used productively? Among our focus group participants, graduate students especially had suffered, the frustrations of using relatively powerful programs on low-end machines. In the end, they became adept at finding other things to do while waiting for the computer to finish some lengthy process. One respondent, for example, reserved the early morning for certain procedures. After setting the computer to work, she had breakfast, leaving it to grind away at its allotted task. Eventually moving to a more powerful machine, this informant reported that she could now do comparable analyses in minutes.

On the other hand, although software costs are hardly negligible, Russell and Gregory's (1993: 1810) suggestion that the 'cost of ... qualitative analysis programs is often prohibitive' seems wide of the mark. Weitzman and Miles (1995) provide information on 23 packages potentially of use to qualitative researchers. Taking into account educational discounts, the average cost of a package described in their book is
US$238.25. (The cheapest package, priced at $25 and of a somewhat specialist character, was excluded from the calculation to avoid artificially depressing the mean.) On average, text retrievers, almost all of which are commercial packages, are somewhat more expensive than dedicated qualitative analysis packages produced by academic developers. Perusal of advertisements in computer magazines suggests that high-end word processors had a 'street price' of around US$290 around the time Weitzman and Miles were writing. Given the relatively small size of the overall market for qualitative analysis software, one judgement might be that their pricing is in fact quite competitive.

Another way to look at this question is to explore the issue of software piracy. We found little evidence that CAQDAS software is widely pirated in Britain (Lee and Fielding, 1995). Obviously in our focus group research we dealt with people who were non-pirates by definition; they had registered their software. We enquired, however, if they knew of pirated copies of software and asked them about their own reasons for not pirating software. There seemed to be general agreement on three points. First, the generally low cost of CAQDAS programs – at least relative to commercial business-oriented packages – discourages piracy. To put the matter bluntly, at current pricing levels users do not feel ripped-off. Second, in the British context at least, recent legislative changes seem to have shifted the culture in universities towards one that is anti-piracy. Departments, it seems, are increasingly concerned about illegal copies of software and are policing their use. Having said that, the temptation to pirate software seems in part to be related to the availability of the software. Several focus group participants noted that packages had taken a long time to arrive from abroad and one researcher had resorted to using a pirated copy of unknown provenance while waiting for a legal copy to be delivered.

Russell and Gregory observe that there are difficulties associated with the portability of specific packages and the transferability of data. Researchers reliant on mainframe packages for qualitative analysis may have difficulties, for instance, if they move to a site which does not support the package. Russell and Gregory (1993: 1811) also suggest that 'manual data management systems are often more flexible with regard to data sharing' than are computer-based systems. Qualitative analysis software came of age as, or even after, the desktop revolution in computing occurred. Since mainframe packages are presumably now used only by a small minority of users, portability would not seem to be a problem of major dimensions. Certainly, it was not an issue that had affected any of our focus group participants. Given the physical awkwardness of paper-based data, we would question Russell and Gregory's contention that manual systems allow data to be shared more easily than computer-based systems. Having said that, as some of those in our focus groups reported, problems in transferring data occur in mixed hardware and software environments. The co-ordinator of the project mentioned elsewhere in
which four different researchers collected data for each other recorded, for
instance, that 'researchers in other institutions are supposed to send me
disks with the file in Word which I can read . . . In fact they don't use [an
interchange format], they send it in other things, which then have to be re-
done' (FG3). Others reported difficulties where some team members used
Macintoshes and others used PCs, while for one research team program
use was impeded because they had to share a single PC between three
researchers.

The Possible Exaggeration of Practical Benefits

How far the potential benefits of CAQDAS have been realized in practice
can be questioned. Implicit in claims that computer use in qualitative
research fosters paper reduction, the speedy finding of material, and flexi-
bility in the coding process is an assertion that using a computer increases
access to one's material. As Weitzman and Miles (1995: 335) point out,
however, in practice things were somewhat different. Some programs
have awkward interfaces. Others have features like facilities for quasi-
statistical output or for hypothesis testing which leave many researchers
feeling uncomfortable. In some cases, too, developers fitted 'governors' to
their programs. These were ways of working, such as the need to code data
off-screen before input to the program, which were built into the software.
Almost always the intention was to make matters slightly more difficult
for the researcher so as to inhibit facile analysis. Very often, however, as
our focus group data clearly show, these were simply perceived by users
as a hindrance.

There are no clear estimates of how long it will take a researcher to
obtain a good working knowledge of a particular package. The experi-
ences of our focus group participants suggest, however, that at their
present stage of development CAQDAS programs offer troublesome,
though by no means insurmountable, barriers to self-teaching. One user
testified to being 'stuck at the first hurdle, had many phone calls to [devel-
oper] before I sorted it out, had to write myself an idiot's guide, which was
stained with blood, sweat and tears . . . not knowing whether things were
going wrong through my own stupidity or mistakes I was making' (FG3).
Another user described encountering bugs in the program she used,
which even the developer proved unable to resolve. Told of a fault, his
disarming response was simply to express surprise – 'he wasn't even
aware that it wasn't doing it, which was even worse!' (FG3). In some
instances, users found themselves being thrown back onto local support
services. In many cases this proved to be unsatisfactory. Although we are
aware of sterling exceptions to the general rule, computer support staff are
often not entirely helpful. CAQDAS programs are not well known outside
the qualitative analysis community. Moreover, many computer support
staff have backgrounds in the natural sciences or in statistics and sometimes have difficulty in relating to qualitative analysis.

Since later versions of programs are generally easier to use than the earlier ones, it might be that some of our users paid a penalty for being relatively early adopters. Weitzman and Miles (1995) rated each of the 23 packages they reviewed on a three-point scale: strongly user-friendly, OK, and weak. Inevitably, of course, judgements of this kind are subjective. What is to an expert a minor inconvenience capable of a 'workaround' can seem a terrifying and insuperable problem to a novice. Nevertheless, Weitzman and Miles rated 14 out of the 23 programs as 'strongly user-friendly'. In only four cases did they rate the user-friendliness of a package as 'weak'. The proportion of strongly user-friendly programs falls somewhat if one counts DOS versions of programs rather than their Windows counterparts. This in itself points to a trend which has become increasingly apparent in recent years. As is true more generally in the computing world, the developers of qualitative analysis software are moving in the direction of interfaces which are easier to use and which conform to industry 'look and feel' standards. These developments make considerably easier transfer of learning from one package to another. (Of course, developments of this kind might have to be bought at the cost of more substantial hardware requirements.)

The computer's ease of use can have possible paradoxical consequences. Russell and Gregory (1993) contend, for example, that because the computer is easy to use 'researchers may perform meaningless queries instead of focusing time and energy on productive directions'. None of our respondents reported this as having been a problem for them. Some clearly recognized, though, a downside to the creative play aspect of using a computer: 'I can spend an afternoon just playing with it, put together codes … and see what I get' (FG3). (It might be possible to argue that the unproductiveness of a particular line of enquiry is likely to become more evident earlier when the computer is used than in a manual analysis.) To take another example, claims that computer analysis takes less time than manual analysis might not be entirely justified. Because coding data is easy in a computer-based system, the time saved during analysis can be lost again if users do more extensive or more sophisticated coding than they would have done manually. (One of our respondents reported that coding in full an admittedly lengthy transcript – over 100 pages – took seven hours.)

Subject to time pressure, use might be suspended (Lee and Fielding, 1995). But users were loath to say they were abandoning CAQDAS. While there was a degree of disillusionment on the part of some respondents, most said they would continue use. One user remarked, 'I always intend that I am going back and use this thing and interrogate it more. I am sure I will' (FG4). Another who had been very frustrated with her experience of CAQDAS commented, '[Y]ou were asking whether [package] has
WHY USE COMPUTERS IN QUALITATIVE RESEARCH? 73

actually affected the project ... [It] has affected us detrimentally, and I think it is partly because of our inexperience, but the psychological block that we have developed as a team ... killed a lot of the enthusiasm and interest ... The whole thing became coloured by our hatred and resentment [of] the bloody program' (FG1). (We consider this case in detail in Chapter 5.) Despite these remarks, however, the respondent was still strongly committed to CAQDAS and saw her problems as arising as much from the team research environment as from the software: 'I can still see it has got value and I still am irritated by this thing and I don't like being beaten by it.' She had 'always had a love/hate relationship with computers' and planned to use the package again on her next project, but allowing more time for data entry and coding. Others, who had struggled with large data sets, were less likely to consider using the computer again for small-scale projects, because the set-up time is considerable with CAQDAS and manual methods can cope with limited data sets without being swamped. However, even here there were exceptions. 'Now I think with small data sets I wouldn't [work manually], because I would think that's easy on the machine' (FG1).

For some applied users, the advanced analytic features available in some packages were 'overkill, [package] was too good, it was often too much for the sorts of work you were doing' (FG6). Others worried, however, that they were failing to realize the full potential of the software. Some referred to using the package in 'quite a manual way' (FG3). Thus, 'I have been totally unadventurous with it, all I have been asking it to do is to give me a print-out of all the answers to question one together, all the answers to question two, and there I stop. Last year I tried to categorize it and it took such a long time to sort out I wouldn't dare do it again' (FG3). Some of our focus group participants felt that program documentation often did not give enough emphasis to more sophisticated features. Many of the process descriptions offered were of a straightforward retrieval strategy, recovering the data coded with a given category. There appeared to be a gap between this approach, which is the basic use to which a code-and-retrieve package can be put, and work on category development. T find I use the very basic sort [procedure] ... Partly coward's way out but also some of the things are fairly subtle and complicated and it would take so long to organize the material' (FG3).

Closeness to Data

A worry for many qualitative researchers is that computer methods discourage involvement in and engagement with data, inviting the analyst instead to skim the surface of even the richest material. Agar, for example, encapsulates this anxiety by recounting a metaphoric nightmare. Two studies take place side by side, 'a lousy computer analysis and a beautiful
analysis done by hand'. In the nightmare, however, the 'community of scholars would immediately gather round the printout and celebrate its form rather than its content' (1991: 185). This is not an idiosyncratic concern. Fears that using a computer distances analysts from their data are recurrently expressed on electronic bulletin boards relating to computers and qualitative research. As Weitzman and Miles (1995: 335) comment, 'Judging from the electronic mail we've seen, many users place high value on "staying close to my data".' In assessing these concerns, we concur with another observation by Weitzman and Miles; what we are dealing with here is a complicated issue'. Raising the issue of closeness to the data in relation to computer use implies that not using a computer will ensure that the analyst stays close to the data. Yet dissatisfaction with manual methods and associated worries about being overwhelmed by the data suggest that manual analysis does not of itself guarantee a rich and fruitful closeness to the data. A further complication is that the phrase 'being close to the data' has a wide range of possible meanings most of which are rarely explicated.

Staying close to the data can mean being able to recover the sights, sounds and experiences of being in the field. Data, in other words, have an evocative character for some researchers. This is nicely, indeed evocatively, captured in the comment made by one of the anthropologists Jackson interviewed: 'For example, you write about a sacrifice, how it's done. When reading my notes I remember how it smelled ... everyone's really pleased when it comes time to eat it' (1990: 13). For others, 'closeness to data' seems to be related, not necessarily in a very explicit way, to ideas about retaining the holistic character of data. At least for some of our focus group participants there remained an ineffable quality to data which software could not capture and that they experienced as a loss. "There is something missing from data perceived in one way which was present in data perceived in another way, and software can't manage that' (FG3). For others, computerized code-and-retrieve based procedures fragmented the data in ways that made it difficult to see the wider picture. One said: 'because it was much easier to churn it out in little segments, that took away your interest with the whole transcript' (FG2). As Dey (1995) points out, the problem here is that in code-and-retrieve procedures text segments coded in the same way bear an apparent conceptual relationship to one another. Once retrieved, however, each segment appears divorced from the context in which it originally appeared. Focus group participants clearly felt that seeing results in context remained an important part of the craft of qualitative research. One user who declared 'I don't trust computers. I still do manual calculations' felt that the same held true for CAQDAS: 'if you just go into it and get the results you can't understand them unless you can contextualize the information you get' (FG3).

Having said that, to our mind some of the concerns about data fragmentation expressed in the literature might be overdrawn. The notion of
'context' implies boundedness. If it did not, the only context we could have for a segment of text from an interview transcript, say, would be the entire transcript. Presumably, in fact, most analysts work with an implicit idea of local context, and it is this which determines where the boundaries of a given segment are located. Thus in classical grounded theory open coding is based on an indicator-concept model (Glaser, 1978). Presumably, the boundaries of coded segments are determined in this case by how far the degree of fit between concept and indicator is compromised by extending the segment. Alternatively, in the kind of 'loose, inclusive coding' advocated by Becker and Geer (1960), segment boundaries are left deliberately wide, but subject to revision – presumably based on an understanding of context – as an analytic model is pieced together. What we would concede is that the concerns about fragmentation are clearly relevant whenever sequence is a consideration in data analysis. When something of interest is embedded within a narrative, refers back to something said or observed at an earlier juncture, or where one is interested in transitions from one topic to another, fragmentation is a problem. One solution might be to develop more complex coding approaches (Riessman, 1993). On the technical side, developers have experimented both with proximity searching (Drass, 1989) which allows the recovery of sequences of coded text, and the use of hyperlinking techniques (Dey, 1995), the non-linear and associative character of which readily permits the recovery of context.

Sometimes, when researchers refer to being close to their data what they have in mind are the tactile and perceptual aspects associated with data handling. Jackson (1990: 13) notes that 'quite a number' of the anthropologists she interviewed reported that they derived pleasure from handling their fieldnotes. In a similar vein, Seidel records that a researcher who had begun to use ETHNOGRAPH 'called me up and said he missed having his piles of xeroxed copies and note cards lying around' (1991:114). Staying close to data in this sense implies that material gathered in the field retains a tangible character capable of being experienced in a tactile way. Receiving positive tactile feedback from a task presumably makes that task more pleasurable. This in turn sustains interest, aids motivation and increases one's sense of accomplishment. Interestingly in this regard, Converse (1987: 383) indicates that when in survey research computers replaced counter sorters as an analytic tool, researchers felt that some of the craft was lost. Presumably part of the source of these feelings was that analysts could no longer handle their punch cards or literally see distributions emerge as the machine differentially sorted cards into receiving bins.

The depth of the perceptual field an analyst is able to bring to data is affected by the form within which those data are encapsulated. A computer screen contains only something like 24 lines of text. It becomes almost literally a window, framing data in possibly useful, but always
constraining ways. It is less easy to see, for example, the structure of a text on a screen than on a set of printed pages which can easily be flipped back and forth (Pfaffenberger, 1988:19). In an article provocatively entitled 'The right brain strikes back', Michael Agar describes how in the early stage of a fieldwork project he eschewed computer use in favour of finding an empty university classroom with a number of blackboards in it. He would write and erase ideas on the boards, turning from one to another to scan the current state of his thinking on the assumption that 'simultaneous visual access to material is what makes ideas happen' (1991: 193). C. Wright Mills advocated periodically spilling file folders out onto the floor and re-sorting them as a method for stimulating creative reflection on one's data. Such a procedure allows for 'perspective by incongruity' (Mills, 1959; Burke, 1969), the possibility that by juxtaposing seemingly dissimilar objects one will discern a non-obvious relationship between them. It is hard to think of a computer analogue of this procedure. Many common types of computer program contain a sort facility. Rarely, however, do they allow one to sort in a random order.

Quite a number of our focus group participants were clearly anxious about the impact computer use might have on their tactile and perceptual relationship with data. Speaking of an early concern, a user reported worrying 'that I might end up looking at the screen more than I was looking at my transcripts, the hard copy was the thing that I clung to, I was very concerned that I would be diverted to getting lots of neat piles of categories with this, that and the other, and forget about that first interview' (FG2). Or: 'unless I can hold this in my hand it's not real. I find it very difficult to read things for meaning on the screen' (FG6). The solution was to balance on-screen work with work using hard copy. Several users suggested that this issue could be resolved if one had done the transcription oneself. If so, there was sufficient familiarity to work largely on-screen. 'If you have done the transcription you can code it straight into the screen much easier but if you haven't then you need to go through the process of coding it all off-screen to get in touch with it' (FG1). In any case, as a number of users indicated, 'manual' elements were far from absent when they used the computer to analyse their data. 'I spend a lot of time reading the transcripts and getting them into shape [so] they are ready for [package], and I still do a lot of memoing and scribbled-down notes' (FG3). Others said: 'I don't think it is possible to rely on it [package] entirely' (FG3) and 'does anybody actually advocate doing everything on the screen? Never picking up your pen, that would be impossible' (FG3). Interestingly in this context one focus group participant who had a computer support role made a contrast with SPSS. In her view, SPSS was a good instance of software that required a high degree of abstraction. CAQDAS, on the other hand, was helpful because it supported either formal abstraction or close-to-the-data approaches. In general, 'one has got to evaluate the context of any given project, how far one wants to take
WHY USE COMPUTERS IN QUALITATIVE RESEARCH? 77

the application of a rigorous structure. There must be room for accommodation of research styles' (FG5).

Some of the distancing effects which inhere in current packages might be ameliorated by developments in operating system capability and interface design (Weitzman and Miles, 1995). Such developments would include the simultaneous display on screen of codes and data and what Weitzman and Miles call 'pencil-level richness', the ability to mark and annotate screen displays in flexible ways. We would also add to this what we suspect will be an important development in promoting a rich and close involvement with data: the substitution of digital video for analogue audiotape as a data-logging medium in qualitative research. Beyond this, as Bryan Pfaffenberger (1988) has astutely observed, the only protection against the power of the computer to shape work in hidden ways is critical awareness. This requires reflection on the analytic process itself, independent of the computer, as well as awareness of what it is about the technology that operates in undesirable ways.

One final comment about the issue of closeness to data is that, whether or not the computer does distance researchers from their data, it challenges researchers to be clear about their assumptions. In earlier writing, we used the metaphor of Frankenstein's monster to refer to the fear some researchers felt about the impact of computer methods on the craft of qualitative research (Lee and Fielding, 1991: 8). When we used this metaphor we had in mind the standard cinematic representation of a mute and horrible creature visiting mayhem and destruction on those who had created him. As many commentators have pointed out, this is far from Mary Shelley's original conception. That story can be read as counterposing an enlightenment view of science, reason and progress typical of the eighteenth century with 'early nineteenth century Romantic perspectives that probed beneath the tranquil face of Nature and man to confront the dark and passionate powers therein' (Smith, 1994:58–9). The desire not to be distanced from one's data in part reflects a stance which qualitative research inherited via the Chicago School from later German Romanticism. Faced with the disappearance of peasant culture under the impact of industrialization, writers in the Romantic tradition undertook direct observation of rural life. As Gouldner (1973) points out, by regarding the marginal, the dispossessed and the lowly as worthy of study, Romanticism implanted in the social sciences an 'open' or 'democratic' conception of data in which nothing social falls beyond the purview of the social scientist. Alongside this democratic conception of data, however, Romanticism also encouraged a celebratory stance towards data; one which 'wants and appreciates the object in its concrete totality, in its uniqueness and individuality' (Gouldner, 1973: 352). Grounded in what Gouldner terms a 'collector's aesthetic', this celebratory stance can encourage a fascination with the data themselves at the expense of the analytic process. Thus Anselm Strauss records his irritation with students in his seminars being
'entranced by data' (1987:162) in ways which inhibit their capacity to get on with analysis. In a critique that is partially self-directed, Lofland (1970), too, criticizes qualitative researchers for a failure of analytic nerve which discourages them from the task of transforming richly descriptive data into wider conceptual and theoretical understanding. Such failure is self-perpetuating because it inhibits the search for alternative, and potentially richer, conceptual frames. It is also, according to Lofland, abetted by the difficulties which inhere in the analysis of qualitative data. From this point of view, anxieties about computer use might not reflect reluctance to use new technology so much as an ambivalence about the analytic process itself. The computer promises to make the analysis easier, but it also forces the researcher to confront the need to embark on the analytic journey. In at least some cases, we would contend, concerns about the alleged distancing effects of the computer also embrace a reluctance to take that fateful step.

Related to this, it is interesting to note that some of the anthropologists Jackson (1990) interviewed felt embarrassed about their fieldnotes. One of them said, for example, 'Rereading them, some of them look pretty lame. How could you be so stupid? Or puerile?' Although we found nothing comparable to this in our focus group material, it is at least conceivable that some researchers are actually resistant to looking back over field materials because these remind them of their mistakes. Computers are, of course, both systematic and without feeling. Computer-based methods might ensure that particular materials are not discarded or discounted because of the researcher's negative feelings towards them.

Unintended Consequences

There have been some worries that instead of studying social settings and processes in depth, qualitative research, under the impact of the computer, might come eventually to mimic survey research, taking a broad sweep across a large number of cases. In other words, computer use might encourage qualitative researchers to trade resolution for scope, to use John Seidel's phrase. Seidel does not argue that working with large amounts of data is intrinsically wrong. His fear is that computer use will foster an illusion that large data volumes can be handled adequately. In exchanging resolution for scope, however, researchers 'will end up missing interesting and important things in the data' (1991:109).

We sought in a number of ways to explore how far the availability of CAQDAS might have affected sample sizes. We have fairly detailed information about 15 of the projects undertaken by our respondents. In Table 3.1 projects are ranked by sample size within data collection method. Methods are ordered from those we take to be less structured to more structured approaches. We have distinguished projects which we
considered to be academic in scope from those having a more applied focus. We have also indicated the broad substantive area within which each project was located.

These projects are not, of course, representative of any wider population. But it is worth making several comments about the distribution of sample sizes. First, the median sample size for the 15 projects is a surprisingly large 100 cases. While the smallest sample had 10 cases, the largest had 646. Academic studies (median = 56) tended to be smaller in terms of sample sizes than applied projects (median = 150). Large sample sizes are apparent for each data collection method. As one might expect, though, the general tendency is for less structured methods to be based on fewer cases than more structured methods.

We found little in the testimony of our focus group participants to suggest that packages were encouraging researchers to feel they could handle larger samples than usual in qualitative research. Instead it seems that what had led people to look into CAQDAS in the first place was that they had to deal with a large or varied amount of data. Rather than the use of large sample projects being attributed to CAQDAS, decisions about sample size more clearly reflect sponsor or peer expectations or are simply the product of the specific methodological stance taken by the researcher. ‘You have to make some commitment to representativeness and you are talking about housing, the range of people’s homes is enormous, number of categories, so in fact it is driven by some sense of sample size, and there are certain objectives you can make within a research project and that drives the number of interviews rather than the qualitative software …
before I started interviewing I knew that I was going to have so much information I would need help managing it' (FG3). Someone who gave computing and research methods advice to students and others at her institution replied to a direct question about how far computer use might be driving up sample sizes as follows: 'It's tending to be the other way round. People have done collections of large data sets and they face the risk of underreporting them. Most research projects get their funding first … and they say "Can you find me some [software to handle the data]?"' (FG4).

Because concern about sample sizes is often expressed as a worry about overall trends in qualitative methods, we decided to supplement our focus group data with material we hoped would give us a wider picture. To do this we downloaded from Sociofile abstracts in which the phrase 'depth interviews' appeared (n = 556). Depth interviewing does not of course exhaust those methods regarded in the social sciences as qualitative. There were two reasons for focusing on interviewing rather than say participant observation. First of all, we wanted to keep the analysis simple. We also reasoned, however, that, if there was a trend towards a rise in sample sizes, interview studies were likely to be sensitive to it. Abstracts not directly reporting an empirical study were excluded from the analysis, as were those which contained no specific information about sample size. Table 3.2 shows details of the distribution of sample sizes for 1991, 1992 and 1993 (the last complete year for which data are available). Three earlier years, 1977, 1983 and 1987, are shown for comparison.

The average median sample size for the years shown is 37.4 (s.d. = 13.79). Inspection of the table does not suggest a linear trend towards sample size inflation; if anything the median sample size is declining slightly. In each year, however, a small number of studies appear which have rather large sample sizes. All abstracts which revealed a sample having over 100 cases were inspected individually. It seems that in many cases these were quasi-quantitative studies. Of the 21 abstracts we inspected, eight (38 per cent) indicated a definite use of quantitative techniques such as cluster or factor analysis. In another four cases, while analytic techniques were not mentioned, descriptions of the findings were firmly couched in distributional or associational terms. (In the remaining

<table>
<thead>
<tr>
<th>Year</th>
<th>Minimum sample size</th>
<th>Median sample size</th>
<th>Maximum sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>3</td>
<td>50</td>
<td>180</td>
</tr>
<tr>
<td>1983</td>
<td>3</td>
<td>40</td>
<td>403</td>
</tr>
<tr>
<td>1987</td>
<td>5</td>
<td>37</td>
<td>332</td>
</tr>
<tr>
<td>1991</td>
<td>1</td>
<td>28</td>
<td>575</td>
</tr>
<tr>
<td>1992</td>
<td>5</td>
<td>33</td>
<td>150</td>
</tr>
<tr>
<td>1993</td>
<td>5</td>
<td>35</td>
<td>150</td>
</tr>
</tbody>
</table>
WHY USE COMPUTERS IN QUALITATIVE RESEARCH?

Finally, we decided to look at all those instances where the words 'depth interviews' appeared in an abstract which referred to journals largely devoted to reporting qualitative work. Specifically, we looked at the *Journal of Contemporary Ethnography* (and its forerunners, *Urban Life* and *Urban Life and Culture*) and at *Qualitative Sociology*. The numbers involved were not large enough to explore trends. However, the overall picture is quite consistent with the patterns already seen. Some 18 articles were identified which reported the results of empirical research and which indicated a sample size. (No information on the size of sample was available for one further case.) The median sample size was 30 with a mean of 43 (s.d. = 36.05). The smallest sample size was seven; three studies reported sample sizes of 70 or over.

One implication of our focus group data is that researchers involved in such studies based on large samples adopted software early in the hope that it would ease potential data management problems. An obvious way to test this hypothesis would be to look at studies in which software was used and to compare the sample sizes from those studies with sample sizes in a more general sample of qualitative studies. There is no very easy way to identify studies of the former type. We did, however, inspect the *Social Science Citation Index* for studies citing John Seidel's descriptive writings on ETHNOGRAPH, one of the earliest and best-known packages. We did not have the resources to track down these studies and examine them. It is interesting to note, however, that of the 163 articles we identified only about a third (35.6 per cent) were single-authored, slightly more than a quarter (29.5 per cent) had two authors and 35 per cent had three or more authors. For comparison, we took a one-in-ten sample of our depth-interview studies. Of the 56 articles we looked at, just under two-thirds (64 per cent) had only one author, 18 per cent had two authors and a further 18 per cent had three or more authors. If we make the assumption that multiple authorship is more likely in large-scale studies then this provides additional evidence that researchers needing to collect large amounts of data have been particularly attracted towards computer use to analyse their qualitative data.

In addition to exploring the impact of CAQDAS on sample size, we wondered if computer-based methods were chiefly seen as relevant to the analysis of interview data. We were aware that some packages were influenced by approaches developed for the analysis of documents (in particular, some European packages have strong links to the tradition of hermeneutic analysis). Interviews are likely to have more thematic coherence than fieldnotes and/or at least some kinds of documents. While there might well be interpretive problems, respondents are responding (if only nominally) to similar stimuli on the same kind of occasion and often under similar conditions. This tends not to be true either of fieldnotes or of
documents prepared by third parties where there are processes of selection and expression which are more or less non-reviewable by the user. Is CAQDAS, then, necessarily of restricted application in the analysis of data from sources other than interviews?

One should not underestimate the value of CAQDAS as a data management device; it can take considerable effort to organize and keep track of data from multiple sources, even when it is specific to a particular case. One project, for example, involved multiple interviews concerning a particular work activity as well as independent but linked observations of a separate activity. As noted above, purely observational studies and purely document-based studies were in the minority in our sample, although field observation was sometimes a component of the mixed method studies as well. Several users reported holding fieldnotes in CAQDAS as well as interview data. Thus, 'it is just as appropriate for doing fieldnotes if you know what you are going to use your fieldnotes for... It is useful to put on [package] just as a marker that there are fieldnotes on this person so that when you are looking at your transcripts you know there may be other relevant material. Because things like that get lost' (FG3). In fact where mixed data sources were used, researchers seemed to entertain no inhibition about using these kinds of data alongside interview data in the package. Thus 'the way I then back up the interviews is by looking at project files, so correspondence files, project diaries which firms hold, minutes of meetings, site meetings, briefing meetings... that does go into the package as well. It's confirming. One thing I ask is, "how much time do you spend working on a project?" ... and I can then back that up by looking at their time sheets' (FG5). This is a very explicit account of use for confirmation and this verificationist approach to triangulation has to assume equivalence of meaning to fulfil its logic. Thus 'really the only reason I do look at the other data is as a back-up to the interviews, so somebody says "how do you know we are telling the truth?" or "how do you know you are getting the right type of answer?" [and] I can say "well, I looked it up, information to back up the case".' Similarly, 'again, I do like [respondent just quoted], using some other, apart from my interviews, background information as justification triangulation, information about the context where respondents are working' (FG5).

As noted, six of our focus group participants were currently working on research that included a quantitative element in the research design. Although we had suspected that qualitative data analysis software might frequently be used to analyse data from free response questions on an otherwise quantitative survey instrument, we found only two such instances. Of the remaining mixed method projects, one involved repertory grid (quantified responses, quantitative analysis) alongside diaries and interviews, one a short survey used to sum up basic information about the sample which was pursued in more depth by interview, another a large
longitudinal quantitative study with a significant but subsidiary use of interviews, while the remaining project was part of a series of educational evaluations which combined quantitative survey data with qualitative data from case studies. There was little evidence, in other words, of CAQDAS being used to 'mop up' messy data on projects that were otherwise statistically driven.

We found both 'basic' and applied multi-method research among the projects on which our focus group participants were working. In other words, it was not only quantitatively based projects that came with an applied focus. However, in line with other work (Caracelli and Greene, 1993), we also found little evidence that different methods were extensively integrated when the time came to analyse the data. This might have some implications for the extent to which 'conceptualizing or 'theory-building' features of CAQDAS are used. Thus, 'there are questions carrying the survey and which touch on the area [covered by qualitative methods] ... but they are not integrated beyond that, the quality is very different really, so they are not integrated beyond that' (FG1). This despite the fact that this informant both endorsed and enjoyed working with multiple kinds of data. It should be noted that most users referred to making some basic quantification of their qualitative data (often referred to as 'content analysis') at least in the sense of counting code frequencies and the like. One statistically trained user who worked in an advisory capacity was able to identify for quantitatively oriented researchers three distinct uses to which they could put qualitative data analysis software. For researchers who ‘insisted that [the data] then went into SPSS I used [text analysis package] to help me construct the sensible categorization scheme. Other people I was able to convince that it was there for their searching, for them to pursue their own lines. And the last perhaps trivial but often very welcome use was the quotable quotes to put alongside the pie charts and the statistical tables’ (FG4).

In the case of a large-scale study of religious beliefs, something more elaborate was being attempted. Consideration had been given to the commensuration of the data from different sources. The analytic strategy was interesting: rather than simply looking for points of connection, the difference between the two types of data was recognized, and reflected in the use of two different packages, a statistics package and a CAQDAS package. The researchers appeared to feel that the difference between quantified and textual, discursive information about the same referent should be preserved rather than seeking integration: they represented different aspects of the phenomenon being studied. 'We don't want the qualitative to duplicate what we have already got on the quantitative – pointless' (FG4). The problem of ensuring the comparability of the data is also manifest in another, perhaps more frequent, occurrence in qualitative analysis, where new data are collected after a conceptualization has begun to be built up. Since many research designs include pilot fieldwork this is
quite a common occurrence, and the issue again arises of the researcher's evolving perspective leading to different aspects of the phenomenon being perceived as relevant. This quality of perspectival emergence can, however, be 'tracked' quite effectively by annotation features allowing memoing and the use of source tags (in some cases, these automatically date an entry or a retrieval). 'Most of the other things I would do would be recording information about indexing categories and you can do that fairly easily, change an index category, you can do that on [package], 'cause you can bring up comments on a category you have created and work out how it has changed' (FG3).

In summary, CAQDAS can be used as a 'gateway' to a statistical analysis, for example through export to a statistical package or by code-frequency counting. There seems little evidence, at least in our data, that this led users to analyse qualitative data quantitatively. Nor does CAQDAS appear to appeal only to researchers with a quantitative bent, or to people who only do qualitative research. Among our sample, it appears that 'pragmatism rules', though we would add the caution that methodological eclecticism needs to keep in view problems relating to the validity of comparisons.

**Conclusion**

The experience of our focus group participants seems not to validate the strong claims of either proponents or critics of computer-assisted qualitative data analysis software. Users experienced difficulties and frustration in learning and using packages. With some exceptions, the computer delivered data management benefits rather than transforming analytic practice. Against this, those who, for various reasons, were uncomfortable with existing manual methods of analysis generally found the systematicity of computer-based methods helpful, and perhaps even in some situations liberating. On the other hand, users also resisted those aspects of computerized methods that seemed to them to compromise the craft elements traditionally associated with qualitative research. For the moment at any rate, we would suggest that a verdict of 'not proven' is entered against the charge that CAQDAS encourages users to mimic survey practice, for example by collecting data from larger samples. As a general rule computer use does seem to lend legitimacy to qualitative research. It allows researchers both to emphasize the traditional strength of qualitative research and to counter misconceptions about qualitative data analysis.

A final point is that for some there was no craft from which to become detached. As one focus group participant who was at the beginning of her career jokingly put it, 'I have never cut and pasted'. In the future many researchers are likely to be introduced to qualitative computing as a
normal part of their professional socialization. As such they might need explicitly to be taught to recognize the tactile and perceptual constraints which technology imposes. By then, hopefully, computer-based methods will more closely emulate traditional practice. Meanwhile, the teaching of qualitative analysis might also have benefited from a greater explicitness about analytic procedures.

**Notes**

1 Trainers testify that, worryingly often, participants on CAQDAS courses seek to learn how to use qualitative software despite having only a limited background in qualitative methodology (Lewins and di Gregorio, 1996).
2 While such conventions exist in quantitative research, David (1991) claims they are less well developed than they might be.
3 Unlike the Anglo-American common law, one of the glories of the Scottish legal system is that it permits this alternative to the stark bifurcation of 'guilty' or 'not guilty'.