

# The Digital Revolution in Qualitative Research: Working with Digital Audio Data Through Atlas.Ti

by Will Gibson, Peter Callery, Malcolm Campbell, Andy Hall and Dave Richards  
The University of London, Institute of Education; University of Manchester; University of Manchester; University of Manchester; University of Manchester; University of York

*Sociological Research Online*, Volume 10, Issue 1,  
< <http://www.socresonline.org.uk/10/1/gibson.html> >  
doi: 10.5153/sro.1044

Received: 9 Aug 2004 Accepted: 17 Feb 2005 Published: 31 Mar 2005

---

## Abstract

Modern versions of Computer Assisted Qualitative Data Analysis Software (CAQDAS) are enabling the analysis of audio sound files instead of relying solely on text-based analysis. Along with other developments in computer technologies such as the proliferation of digital recording devices and the potential for using streamed media in online academic publication, this innovation is increasing the possibilities of systematically using media-rich, naturalistic data in place of transcribed 'de-naturalised' forms. This paper reports on a project assessing online learning materials that used Atlas.ti software to analyse sound files, and it describes the problems faced in gathering, analysing and using this data for report writing. It concludes that there are still serious barriers to the full and effective integration of audio data into qualitative research: the absence of 'industry standard' recording technology, the underdevelopment of audio interfaces in Atlas.ti (as a key CAQDAS package), and the conventional approach to data use in many online publication formats all place serious restrictions on the integrated use of this data. Nonetheless, it is argued here that there are clear benefits in pushing for resolutions to these problems as the use of this naturalistic data through digital formats may help qualitative researchers to overcome some long-standing methodological issues: in particular, the ability to overcome the reliance on data transcription rather than 'natural' data, and the possibility of implementing research reports that facilitate a more transparent use of 'reusable' data, are both real possibilities when using these digital technologies, which could substantially change the shape of qualitative research practice.

---

**Keywords:** *CAQDAS, Recording Technology, Online Publication*

---

## Introduction

**1.1** The aim of this paper is to discuss the use of digital data within qualitative research practice. Since the early Eighties, the revolutionary potential of computing technology in qualitative research has been a hot topic. In the most part, the debate has centred around the use of Computer Assisted Qualitative Data Analysis Software (CAQDAS) and its impact on the analysis process (Becker 1984; Conrad and Schulamit 1984; Drass 1980; Gerson 1984; Kirk 1981; Pfaffenberger 1988; Seidel and Clark 1984; Sproull and Sproull 1982; Tesch 1990). The story of this debate can be summarised, albeit somewhat simplistically, as a move from initial excitement and experimentation, peppered with some scepticism, towards a more reflexive caution over the impact of CAQDAS on the research process. Notably, there has been concern that the design of the programs themselves, and the ways in which they afford the development of coding structures and theory building approaches, may serve to shape the outcomes of the analytic process (Buston 1997; Crowley et al. 2002; Fielding 1995; Lee and Fielding 1996; Lee and Lea 2000; Raymond and Lea 2000; Richards and Richards 1995; Seidel and Clark 1984). This concern has centred around a perceived convergence on 'code and retrieve' software packages which aim to facilitate 'grounded theory' analytic approaches (Coffey et al. 1996; Lee and Fielding 1996). Another, much less prominent, aspect of the debate in this area has concerned the data collection process itself in terms of using email and Internet-chat facilities to conduct interviews and focus groups (Chen and Hinton 1999; Gaiser 1997; Illingworth 2001; O'Connor and Madge 2001; Saris 1991; Selwyn and Robson 1998; Williams and Robson 2004).

**1.2** While such issues continue to be discussed, comparatively little is being made of the potentials of either CAQDAS or online publication formats for the use and representation of audio/visual digital data formats in qualitative research. Currently, the dominant research convention within qualitative research is to

obtain audio and sometimes audiovisual data from interviews, focus groups, observations or participant observations, and transcribe that data for the purposes of both analysis with CAQDAS and subsequent publication in print or online journals. Technological developments now mean that many CAQDAS programs can be used to analyse data sources other than text. Similarly, comparatively recent developments in 'streaming' technologies mean that the use of this data within online publications is an increasingly realistic prospect. 'Streaming media' is a term that describes the delivery of compressed multimedia formats over the Internet. As its name suggests, streaming technology allows data to be transferred in a stream of packets that are interpreted as they arrive on a client machine rather than being reassembled into a file format. Without streaming, the entire media would have to be downloaded in one big package before it could be used; the process can be thought of as 'just in time delivery', with the client PC receiving the next packet of media data as it is needed. Streaming usually requires special servers that deliver the streams to users using protocols such as IP multicast, UDP, and RTSP, whereas standard web servers and the HTTP protocol can be used for downloading. According to figures produced by AccuStream media research unit (<<http://www.accustreamresearch.com/products/streamingmedia2003.html>>), total video streams served/viewed in 2003 rose 104% to 7.87 billion with broadband streams making up 78% of this figure; video streams served during January and February by the top ten Internet sites rose from 292 million in 2003 to 523 million in 2004. Usage of this level demonstrates that streaming technology is fully embedded into standard web architecture and usage patterns and therefore should not be viewed as a marginal medium for academic publication.

**1.3** It is important to emphasise that neither of these features should be regarded as peripheral or 'gimmicky' developments, as the provision of naturalistic (as opposed to transcribed text) that can be displayed within online published literature through streaming media may allow researchers to respond to two important and persistent problems within qualitative research. Firstly, the written representation of speech through transcription is a poor substitute for the data to which it pertains; even complex representation structures such as those used in Conversation Analysis, which are designed to explicate context within the transcriptions, are not adequate for picking up all the relevant inflections that are placed on language and interaction (Ashmore and Reed 2000). The result is that it can be extremely hard to communicate the complexity of meaning embedded within a given interactional event through transcription. This is particularly troubling in qualitative research where the commitment to 'naturalistic' enquiry is very often a central rationale for its employment. Naturalism can be summarized as a view that the social world should be studied in its 'natural state' (Hammersley and Atkinson 1983) by using methods that are unobtrusive (i.e. do not 'transform' the world through the process of their application). That the methods of transcription themselves serve to transform this data has therefore been a longstanding concern amongst qualitative researchers (Mishler 1991). However, by using 'pure' audio/audiovisual data as embedded digital files, researchers are able to bypass the necessity of relying on inadequate representation systems, and can use the data itself to display the phenomena under study.

**1.4** Secondly, because of the limitations of space usually afforded within publication formats, one of the emergent conventions of publication within qualitative research has been to provide a few key examples of particular phenomena that represent a wider data set. A common criticism of this approach has been that it is easy for researchers to misrepresent data through the selective use of quotations (Becker 1998). As researchers do not typically present all their findings but merely a selection of them it is hard for quality controls to be implemented to guard against misuse of such data by, for example, selecting examples to make a case or build a theory that is not necessarily represented in a data set. However, because streamed media makes the possibility of including much larger amounts of data within text, these new technologies may provide ways to change these publication conventions; the process of constructing qualitative research reports could become much more 'accountable' to professional standards of data use (Brent 1984; Fielding 2003; Kelle and Laurie 1995; Muhr 2000) by providing multiple examples of data or by encouraging the 're-use' of data - currently an unusual practice (Albrecht 1985; Gladney et al. 2003), which brings with it important ethical and legal issues (Akeroyd 1998; Parry and Mauthner 2004).

**1.5** The reusability of data has been a key issue for authors interested in the use of streaming media within hypermedia authoring (Coffey et al. 1996; Dicks and Manson 1998). Hypermedia can be thought of as computer mediated texts (including written as well as audiovisual media) that can be linked to other such texts. The medium has been seen as offering significant scope for creating alternative approaches to data analysis and representation for qualitative researchers that go beyond simple code and retrieve procedures, and lead to the creation of flexible narrative forms. In these alternative approaches to data use authors do not construct lineal narratives from their coding analysis, but develop more nuanced presentational forms with multiple pathways that enable readers to negotiate, interpret and create narrative structures in a multiplicity of ways (Mason and Dicks 2001). However, as Mason and Dicks (2004) have argued, the development of alternative approaches to analysis and representation through such media has been problematic as such forms often lack the persuasive power of more traditional narratives. Furthermore, limitations in available software that enables sufficiently flexible hypermedia authoring, and the need for

interdisciplinary collaborative research that gives researchers the necessary variety of skills for such projects, both limit, for the time being at least, the practical application of the approach (Mason and Dicks 2001).

**1.6** The point we wish to emphasise for the moment is that digital technologies are *potentially* revolutionary for qualitative approaches as they may play a key role in fundamentally altering research practice. Given this, the consideration of the procedures of using digital audio/video data can be seen to be of great importance. This paper aims to contribute to such a discussion by reflecting on the collection, analysis and use of audio data in a qualitative research project using code and retrieve procedures to construct lineal narratives.

**1.7** The data for the discussion that follows comes from an analysis of interaction within online learning environments. A part of the research involved obtaining audio recordings of 35 face-to-face seminars with between six and 12 people (including one tutor per group). Because of the necessity of achieving a high quality recording in which individual participants could be identified we decided to use Mini-Disk recorders with good quality stereo microphones to capture the data. The original idea was that each recording would be transcribed for subsequent analysis in Atlas.ti qualitative software. However, as the research progressed it became increasingly evident that providing transcriptions of the group seminars was an extremely time-consuming process. Because of this we decided to experiment with using the functionality of Atlas.ti to code audio data and to analyse these interactive discussions in their audio form. In this way, the transcription process could be much more selective, focusing simply on the data that was being used for research reports. As this analysis continued, however, the paradox of providing transcriptions purely for the purposes of representation became all too apparent. With this in mind, further experimentation was made in attempts to use the audio data for representation purposes rather than merely for analysis.

**1.8** The discussion that follows is, in part, a report on the problems and concerns that emerged in the process of this experimentation, but it is also a summary of the types of issues that researchers who conduct audio data analysis are likely to face. The overall purposes of the paper are threefold: firstly, to describe the basic technical requirements for undertaking this type of recording in terms of both hardware and software; secondly, to compare the process of analysing audio as opposed to text-based data using Atlas.ti; and thirdly, to illustrate the challenges faced in using this data as opposed to transcribed data for subsequent writing and publication purposes. The argument throughout is that both in terms of analysis and publication there are undoubtedly tangible benefits in using naturalistic data, but that the lack of standardisation of technology, the underdevelopment of CAQDAS software such as Atlas.ti as interfaces for audio analysis, and the current structure of academic publication all mean that, for the time being at least, a thorough integration of naturalised data continues to be problematic.

## **Recording Equipment and the Transferal of Digital Data for Analysis**

**2.1** The decision to use Mini-Disks as a format for data collection in this project was one based on the need for high quality recordings and not on their appropriateness for transferring digital data to computers for subsequent analysis. However, having decided to attempt to transfer the data for computer analysis, it became all too evident that not all digital collecting devices are well suited to this process. This section of the discussion reports on the variety of recording technology available for obtaining digital data for computer analysis. It is not our intention to provide a thorough review of the wide range of hardware options, but rather, to draw attention to the kinds of issues faced in selecting one or other device.

**2.2** Analogue and Digital Recording Devices - Many researchers still rely on analogue tape recording devices when conducting interviews, focus groups and suchlike. While in principle it may be possible to transfer this data to a PC and convert it to digital data, this is, in comparison to other digital options, extremely time-consuming and far from straightforward (not least because it involves real-time data transferal). In addition, the nature of the technology itself can often be inferior to digital devices in terms of the quality of recording that is produced (Stockdale 2003). Thus, in terms of both the compatibility of devices and the quality of end product, digital recording is a preferable approach (Stockdale 2003). There is currently no industry standard digital audio data collection device: Mini-Disk recorders; MP3 recorders; and Dictaphone voice recorders are all feasible options, but, as we suggested above, each has its own pros and cons for the data collection process. What follows is a list of important considerations when selecting one or other device:

**2.3 Microphone** - In-built microphones are often of poor quality, and more sensitive external microphones may be required. In such instances, the recording device needs to have an external microphone port.

**2.4 Recording Time** - Many devices enable users to select levels of recording quality. For effective transcription the higher quality settings are likely to be necessary. In our study the slowest recording speed (and highest quality) produced recordings in which it was easy to identify individual voices, even within large

groups of speakers. The duration of recordings will vary considerably between the types of devices, from around 60 minutes per (changeable) disk on Mini-Disk players to hundreds of hours on high-end MP3 recorders. Devices that record to a 'hard disk' such as voice recorders or MP3 players may have memory cards that could expand the recording time.

**2.5 Method of Data Transfer** - While in principle digital files can be transferred easily between recording devices and computers, some digital recording equipment is designed only to transfer files (typically music files) that have been purchased from particular websites. The result is that sound files recorded on the devices themselves may not be very easy to transfer with the digital connections provided such as USB or Firewire. For example, while some Mini-Disk recorders do come with USB connections to music filing software, these connections only support the transfer of files purchased through particular channels. Some Mini-Disk players come with optical outputs that can be used on some computers; otherwise the files need to be played into sound recording software (such as Windows Media Recorder) through a high quality sound card (which are not provided as standard on all computers). Similar problems may occur with some MP3 recorders; Dictaphone sound recorders are much more straightforward in this respect as they are specifically designed to facilitate the transfer of recorded sound onto computer software. However, many of these have a very low recording time. When selecting recording equipment then, users need to make sure that the data can be moved from the device in a straightforward way. Within our experimentation of using Mini-Disk technology, we found a variety of ways of getting the data off - connecting the unit to an external CD burner and moving the data to a computer from the CD; hiring professional editing facilities to move the data from professional Mini-Disk equipment - which were either expensive or time-consuming and, in all cases, circuitous.

**2.6 File Format** - Digital sound files come in a number of formats, not all of which are necessarily supported by the CAQDAS being used. For example, Atlas.ti version 5 only supports WAV (\*.WAV) and NEXT/SUN (\*.AU and \*.SND) audio files. The type of format that the recording devices use will vary and the format may change depending on the method of data transfer from the unit to the PC. When using normal digital transfer formats such as USB or Firewire, the formats will likely be the default formats of the unit. However, if an alternative transfer method is used, such as playing the file into a sound card or burning the file onto CD, the format may well be different. Whatever the transfer method, the point to be aware of is that the purpose is to gain a format that is compatible with the software being used. In our case, the Mini-Disk players only produced a compatible format when the files were played into a sound card. The default format of files when transferred through USB (which was not possible anyway for the reasons mentioned above) would not have been compatible. When burnt directly onto a CD from the unit, the files were converted to music file formats (which were not compatible with the software used). In contrast, digital Dictaphone recording device software such as the Sony ICD-ST10 provide software that enable the conversion of files to a number of formats, including WAV, which can subsequently be altered to a variety of formats through other software.

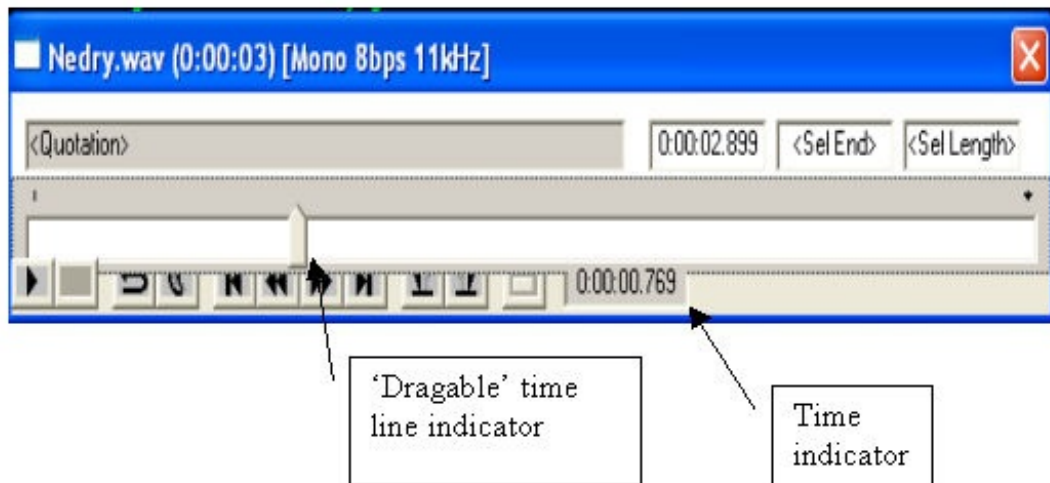
**2.7** In summary, the general aim is to select a device that can create sufficiently high quality digital files of an appropriate length, which can be transferred to a computer and which are, or can be made to be, compatible with the software being used. In our research we found that while the quality and duration of recordings on Mini-Disk recorders was high, there were some serious although not insurmountable problems with moving the data from this format to a PC. In the absence of an industry standard device for obtaining and transferring recordings, researchers need to be careful when choosing recording hardware and software.

## **Data Analysis**

**3.1** Once compatible digital computer based files have been produced, researchers can begin analysis using the CAQDAS format selected. In this section we provide some reflections on our experiences of using Atlas.ti to analyze our data set. The point of the following discussion is not simply to evaluate the current benefits of the Atlas.ti interface, but to highlight important features of the process of coding audio data that may be relevant to the development of future interfaces. While there are some less well-known programs that provide specialized frameworks for audio analysis (e.g. C-I-SAID), Atlas.ti has been, and continues to be, one of the most well-used interfaces (Weitzman and Miles 1995). As such, reflection on the process of trying to use this software for coding audio data is useful in explicating some of the issues that researchers face when undertaking audio analysis.

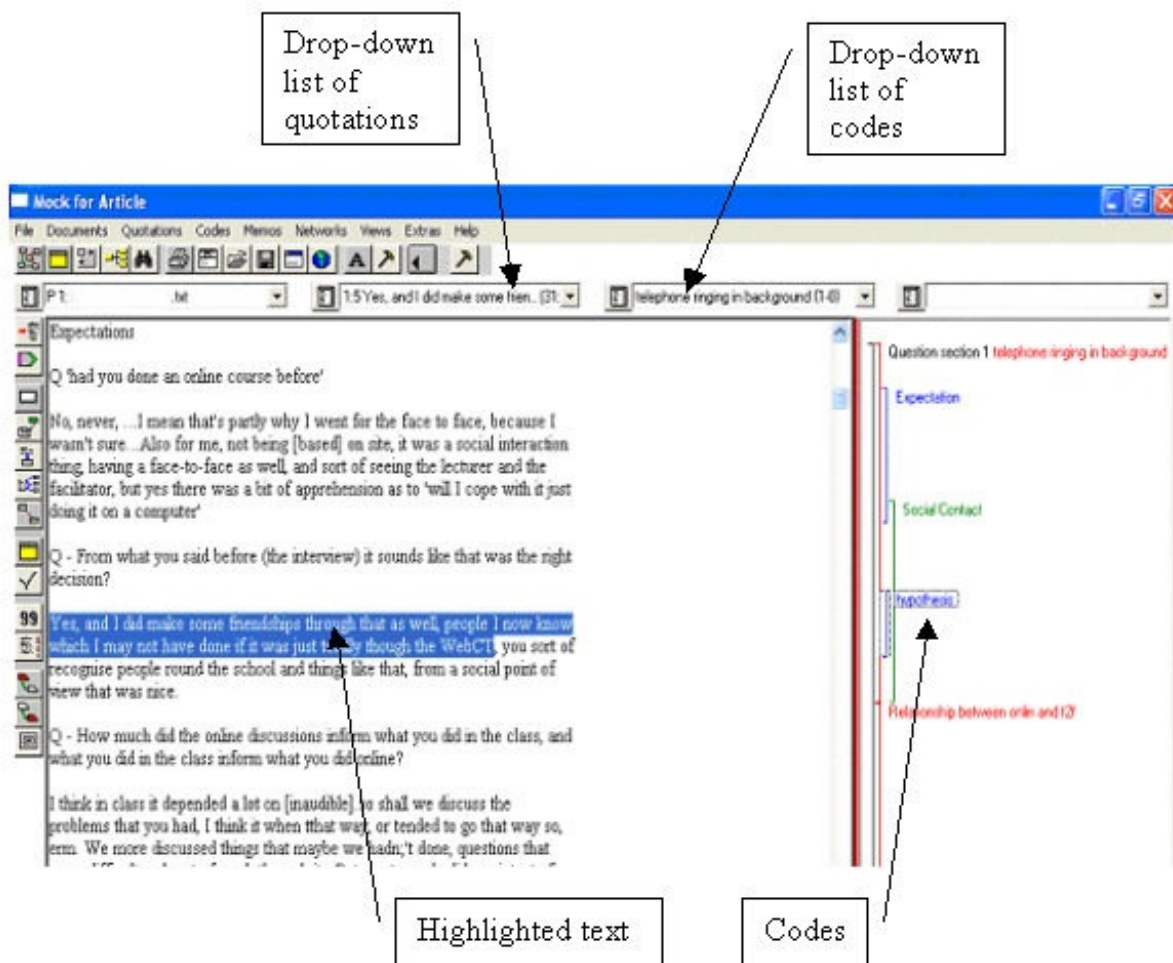
**3.2** Atlas.ti works by linking the files being worked on with what it calls the 'hermeneutic unit' (in other words, the coding structure being produced). Once the files have been linked (a simple process of putting the file in an appropriate folder and clicking on an 'assign' function) researchers are able to start analysis. In Atlas.ti, sound files are played through an editing interface that uses a pointer that scrolls across the screen as the files are being played, which, along with a time counter, creates a visual time line of the file (see figure 1). This interface is used to navigate through the sound file and to select particular sections of

the file for coding. The selection is made by specifying start and end points within the file and then applying a code to the selected section. In Atlas.ti, the code defining and application process is the same for audio as it is for text files involving drop-down menus to create codes and assign them to the data.



**Figure 1.** Screenshot of audio interface

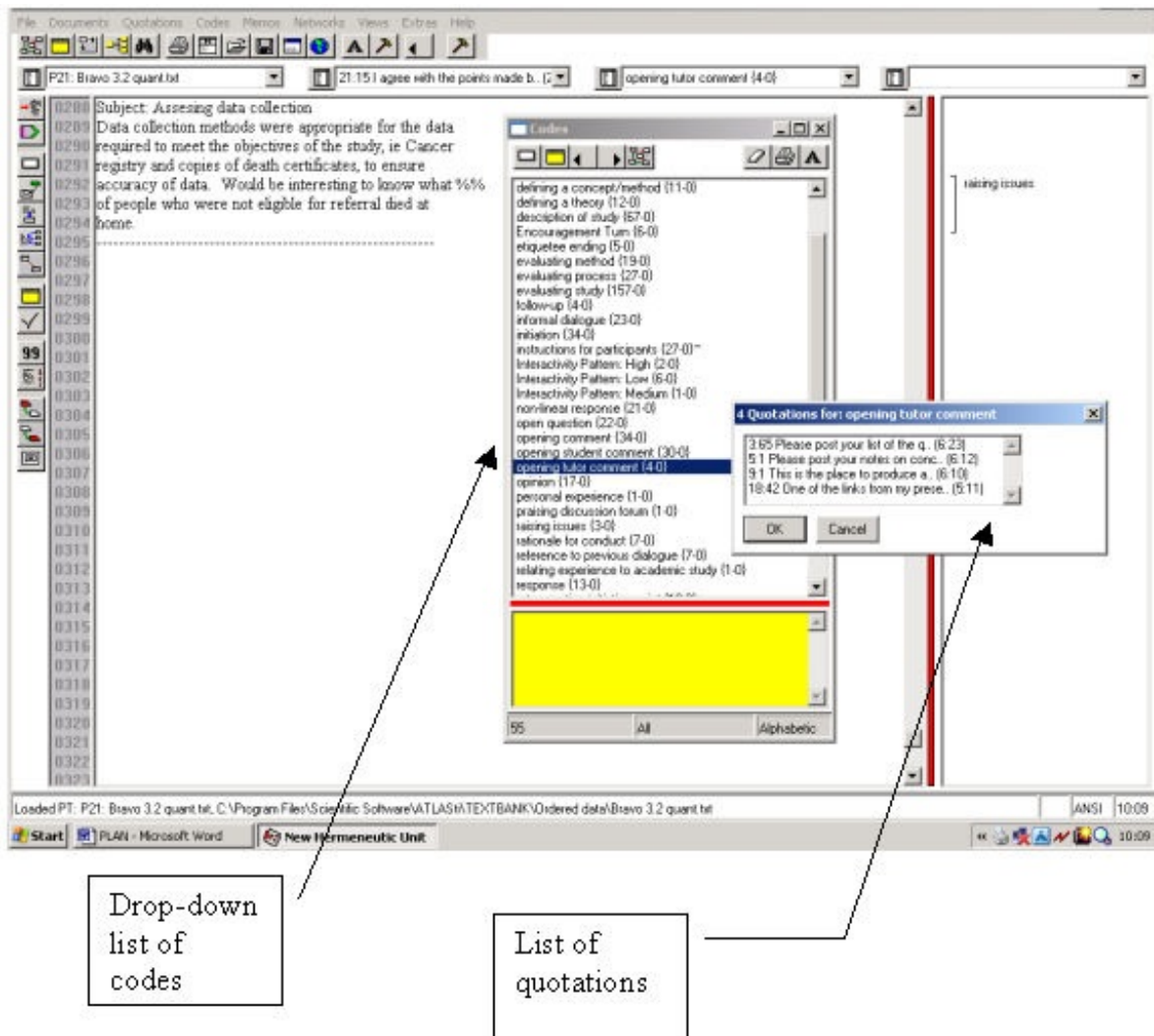
**3.3** However, the system of viewing those assigned codes in relation to audio and visual data in Atlas.ti is very different. With text analysis, codes are displayed in an individual section of the interface on the right hand side of the screen with coloured lines displaying the text sections that have been coded (similar to handwritten text selection lines in paper-based analysis). When a user clicks on the codes on the right hand side of the interface, the relevant text to which it pertains is highlighted on the left (see figure 2).



**Figure 2.** Screenshot of text coding interface

**3.4** In addition, the interface uses a drop-down list of the quotations that the user has defined; when the

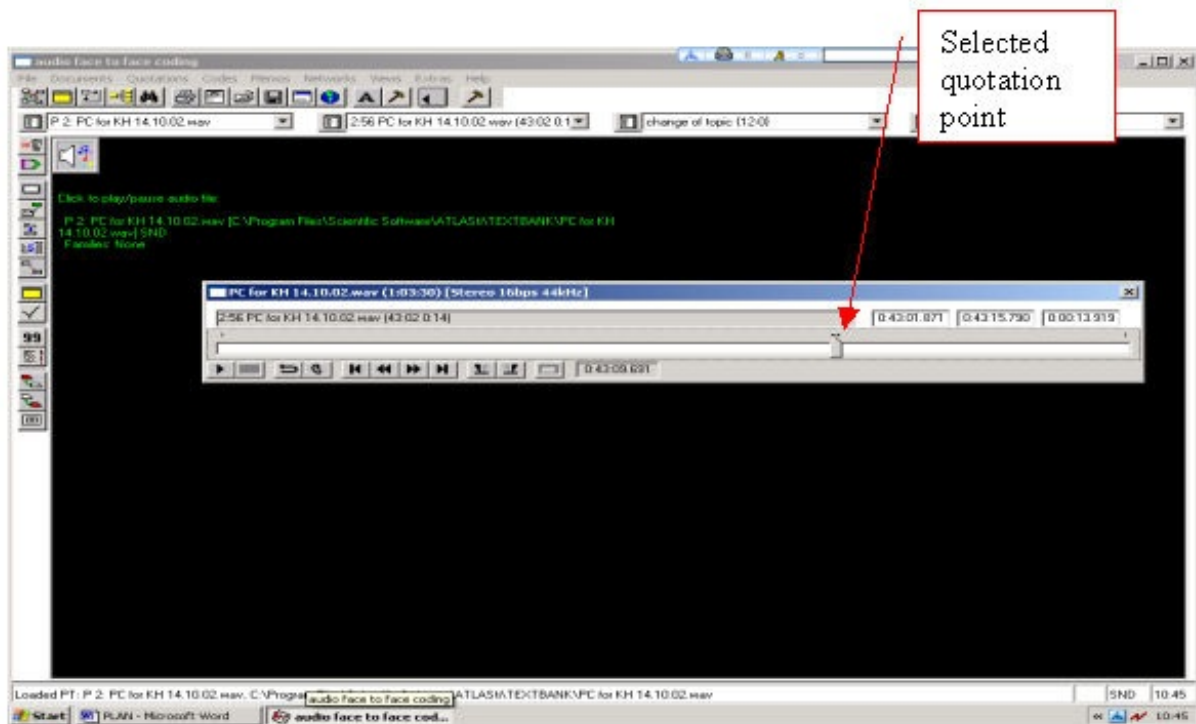
user clicks on these quotations the software takes the user to the relevant quotation and highlights it. In a separate window, Atlas.ti provides a drop-down list of the defined codes (see figure 3). When a code is selected that has been applied more than once a window is displayed that contains all the quotations that have been assigned to that code. When the user clicks on these the program displays the quote as highlighted text. Where only one quotation has been assigned to a code, Atlas.ti displays that quotation as soon as the code label is selected.



**Figure 3.** Screenshot of codes and quotations in text coding

**3.5** These methods of using 'hypertext' links to navigate between codes and quotations is a key benefit of computer based qualitative analysis (Dey 1995), and has been referred to as 'arguably the best of all the code-and-retrieve programs, allowing simultaneous viewing of all relevant information' (Weitzman and Miles 1995). Our experience of using Atlas.ti in this and other projects has borne out the value of this interface structure; the ability to see the codes that have been applied to text while viewing the data itself is a particularly strong feature of this aspect of the software design as it enables users to view the development of a coding structure in context - i.e. with reference to both the other codes that have been applied to the text and the text itself.

**3.6** In contrast, audio coding in Atlas.ti has a much more basic interface. The two drop-down code and quotation menus displayed in figure 2 are still present, but the system of visual code representation through colored lines is not. Thus, when a code is selected through either the drop-down quotation or code menus the quotation is only displayed with its start and end points (see figure 4); unlike text-based coding and its use of colored lines, in audio work only one code can be viewed at any given time. This is a significant departure from the rich context of code viewing in text-based work where the wider context of code construction can be viewed while coding work is going on. This is particularly problematic because, unlike text-based work, there is no visual reference for the audio data, which can make it much more awkward to handle.



**Figure 4.** Screenshot of quotation selection in audio coding

**3.7** In our research we found that this basic difference in representational structure had a significant impact on the approach to conducting the analysis. It was difficult to visualize the overlap of codes within the interface itself so that researchers became reliant on memory and note taking to keep track of the coding structure that was developing. It was not easy to tell, for example, whether or not a given piece of sound had been previously coded in another way or was part of another code section. Unlike text-based work, the audio interface did not offer researchers a way of conceptualizing the coding structure in relation to this wider context. One of the key features of Atlas.ti is that it is intended to facilitate a 'grounded theory' approach to data analysis in which theoretical statements can be elaborated from the data analysis (Glaser and Strauss 1967); the basic visual representational format of text coding outlined above is a key part of the process. In our use of the data we found that its absence made the analysis process much more problematic. For example, one part of our coding strategy involved categorizing the answers that students gave to questions within the seminars. Answer sections had a six stage coding process in which a given answer section could have up to six different codes, all pertaining to its status as an answer (for example, whether it directly answered the question; introduced new issues; referenced course material; oriented to other students, and so on. See table 1 for more details). Individual sections of a sound file, then, were often coded in multiple ways, but the interface made it extremely difficult to visualize this overlap. A key of the limitations of the interface was that little visual compensation had been provided for the move to a non-visual form of coding, and being able to 'see' an emerging coding structure in relation to data segments is, we suggest, fundamental to creating an intuitive software package.

**Table 1.** Coding strategy for answer sections within online seminars

Level One	Dealing directly with issues raised by tutor	Skirting the main issues raised	
Level Two	Introducing new issue or perspectives	Orientation to arguments previously made in the seminar	
Level Three	Reference to course material (readings/lecture notes)	Non-referenced response	Critique of materials
Level Four	Orientating to other students	Orientating to tutor	
Level Five	Raising new questions	Closed ending: no questions	
Level Six	Questions directed to particular student	Questions directed to whole group	Questions directed to tutor

**3.8** However, the potential benefits of this system of working were, equally, very clear: the richness of the data as 'audio text' meant that features of the data that may otherwise have been lost in transcription - such as tone of voice, intonation, background noises and so forth - were all present, bringing a much clearer conception of context than can be provided through text transcription. We therefore found that it was easy to identify, and to analytically take account of, more subtle interactional features such as rhetorical questions; interruptions within the talk; shifts in emphasis and levels of interest; background talk; or physicality within interaction. For example, audio extract 1 is a piece of dialogue in which a tutor uses humour and bodily enactment (walking around) to illustrate the problems that can arise when supervising student nurses. Because the data analysis was not dependent on transcription, these minutiae of social action were still present within the data being coded and therefore could be included within the analysis.

[Click here for audio extract 1](#)

**3.9** The problem of the limitations of the Atlas.ti interface are certainly not insurmountable and it is easy to envisage that a more 'visually rich' platform may be developed that may facilitate a more user-friendly coding approach. However, it is interesting to note that the recently released updated version of Atlas.ti (version 5) has made no significant changes to this aspect of the program. For the time being then, the limitations of coding audio data within the Atlas.ti interface remain.

### **Data Usage/Representation and Publication**

**4.1** As we saw in the introduction of this paper, one of the key uses of CAQDAS software has been to aid the ordering of chunks of data into distinct themes or codes. In Atlas.ti it is easy either to print out or to save electronically a list of all the quotations that have been assigned to a given category. In terms of the practicalities of working with qualitative data of this kind, the ability to cut and paste quotations is, in comparison to paper-based use, convenient and time saving (Buston 1997; Creswell 1998). Users are able to paste these coded quotations into files in other formats such as Microsoft Word either for further analytic work or for writing. When working with audio data, however, the usage strategies are quite different. We applied two different methods for using the data; firstly, once selected, quotations were transcribed to produce a list of text transcripts of the relevant audio files. Such selective transcription of codes (i.e. only coding the data that is to be used rather than the entire data set) significantly reduced the overall labour of transcription. Furthermore, the ability to play, pause and repeat the audio clips within the Atlas.ti interface makes this transcription process reasonably straightforward. To the extent that reducing transcription time was a significant part of the rationale for coding the data in an audio format, our experimentation with audio



coding can be viewed as being successful.

**4.2** However, further experimentation was made in attempting to use the data in a naturalistic format within the context of the writing process. Because Atlas.ti is merely a code-mapping interface, separate editing software (in this case 'Audacity') was needed to cut the files into the chunks designated in the coding process. Though time-consuming, the process was reasonably straightforward, involving splicing the text at the start and stop points designated within the coding framework. These spliced sections were saved as individual file sections and either inserted into or linked to text documents to be played in place of, or alongside, transcriptions. Again, although the procedure for getting to this stage of embedded text was laborious, the benefits were evident: the sound files brought a new dimension of representation that made the 'experience' and understanding of data much richer. Figure 5 is a transcription and accompanying audio clip of a short piece of talk from one of the seminars, which displays some of the key interactional features of the talk. The clip, we suggest, offers a much more intuitive understanding of the interaction than the transcription does, and enables the reader to get a clearer understanding of the event being depicted.

1	Student 1	[...] biggest cons I think is erm . you go to most A and E's now and they	
2		say, er, 'you'll be seen within twenty minu[tes] . and its just someone...	
3	(Unidentified)		[mmm]
4	Student 1	... triaging [you know] that's all it is. And you then you go in a four hour..	
5		[murmurs of agreement]	
6		cue don't you so. If you look [under these]. Devices, you know [its]	
7	(Student 2)		[Eight hours]
8	(Tutor)		[This is]
9		what [...]	

**KEY**

- [ ]** Text in square brackets depicts overlapping talk
- .** A single dot indicates a brief pause
- ...** Three dots indicate continuous text
- [...]** Three dots in square brackets show the start and stop points of the transcripts

**Figure 5.** Transcription extract

[Click here for audio extract 2](#)

**4.3** Further, as we suggested in the introduction of this discussion, audio files such as this provide a convenient way of providing multiple examples of phenomena, which clearly demonstrate the continuity of phenomena under study but without taking up significant text space.

[Click here for audio extract 3](#)

[Click here for audio extract 4](#)

[Click here for audio extract 5](#)

[Click here for audio extract 6](#)

**4.4** This format of data presentation gives authors a useful way of presenting data in a way that preserves the richness of interaction, without the need to provide lengthy transcripts. Moreover, the possibility of including links to lengthy data sets that contextualize specific events or phenomena makes opening analysis to detailed scrutiny a real possibility. Instead of merely including extracts, then, streaming media facilitates the integration of large data sets into the reporting process.

**4.5** There were, however, some key problems with this form of data representation. When using text transcriptions researchers can 'edit' the text in order, for example, to preserve anonymity (by erasing names) or to cut elongated sections of talk that may be distracting or irrelevant to the particular analytic point being made. While audio clips can also be edited to cut or 'bleep' out particular sections of talk, we found that the speed and 'messiness' of natural talk meant that such editing could make the audio clips difficult to follow. This was particularly problematic when more than one person was talking, as the 'edit' would cut all the sound at the selected point rather than just the problematic part. A further and perhaps

more profound problem, however, is that, as yet, the actual use of such data for the purposes of publication remains limited. In spite of the possibilities afforded by technological innovations, most publication formats remain fairly conventional, even in online journals, and do not support the use of these kinds of integrated technologies. Where the use of such audio visual data is supported as an integrated feature of publication (Sociological Research Online is unusual in that it does support such use), authors rarely actually use the technologies in the ways described, and instead tend to rely on conventional transcription (Sociological Research Online contains very few examples of authors employing audiovisual data within the articles - see Lomax and Casey (1998) for a rare example). Thus, having worked with audio data in an attempt to preserve the 'naturalness' of the events being examined, our choice of journals for using such data was extremely limited. Given that the technologies that facilitate the use of this type of streamed audio and video media are comparatively recent (emerging only in the last nine years with significant improvements over the last four years - see the RealNetworks website: <<http://www.realnetworks.com/>>), it is perhaps not surprising that there has yet to be a large move towards its use within published literature. In the near future the conventions of online publication and the uses of media rich data may be a real possibility, but for the time being current approaches to publication represent a significant barrier to the integrated use of naturalistic digital data.

## Conclusions

**5.1** This paper has suggested that there are key benefits to using digital audio/video data as an integrated feature of qualitative analysis and publication. In particular, such usage may enable researchers to produce significantly more naturalistic research reports that draw on the data in its 'natural form' rather than relying on transcribed versions. Furthermore, such data can be used in a more rigorous manner and help qualitative researchers to achieve 'accountable' research processes in which the uses of data are transparent and where the re-use of data becomes a real possibility. In our research project, however, we found that in spite of the significant technological developments in the field of digital and streaming web technologies, the integrated use of this data is far from straightforward: the lack of 'industry standard' recording technology, the apparent underdevelopment of Atlas.ti (as a prominent CAQDAS interface), and the conventionality of online publication formats make the systematic use of digital audio data extremely problematic. It seems, then, that significant research and development is still required before qualitative researchers can make the most of this digital age.

---

## Acknowledgements

This paper is part of a project funded by the General Nursing Council (GNC Trust). We are extremely grateful for their support throughout the project.

---

## References

- AKERROYD, A. 1998. "Personal Information and Qualitative Research Data: Some Practical and Ethical Problems arising from Data Protection Legislation." Pp. 89-106 in *Using Computers in Qualitative research*, edited by R Lee. London: Sage.
- ALBRECHT, G. L. 1985. "Videotaping Safaris: Entering the Field with a Camera." *Qualitative Sociology* 8: 325-344.
- ASHMORE, M, and D Reed. 2000. "Innocence and Nostalgia in Conversation Analysis: The Dynamic Relation of Tape and Transcript." *Forum: Qualitative Social Research* 1.
- BECKER, H. 1984. "Fieldwork with the Computer." *Qualitative Sociology* 7: 16-33.
- BECKER, H. 1998. *Tricks of the Trade: How to Think About your Research While You're Doing It* . Chicago: University of Chicago Press.
- BRENT, E. 1984. "Qualitative Computing: Approaches and Issues." *Qualitative Sociology* 7: 34-60.
- BUSTON, K. 1997. "NUD\*IST in Action: Its Use and Its Usefulness in a Study of Chronic Illness in Young People." *Sociological Research Online* 2.
- CHEN, P, and S. M. Hinton. 1999. "Realtime Interviewing Using the World Wide Web." *Sociological Research Online* 4.
- COFFEY, A., B Holbrook, and P Atkinson. 1996. "Qualitative Data Analysis: Technologies and Representations." *Sociological Research Online* 1.

- CONRAD, P, and R Schulamit. 1984. "Computers and Qualitative Data: An Introduction." *Qualitative Sociology* 7: 3-15.
- CRESWELL, J, W. 1998. *Qualitative Inquiry and Research Design: Choosing Among Five Traditions* : Sage.
- CROWLY, C, R Harre, and C Tagg. 2002. "Qualitative research and Computing: Methodological Issues and Practices in Using QSR NVivo and NUD\*IST." *Social Research Methodology* 5.
- DEY, L. 1995. "Reducing Fragmentation in Qualitative Research." in *Computer-Aided Qualitative Data Analysis: Theory, Method and Practice*, edited by U Kelle. London: Sage.
- DICKS, B, and B Manson. 1998. "Hypermedia and Ethnography: Reflections on the Construction of a Research Approach." *Sociological Research Online* 3.
- DRASS, K. 1980. "The Analysis of Qualitative Data: A Computer Program." *Urban Life* 53: 332-352.
- FIELDING, N. 1995. "Choosing the Right Software Package." ESRC Data Archive Bulletin.
- FIELDING, N. 2003. "Qualitative Research and E-Social Science: appraising the potential." A Report to the ESRC.
- GAISER, T. 1997. "Conducting Online Focus Groups: A Methodological Discussion." *Social Science Computer Review* 15: 135-144.
- GERSON, E. 1984. "Qualitative Research and the Computer." *Qualitative Sociology* 7: 61-67.
- GLADNEY, A, C Ayars, W Taylor, P Liehr, and J Meininger. 2003. "Consistency of Findings Produced by Two Multidisciplinary Research Teams." *Sociology* 37: 297-313.
- GLASER, B., and A Strauss. 1967. *The Discovery of Grounded Theory : Strategies for Qualitative Research*. Chicago: Aldine.
- HAMMERSLEY, M, and P Atkinson. 1983. *Ethnography: Principles in Practice*. London: Routledge.
- ILLINGWORTH, N. 2001. "The Internet Matters: Exploring the Use of the Internet as a Research Tool." *Sociological Research Online* 6.
- KELLE, U, and H Laurie. 1995. "General Methodological Issues." Pp. 19-28 in *Computer-Aided Qualitative Data Analysis; theory methods, and practice*, edited by U Kelle. London: Sage.
- KIRK, R. 1981. "Microcomputers in Anthropological research." *Sociological Methods and Research* 9: 473-492.
- LEE, R, and N Fielding. 1996. "Qualitative Data Analysis: Representations of Technology: A Comment on Coffey, Holbrook and Atkinson." *Sociological Research Online* 1.
- LEE, R, M, and E Lea. 2000. "Computer software and qualitative analysis: trends, issues and resources." *Social Research Methodology* 3: 231-243.
- LOMAX, H and N Casey. "Recording Social Life: Reflexivity and Video Methodology." *Sociological Research Online* vol. 3, no. 2 <http://www.socresonline.org.uk/3/2/1.html>
- MASON, B, and B Dicks. 2001. "Going Beyond the Code: The Production of Hypermedia Ethnography." *Social Science Computer Review* 19: 445-457.
- MASON, B, and B Dicks. 2004. "The Digital Ethnographer." *Cybersociology* 6.
- MISHLER, E. 1991. "Representing Discourse: the rhetoric of transcription." *Journal of Narrative and Life History* 1: 255-280.
- MUHR, T. 2000. "Forum: Increasing the Reusability of Qualitative Data with XML." *Qualitative Social Research* 1.
- O'CONNOR, H, and C Madge. 2001. "Cyber-Mothers: Online Synchronous Interviewing using Conferencing Software." *Sociological Research Online* 5.
- PARRY, O, and N Mauthner. 2004. "Whose Data Are They Anyway? Practical, Ethical and Legal Issues in

Archiving Qualitative Research Data." *Sociology* 38: 139-151.

PFAFFENBERGER, B. 1988. "Microcomputer Applications in Qualitative Research." *Qualitative Research Methods: A Sage University Paper* 14.

RAYMOND, M, and E Lea. 2000. "Computer Software and Qualitative Analysis: trends, issues and resources." *Social Research Methodology* 3: 231-243.

RICHARDS, T, and L Richards. 1995. "Using Computers in Qualitative Research." in *Handbook of Qualitative Research*, edited by Y. S Lincoln: Sage.

SARIS, W. E. 1991. *Computer-Assisted Interviewing*. London: Sage.

SEIDEL, J, and J Clark. 1984. "The Ethnograph: a computer program for the analysis of qualitative data." *Qualitative Sociology* 7: 110-125.

SELWYN, N, and K Robson. 1998. "Using E-mail as a Research Tool." *Social Research Update* 21.

SPROULL, L, and R Sproull. 1982. "Managing and Analysing Behavioural Records: explorations in non-numeric data analysis." *Human Organisation*: 283-290.

STOCKDALE, A. 2003. "Tools for Digital Audio Recording in Qualitative Research." *Social Research Update*.

TESCH, R. 1990. *Qualitative Research: analysis types and software tools*. London: Falmer.

WEITZMAN, E, and M Miles. 1995. *A Software Source Book: Computer Programs for Qualitative Data Analysis*. London: Sage.

WILLIAMS, M, and K Robson. 2004. "Reengineering Focus Group Methodology for the Online Environment." Pp. 25-45 in *Online Social Research: Methods, Issues and Ethics*, edited by J Hall. New York: Peter Lang.

<<http://www.accustreamresearch.com/products/streamingmedia2003.html> >

<<http://www.realnetworks.com/>>