Computing and the 1930s Cinema Culture Project

1. Introduction

The stated objectives of the project are: "*To gather data, in various forms, on cinemagoing and other popular leisure pursuits of the 1930s - habits, tastes, preferences - across Britain, and to document regional, gender and class variations in these*" and "*To devise and put into effect a set of 'ethnohistorical' methodologies appropriate to the collection, analysis and interpretation of oral history and archival source materials*".

These objectives were formulated initially with outputs in mind: it was considered that interview and other data collected during the course of the project might eventually be made available to other researchers as, say, an interactive CD-ROM. For this, the material would need at least to be indexed in some way and/or some analysis of the material included alongside the raw data. The proposal was made in ignorance of what would be involved. At the same time, while I did not formulate the issue in these terms, I realised there would also be a need for data management of some sort, so that the researchers would be able to extract bits of information from large volumes of interview and other data, a task for which an interpretive rather than a quantitative approach would be required.

Even this overstates the degree to which the computer usage/application side of the project had been thought out in advance, however: although one of the people interviewed for the Research Fellowship mentioned NUD-IST (see below) and another was a user of Hypersoft, I had little sense of what these programs could do, beyond the fact that they had been developed specifically for working with interview data. My formulation of things at this stage was that we needed to find, adapt or develop databases for archival and oral history material. When it became apparent early on in the project that approaches to computer specialists were not proving very fruitful (Glasgow University's Computer Advisory Service could only suggests Filemaker), I realised I did not know enough even to ask the right questions.

The first real breakthrough came as I was browsing in the Sociology section of Dillons in Malet Street, where there is a whole bay devoted to social science research methods. I picked up a copy of Miles and Huberman (1994), and saw that it included an Appendix on choosing computer programs for qualitative data analysis. This includes a table (p.316) setting out the characteristics of 22 programs available or developed for the analysis of qualitative data: a couple of these (NUD-IST and The Ethnograph) I had heard of already: Hypersoft was not included in the table. I sent off for demos of those which seemed most appropriate, and pursued inquiries about Hypersoft. The NUD-IST demo arrived first, and raised a new problem: that of getting software running in the first place if you are not very computer literate and have no-one to turn to for advice. Luckily, the documentation accompanying the NUD-IST demo included information on potential sources of advice, including two discussion groups devoted to computer assisted

qualitative data analysis (henceforth CAQDAS), one of them run by the Melbourne-based developers of NUD-IST and the other much closer to home, by the Sociology Department of the University of Surrey.

With inquiries among the computing community at Glasgow University continuing to prove fruitless, this was the exactly what I needed. Simply listening in to these discussion groups proved most informative: I began to get a sense of which of the programs were most in use (at least by this group of people) and for what purposes and in which disciplines. Discussion group subscribers seem to come mainly from Social Psychology and Geography departments, and there is an interest in areas such as discourse analysis among discussants. A discussion I initiated on CAQDAS and oral history interview material produced some interesting, if not entirely pertinent, ideas and generated useful contacts (NUD-IST users in Glasgow and Stirling, for example) and an invitation to a Geographers' Conference. This was very encouraging and made me feel a bit less useless! Interestingly, those who replied to my query were Ethnograph users.

Information about courses and workshops on CAQDAS is also posted through these discussion groups, and it was through one of them that I learned about a short course at the University of Surrey which I attended from 23 to 25 March. This dealt not only with the nuts and bolts of several QDA software packages, but more importantly opened up key questions which I had not even been able fully to formulate before. This, plus the local contacts I have made among CAQDAS users (the latter through personal networks as well as discussion groups), are the most promising developments so far. What follows is what I have learned from these sources, from my reading, and above all from the extremely timely and useful CAQDAS course.

2. QDA Software and its Uses

There are various uses and applications of computer software in qualitative research.

* Tasks such as making notes in the field, writing up or transcribing field notes, editing of notes and preparing interim and final reports can be performed with word processing software.

*Data management tasks such as coding and storage in an organised database, data search and retrieval, data 'linking' (connecting relevant data segments to each other; forming categories, clusters or networks of information), memoing, content analysis, data display and conclusion-drawing and verification are where the packages developed specifically for qualitative data analysis come into their own, though not all of the packages available can do all of these things well, or at all, and all do what they do in rather different ways.

*More advanced tasks such as theory-building and graphic mapping (creating diagrams that depict findings or theories) are offered by several software packages.

Before deciding on whether and which computer applications will be useful, the researcher must first decide which of these tasks are paramount for the project.

Weitzman and Miles (1994) provide a useful categorisation of types of software which can be used in qualitative research (a copy of their summary table of findings is attached):

Text retrievers and text-based managers (of which there are many, e g Metamorph, Sonar Professional, The Text Collector, WordCruncher, ZyINDEX, Folio VIEWS, Tabletop, MAX). These were not developed specifically for QDA but for organising huge text bases that require limited but fast analytic features. Data is recoverable by keywords already in text, and some packages offer additional features such as word counts, concordances, KWIC). Textbase managers offer more organising and sorting features than text retrievers.

Code-and-retrieve packages (e g HyperQual, Kwalitan, QUALPRO, The Ethnograph) are really where QDA begins, and have in fact been developed by qualitative researchers rather than computer programmers. They can divide text into segments and find and display chunks with a given code or combination of codes. Some programs can also retrieve data on a subject where a given keyword does not appear, and support memoing. The fact that these are not commercial packages means that while they have been developed with the qualitative researcher's needs in mind, there may be difficulties with infrastructure, documentation and technical support.

Code-base theory builders (e g AQUAD, ATLAS.ti, HyperRESEARCH, NUD-IST) can do everything the most advanced code-and-retrievers do but also offer conceptualistion and systematic analysis e g through exploring connections between codes. Some of them offer graphical displays of codes and their relationships; and some can test hypotheses, handle offline data such as unformatted text, videos, etc.

Conceptual network builders (eg SemNet, Inspiration) permit organisation of ideas in the form of a network of topics linked by named relations. While none support coding, some offer limited search and retrieval. The strength of these packages lies in their support of data 'linking'. These packages were developed for brainstorming, development of ideas and concepts and mindmapping rather than for QDA.

Before deciding whether to use CAQDAS, or which type of software to choose, the researcher should be clear about the tasks with which software could assist, and then ask:

*How computerate am I?
*Is there are preferred platform (e g Mac, Windows, DOS)
*What is the project timetable?
*Am I working alone or in a team?
*For what audience is the analysis intended?
*What analytic commitments and orientation do I have?
*Am I choosing for one project or for the next few years?
*What kinds of projects and databases will i be involved with?
*What kinds of analysis am I likely to do?

Other important considerations include nature of data sources (single or multiple); whether records are fixed or will be revised; how structured or open the data is; whether

data entries are uniform or diverse; size of the data base; whether analysis is exploratory or confirmatory; whether coding schemes should be firm at the outset or can evolve as the project proceeds; whether multiple or single coding of data segments is required; whether coding is to be iterative or single pass; how important the context of coded data is; how one wishes to have data displayed; whether analysis is purely quantitative or includes numbers as well.

In other words, fitness for purpose is the key. There is no point in choosing the latest software that does everything imaginable if the researcher needs assistance only with basic tasks, nor in choosing complex software that comes with little technical support if she is not highly computerate or if there is a limited timescale to the project, nor in choosing software that supports only DOS when confined to Mac platform.

Within functions, different packages offer varying approaches and degrees of help. For example, for code and retrieve, some offer multiple coding for any segment of data, others do not; some handle overlapping codes, others do not; some allow the researcher to rename and reorganise codes as the project progresses, others do not; some are faster than others; some offer searches for strings of characters, synonyms or wild cards, others do not. Packages differ in the way the results of searches are displayed, and whether records of searches are kept.

There is therefore no single answer to the question: what is the best QDA package? It depends on the project --its objectives, its circumstances, and the people working on it.

3. QDA Software and Cinema Culture in 1930s Britain

First of all, given that it figures in the project's objectives, we are committed to some sort of computer application. CAQDAS would answer our need to store, manage and analyse the kind of data we are collecting (this is not only interviews but also such material as field notes, letters and other writings from respondents, etc). It is not clear at this stage whether CAQDAS will also lend itself to the objective of making data available to outsiders in electronic form. I don't see why it should not, but have to admit to total ignorance on this point. It may be that we will end up making this a separate project if particular sorts of computer expertise are required for it.

As far as CAQDAS is concerned, however, we will certainly need at least to be able to code and retrieve data and write memos. It would be helpful if offline material could be incorporated in analysis, but this is not crucial in the short term. Whatever software is chosen certainly needs to be user-friendly, though I am not sure if it needs to be capable of being learned quickly. It should probably support Mac--which would limit choice considerably, since few packages are available in Mac versions. I do not feel that the project necessarily requires computer-assisted theory-building, though would not object to buying such a package in case of future need, so long as our basic requirements are met by it.

If these considerations limit the range of choice to selection from the code-and-retrievers and theory-based code builders, it is still not easy to make a final selection in the abstract. To aid the decision-making process and offer some idea of how different packages work in practice, the CAQDAS course included demonstrations of The Ethnograph, HyperRESEARCH and NUD-IST, and hand-on tutorials with The Ethnograph and NUD-IST.

From these, and even though I had difficulties with the tutorial because of networking problems, I found I liked The Ethnograph more than I had expected. The earliest QDA software to be developed, it is quite widely used and seems (in my possibly inaccurate observation) to be favoured by researchers who like to work closely with their data--anthropologists, for example. In this it perhaps significant that it was Ethnograph users who took part in the bulletin board discussion on CAQDAS and oral history. It is not as expensive as most of the other packages, but is more limited than some in terms of technical backup. On the other hand, there is obviously a large community of users out there with experience which could be drawn on. Being a DOS package, it fits with my own preference for this platform; but of course by the same token possibly rules itself out for this project.

HyperResearch, though available in Mac version, offers no memoing facility and is therefore ruled out for this project.

NUD-IST is a popular and user-friendly package, and is highly recommended by users I have spoken to. It is the most commercial of all the QDA software, and its developers offer plenty of support, even to the extent of running a discussion group through which users can raise problems and get quick advice from the developers, Tom and Lyn Richards. The Richardses are also very active in promoting the software through workshops, and regularly travel to Europe to run classes in NUD-IST. I found the tutorial easy to follow, although I was less comfortable with its style than I was with The Ethnograph's. The demonstration of the package suggested to me that it was developed for researchers more interested in the codes and categories emerging from their data than in the data itself; an observation perhaps confirmed by the fact that NUD-IST users I have spoken to are predominantly working on sociological or interview-based social survey sorts of projects. NUD-IST does of course offer code-and-retrieve, as well as quite sophisticated memoing, and is available in Mac. An excellent feature as far as the 1930s project is concerned is its capacity to permit ongoing revision and reworking of codes throughout the life of the project, as data builds up and begins to yield different patterns. I was less impressed by the fact that NUD-IST seems to force the researcher to order codes in a hierarchical 'tree' and display them accordingly. This is certainly not how I look at the relationships between categories when I interpret data of various kinds, and I would feel very hamstrung by that approach. The CAQDAS course included an exercise in (manual) coding of data which confirmed this feeling: I--and indeed my fellow-students--discovered that we conceptualised relations between codes as matrices or networks rather than as hierarchies. This, I have subsequently discovered, describes the 'semantic network' approach taken in conceptual network building software like SemNet and Inspiration. I am not sure whether NUD-IST's hierarchical approach to theory

building can be sidestepped if one is not using the package for theory-building purposes but simply restricting oneself to code-and-retrieve, memoing, etc.

There is a package that appears to offer everything NUD-IST does as a theory builder, while supporting more of semantic network approach to conceptualisation and theory building. ATLAS/ti was unfortunately not demonstrated uring the CAQDAS course, because the current DOS version is about to be updated and made available also in Windows. I liked the idea of this package, but have no idea how user-friendly it is. It probably rules itself out for this project, however, by being unavailable in Mac version.

4. Conclusions, Questions and Recommendations

A number of questions emerge concerning the 1930s Cinema Culture project and computing: do we need it; and if so, what for and when?

Given the project's objectives, it is clear that computing and computer applications are integral. It is clear to me that we should seriously consider computer assistance in data management: a project such as this, which will generate large amounts of qualitative data, needs to manage that data--store it and organise it so that the information we require will be at least readily retrievable. Before microcomputers, the techniques of the qualitative researcher involved a great deal of cutting up of transcripts, manual coding, and slotting of pieces of paper into hanging files. Researchers could soon become overwhelmed with paper, and data analysis could at best be one-dimensional--unless one wanted to be drowned in paper, any data segment would rarely be coded more than once, for instance. Once coding was done, it would also be difficult to regain a sense of the wholeness of particular cases or interviews. CAQDAS can assist with data analysis, permitting greater refinement and flexibility; though it does not necessarily save time. Coding will always be timeconsuming, and the researcher using CAQDAS tends to code more deeply and elaborately than is normally possible with manual coding. Software that supports theory-based code building offers additional functions which we can use if we wish to, but do not have to.

It seems clear, then, that the project could benefit from CAQDAS. At what point, though? CAQDAS allows--some would say forces--the researcher to abandon the temporal separation of data collection and data analysis: it is unnecessary to wait till all the data is in before work on it can proceed, especially if one intends to revise and develop interpretations as data unfolds. We could begin CAQDAS, therefore, as soon as we like. On the other hand, there is no pressing reason to rush into it. We should perhaps discuss a suitable timescale in light of the overall progress of the project to date, therefore.

At the same time, it seems to me for various reasons that we could wait a little before making any final decisions as to which software to go for. The choice appears to reduce to ATLAS.ti, The Ethnograph and NUD-IST. There are a number of factors pushing us towards the latter, which unfortunately is probably not the best as far as this project's needs are concerned. Neither of the others is available in Mac version, though Windows

versions of both are apparently about to be released, so it is possible we could accommodate them by upgrading the project's 630 to Power PC. There are reasons why this might not be a good idea, however. It is a pity that choice should be limited for such a silly reason, and confirms my existing feelings about Mac (but that's another story). My conclusion is that if we decide to get into CAQDAS soon, we have little choice but to go with NUD-IST.

The other area of computer application--the question of outputs--remains open to discussion and investigation.

5. CAQAS: Select Bibliography

Fielding, N. an R.M. Lee (1991, 2nd ed 1993). Using Computers in Qualitative Research. London: Sage Publications.

Miles, M. and A. M. Huberman (1994). Qualitative Data Analysis. Thousand Oaks, CA: Sage Publications.

Reid, A.O. (1992). 'Computer management strategies for text data', in B. F Crabtree and W. L. Miller (eds), Doing Qualitative Research. Newbury Park, CA: Sage, 125-45.

Richards, L and T. Richards (1994) 'From filing cabinet to computer' in A. Bryman and R. G. Burgess (eds), Analyzing Qualitative Data. London: Routledge, 146-72.

Richards, T. and L. Richards (1994) 'Using computers in qualitative research', in N.K. Denzin and Y. S. Lincoln (eds), Handbook of Qualitative Research. Thousand Oaks, CA: Sage Publications, 445-62.

Weitzman, E and M. Miles (1994). Computer Programs for Qulittive Data Analysis. Thousand Oaks, Ca: Sage Publications.

Annette Kuhn
April 1995