Using software for qualitative data analysis: research outside paradigmatic boundaries

Jonathan Tummons, School of Education, Durham University, Leazes Road, Durham DH1 1TA

jonathan.tummons@durham.ac.uk

Abstract

Purpose

This chapter aims to explicate the use of computer software for qualitative data analysis. Drawing on both a review of relevant literature and a reflexive commentary on an ongoing ethnography, this chapter argues that the use of computer software for qualitative data analysis facilitates rigour and reliability in research, whilst also contributing to wider debates regarding the distinctions made between different research paradigms.

Design/methodology/approach

The chapter is divided into two sections. In the first, a review of literature pertaining to the use of computer software for qualitative data analysis is reported. The key themes to emerge from this review are then explored in the second section, which consists of a reflexive commentary on the use of computer software for qualitative data analysis within an ongoing three-year Canadian/UK research project.

Findings

The chapter concludes firstly by foregrounding the methodological benefits of using computer software for qualitative data analysis, and secondly by commenting on wider debates relating to the historical distinctions between quantitative and qualitative research paradigms.

Practical implications

The chapter suggests that the uptake of computer software for qualitative data analysis should be considered as an integral element of the research design process.

Originality/value

The originality of this chapter rests in its focus on methodology rather than method, on a reflexive discussion of the place of computer software within the research process rather than a technical description of how software should be used. This chapter is of value not only to researchers who are using or considering using software for their research, but also to researchers who are engaged in wider methodological discussions relating to qualitative and quantitative research paradigms, and to research quality and generalisability.

Key words

Author note

Jonathan Tummons is lecturer in education and pathway leader for the MSc in Educational Assessment at Durham University. In addition to ongoing research into learning, teaching and assessment in further and

higher education, he is currently acting as co-investigator for **'**Higher Education in a Digital Economy: An Institutional Ethnography**'**, a three-year research project based at Dalhousie University, Nova Scotia, and funded by the Canadian Social Sciences and Humanities Research Council. He is the author of a number of books and book chapters relating to further, higher and adult education, and has published in a number of leading journals, including *Studies in Higher Education*, *Assessment and Evaluation in Higher Education*, *Higher Education Research and Development* and *International Journal of Educational Research*.

Introduction

What is the current condition of the use of specialist computer software for qualitative data analysis? Is it indeed the case that it has now become so (relatively) common for a researcher or a team of researchers to use software such as Atlas-Ti or Nvivo to manage their projects that it does not need to be mentioned in the methods section of a research article (Seale and Rivas, 2012)? Or is it in fact the case that the use of such software continues to be contentious, somehow causing losses in some aspects of the research process or otherwise generating problems of theory and/or method for the researcher (King, 2010)? Does the lack of reference to software in methods sections denote a greater familiarity with and consensus regarding the use of such software, or does it simply reflect the fact that the writing up of much qualitative research is instead characterised by scant regard to method and theory (Tight, 2004; Trowler, 2012; Tummons, 2012), whether or not qualitative data analysis software has been used?

In this chapter I provide an account of the use of specialist software for the analysis of qualitative data that rests on both theoretical/methodological and empirical perspectives. The theoretical/methodological perspective is derived from a discussion of current and recent literature relating to the use of software. In this section of the chapter, I argue that whilst some of the critiques of such software to have emerged within the literature during the last (almost) thirty years continue to make valid points that should be a concern to the reflexive researcher, other longer-standing critiques have outlived their usefulness and their applicability. Following this, I provide an empirical perspective through an account of the ongoing work that I am currently involved with as a co-investigator on a three-year institutional ethnography, *Higher Education in a Digital Economy*. Through providing an account of the use of Atlas-Ti by the research team, I argue that this software provides us, as researchers, with ways of working that can serve to problematise the so-called paradigmatic distinctions between qualitative and quantitative research in terms of not only how data analysis is *operationalised* (Sin, 2008), but also in terms of the *transparency* of the research work being done as an element of research quality (Hammersley, 2008).

A brief note on nomenclature

The term CAQDAS is commonly used within the literature. Somewhat confusingly, this acronym can be arrived at in two slightly different ways. Firstly, there is **C**omputer **A**ssisted **Q**ualitative **D**ata **A**nalysi**S**, leading to constructions such as 'CAQDAS software'. Secondly, there is **C**omputer **A**ssisted **Q**ualitative **D**ata **A**nalysis **S**oftware, the term that is employed here. The term QDAS (**Q**ualitative **D**ata **A**nalysis **S**oftware) is also in use.

Using software for qualitative data analysis: some recent (and not so recent) persistent concerns

A number of themes emerge from a review of literature pertaining to CAQDAS. Some of these – as might be expected when discussing ICTs – have receded or changed over time and new issues have emerged: this can be seen in, for example, the changing emphasis on CAQDAS as a tool for helping with the analysis of not only textual data, but audio and video as well. The ways in which discussion around CAQDAS has shifted over time will be considered. At this time, I also want to unpack and respond to those other themes that have entrenched themselves within CAQDAS literature, that have been reiterated and revised over time but that CAQDAS communities, it would appear, cannot quite shrug off: a straightforward example of this can be seen in the recurring notion that the use of CAQDAS might in some ways serve to distance the researcher from her or his data. I suggest that these kinds of issues can be conveniently if not quite perfectly considered in four ways (acknowledging that there is some overlap between these and that these are not intended to be discrete classifications), in terms of: closeness to the data; driving the research process; theory-building using CAQDAS; and attitudes towards CAQDAS.

[i] closeness to the data

It might at first seem rather odd to discuss the researcher's relationship to her or his data using some of the terms that reoccur frequently within the CAQDAS literature. Terms such as 'separation' (Smith and Hesse-Biber, 1996) 'closeness' (Weitzman, 2000) 'distance' (Gibbs, 2007) and 'reduced proximity' (Roberts, Breen and Symes, 2013) are used to describe the first aspect of CAQDAS to be unpacked here, namely the notion that the use of CAQDAS in some way serves to separate or distance the researcher from her data. There is something inherent in the use of a computer, as distinct from reams of paper that have (one assumes) been sifted, cut up, colour-coded, stuck onto notice boards and so forth, that prevents the researcher from achieving the required level of 'closeness' to their data (whatever that might actually mean).

Where does this aversion to the use of computers to do work (research is, after all, a form of work) come from? The current ubiquity of ICTs at conferences, in offices and in workplaces more generally would seem to render such an aversion somewhat idiosyncratic at best. One of the earliest statements regarding this concern over closeness in the use of computers to assist in qualitative research can be found in a research paper from 1988, when Tesch (1988: 179) argued that researchers suspected that using computers might distance them from their research material. Certainly, early adopters of computers for qualitative data analysis (including, but not restricted to, the use of early specialised software programs such as The Ethnograph) were choosing to do their work in a very different manner from those researchers who were continuing to use needles to sort their edge-notched cards into piles for the purposes of theory building. At the same time, these latter researchers could, quite correctly, point to the fact that the hardware and software of the time posed restrictions (in comparison to current CAQDAS) in terms of the number of interview transcript files that they might be able to manage within their software programs or the maximum permitted length of a coded segment of text. And yet despite these restrictions (to which I shall return), early users of computers for qualitative research were under no illusions as to the labour-saving potential that computers offered, closeness to the data notwithstanding, highlighting a number of mechanical tasks that

might be ameliorated through the use of computers, such as searching notes for specific passages (Brent, 1984).

[ii] driving the research process

Recent commentators have stressed the ways in which CAQDAS can, for example, allow the researcher to search very quickly for specific segments within large bodies of text (Lewins and Silver, 2007). But the claim (above) that using CAQDAS might dictate the process of data collection and analysis is more serious, and perhaps might get us closer to understanding the concern for closeness. The notion that the software drives the research process in some way therefore becomes the second aspect of CAQDAS to be unpacked here. According to these concerns, CAQDAS provides a series of methodological straightjackets that hamper the work of the researcher. Examples include the ways in which particular software packages impose particular coding structures on the researcher. The two most frequently cited examples of this phenomenon are the imposition of coding hierarchies in Nvivo, and the flat coding structures within Atlas-Ti that promote a grounded theory approach by the researchers who use it (Coffey, Holbrook and Atkinson, 1996; Weitzman, 2000; Willis and Jost, 1999). Certainly it would be an undesirable consequence of CAQDAS uptake if particular forms of or approaches to qualitative research were to be lost sight of.

But to what extent does any social practice (research is, after all, a form of social practice) shape or get shaped by the tools and artefacts that are used by practitioners? Consider the facilities offered by current CAQDAS in terms of the different file formats that they support. The most recent version of Atlas-Ti (version 7 – this is the version used by the *Higher Education in a Digital Economy* research team) can support not only Microsoft office files but also pdf files and web pages, all the time maintaining the format and colour of the original files as they are loaded into an hermeneutic unit (HU – the name given to a project in Atlas-Ti, which acts as a kind of digital 'container' for all of the primary documents being used). Sound, picture and video files in a variety of formats (.wav, .mp4, .jpeg and so forth) can also be loaded into Atlas-Ti. All of these documents can be coded and searched. This is in stark contrast to early iterations of CAQDAS, which could only support plain text files. Facilities for collaborative work have also been greatly expanded in recent software releases. In Atlas-Ti it is now a simple task to bundle an entire HU (primary documents, codes, memos and all) and email it to a colleague. And both Atlas-Ti and Nvivo can – after that the correct software licenses have been purchased, be installed on network servers, allowing for the simultaneous reading, coding, analysis or memoing of a document by multiple researchers.

Of course, not everyone will use her or his software in quite the same way. Some researchers will make use of only a relatively small number of CAQDAS features in order to perform relatively simple 'code-and-retrieve' functions (arguably the most ubiquitous of all CAQDAS features (Kelle, 1997)). Some will perform searches using Boolean operators in a manner akin to conducting library catalogue searches (OR, XOR, AND, NOT) – for example, by searching a data set for quotations that have been tagged with both one code AND another, or for quotations that have been tagged with either only one code or only another, but not both (XOR) (Friese, 2012). Others will create network views in order to visualise their data (Lewins and Silver, 2007). As with any technology, there will be some users who operate at an instrumental level, only drawing on a small number of available functions, and other users who operate at a more fluent or expert level, who use a wider range of functions in a more systematic and critical manner (Mangobeira, Lee and Fielding, 2004; Odena, 2013). And software functions are just one issue to consider when reflecting on the practices of the researcher: what

about the methodological assumptions upon which the software rests? Is it indeed the case that CAQDAS in some way implicitly supports and promotes a grounded theory approach to qualitative research (van Hoven and Poelman, 2003)? Certainly, the predominance of code-and-retrieve functions in early iterations of CAQDAS might be seen as encouraging a predominantly 'grounded' approach to data analysis (Coffey, Holbrook and Atkinson, 1996).

Such concerns are in fact robustly countered in CAQDAS literature. The perception of a grounded theory bias within CAQDAS is countered not only by a reminder that the functions offered by CAQDAS are employed within other methodological frameworks (Kelle, 1997) but also by a reminder that grounded theory is an ambiguous methodology at best that has only ever been attached to CAQDAS on an erroneous basis (MacMillan and Koenig, 2004; Roberts, Breen and Symes, 2013; see also Thomas and James, 2006). Indeed, the theoretical vagueness that troubles qualitative research more generally can be seen as contributing to the conceptual confusion that surrounds the use of CAQDAS (MacMillan and Koenig, 2004). CAQDAS users also remind us that software is just a tool, not a methodology: the software does not do the analysis; it simply facilitates the analysis done by the researcher. Therefore the fact that many CAQDAS users have chosen to draw on grounded theory should not be considered as consequent to the use of the software (Dormady and Byrne, 2006; Van Hoven and Poelman, 2003). CAQDAS does not drive the researcher towards grounded theory, therefore.

[iii] theory-building using CAQDAS

The term 'theory-building' originates in an early classification of types of CAQDAS, which were divided up into categories such as 'text-retrievers' (used for operations such as frequency counting), or 'code-and-retrieve programs' (Weitzman, 2000). A further category, 'code-based theory builders', was used to describe those programs that offered functions such as hyperlinking or graphical network modelling to help the researcher draw links between text segments, codes and memos, although such classifications were by no means accepted by all CAQDAS users (Kelle, 1997; Lewins, 2001). And although the increasing sophistication of CAQDAS has over time rendered these categories obsolete, it can be argued that the term 'theory-building' has over time been the victim of conceptual slippage, with software being given an undue prominence in the processes of analysis and theorisation, leading to the (so-called) *mechanism* (Garcia-Horta and Guerra-Ramos, 2009: 163) or *mechanisation* (Roberts, Breen and Symes, 2013: 280) of qualitative data analysis through a perceived over-reliance on software tools, in particular auto-coding tools. But to assume that the software somehow does the thinking for the researcher is a mistake and the literature reminds us time and again that CAQDAS can provide the tools, but it cannot do the analysis – a concern that has been dismissed by one commentator as a form of slight paranoia about technology more generally (Seale, 2005: 197).

I shall return to this theme of 'paranoia' shortly. At this time, it is important to stress that the tools and opportunities offered by CAQDAS are a reflection of software construction, not methodology, and that what is being constructed are tools to help the researcher do her or his work and nothing more. Atlas-Ti no more 'does' the analysis within *Higher Education in a Digital Economy* than Adobe Acrobat (one of the other programs that we use) does. It is therefore necessary to cut through the hyperbole that has sometimes surrounded CAQDAS in the past, hyperbole that has had the effect of using computers to add a "sheen of scientific rigour" to the analysis of qualitative data (Darmody and Byrne, 2006: 123), or a "wow factor" of

mystique surrounding the use of software that serves to generate unrealistic expectations (MacMillan and Koenig, 2004: 180).

The possible impact of auto-coding functions is more problematic, however. The use of auto-coding is described as debateable at best by Lewins and Silver (2007), who argue that it risks limiting the analytical process and lowers the status of any text excluded by the process, thereby echoing earlier concerns that an excessive focus on coding at the expense of other methodological tools risks decontextualizing data (Seidel and Kelle, 1995; Kelle, 1997). Such criticisms rest on the notion that the use of CAQDAS might lead researchers to do too much coding because it is so straightforward to accomplish, leading to "data-fetishism" (Garcia-Horta and Guerra-Ramos, 2009) or a "coding trap" (King, 2010) in which the researcher is surrounded by an excess of codes which distort the rest of the research process. However, it is the researcher and not the computer who defines auto-coding parameters (Odena, 2013). Auto-coding is a function like any other within CAQDAS, that needs to be used properly and carefully a part of a well thought through methodology, for which CAQDAS cannot provide a substitute.

[iv] attitudes towards CAQDAS

The fourth and final concern that I wish to unpack is what might be termed attitudinal responses towards CAQDAS. By this I mean to draw attention to the ways in which CAQDAS 'sceptics' (Odena, 2013: 355) are positioned in relation to the use of software for qualitative data analysis. To some extent, the underlying concerns that will lead to some researchers being seen as apprehensive (Tesch, 1988: 179) or cautious (Bathmaker, 2004: 175; Van Hoven and Poelman, 2003: 114) can be understood in the light of the kinds of issues already unpacked – the concerns that proximity to data will be affected, that software design affects methodological choices, and that computation might lead to automation.

It is also important to acknowledge the historical context of CAQDAS usage. Seale and Rivas (2012), Smith and Hesse-Biber (1996), Tesch (1988), and Willis and Jost (1999) all note that early adopters of CAQDAS were working at a time where prevailing attitudes towards the use of computers in social research might be seen as resting within a positivist paradigm, aligned to the early predominance of software for statistical data analysis in relation to software for qualitative data analysis. In such a climate, the scepticism that can surround the use of computers in the "naturalistic, phenomenological" real of qualitative research can all too easily be seen as distancing researchers from real-world ethnography or other forms of qualitative work, a sentiment added to by the functions of those CAQDAS programmes categorized as "text retrievers" and "textbase managers" (Weitzman, 2000), which were used for functions such as word frequency counting or string searching that generate statistical outputs which appeared to 'belong' to a quantitative paradigm. I shall return to the notion that CAQDAS might serve to blur the boundaries between qualitative and quantitative work later, but for now I wish to stress that the personal preferences of the researcher should be acknowledged and that these preferences will have in part been shaped by that researcher's history, prior research experiences and so forth, a history in which the distinctions between qualitative and quantitative work will almost certainly have been reified within training programmes, methods textbooks and the like (Cooper et al., 2012: 2-3).

Using software for qualitative data analysis: perceived advantages

A number of themes emerge repeatedly from a review of literature pertaining to CAQDAS. There is, arguably, a strong consensus in literature relating to the advantages or benefits of using CAQDAS for qualitative research. Where there is some variation is in the extent to which some of these advantages can be seen as operating at a more than technical level. That is to say, there are some advantages to using CAQDAS that might be seen as producing positive effects that go beyond being related to efficiency or productivity, for example, and instead produce effects that have a meaningful impact on research method and quality, for example. I shall discuss each of these in turn (again, mindful of the fact that these two categorisations are loose and not intended to be discrete). At the same time, it is important to remember that this discussion rests on the assumption that the nature of research work or analysis being done is such that the use of CAQDAS is meaningful and worthwhile – some forms of qualitative analysis such as conversation analysis or discourse analysis where quite small data sets are used being examples of research work that would not particularly benefit from using CAQDAS.

The advantages that can be gained from using CAQDAS include: the convenience of being able to store and manage data sets in one digital location (which has been further facilitated in recent years as updates to CAQDAS have included support for an increasingly wide variety of file formats); quick and easy access to data; systematic and consistent data management; greater speed in searching and re-searching texts; increased and more convenient access to both whole data files and segments of files; simple and convenient tools for keeping track of developing ideas during the research process; the gradual incorporation of new documents within existing data sets; and portability (further facilitated in recent years as laptop computers have become more powerful – in addition, a version of Atlas-Ti for ipad has also been announced). All of these facilities in turn make it relatively straightforward for tams of researchers to work on the same data sets: the ability to easily and quickly export and share bundles of data, graphical representations or networks, groups of memos or code families (or, indeed, all of these) allows for multiple researchers to work on the same project without the need for physical proximity. In turn, these tools also afford such collaborative work a high degree of consistency and hence reliability, which will be discussed in more detail below.

However, whilst the efficiency and productivity effects of using CAQDAS can be seen as being relatively uncontested and certainly uncontroversial, other aspects of using CAQDAS can be seen as having a more profound impact in terms of firstly, method and secondly, quality (once again, with overlap between these two categories). I shall discuss four key issues: system closure; visibility and transparency; rigour and reliability; and qualitative/quantitative blurring.

[i] system closure

System closure is the term used to refer to the practice of including not only primary documents but also secondary documents such as memos, graphical representations, search results, notes or other working papers within the analysis process. The use of CAQDAS makes it a simple task for the researcher to search and then code her or his ongoing analytical or explanatory material using the same coding structure as has been used for the primary data (Bathmaker, 2004: 168; Richards and Richards, 1994: 449; Weitzman, 2000: 809).

[ii] visibility and transparency

The idea that using CAQDAS renders the work of qualitative research more transparent and more visible runs strongly through the literature (Gibbs, 2007). This is not due to any nebulous "wow factor" that CAQDAS can – erroneously – be seen to apply to the research process. Using CAQDAS to do qualitative research requires the researcher to engage in careful and thorough research design and in precise conceptual thought and analysis to the exact same degree as the researcher who choose not to use CAQDAS (MacMillan and Koenig, 2004). We should not therefore be fooled by any "sheen of scientific rigour" that CAQDAS might – equally erroneously – be seen to apply (Darmody and Byrne, 2006: 123). Rather, the use of software provides the technical means by which different elements of the ongoing research process – such as memoing or coding – can be straightforwardly captured and made visible to research users. As such, it becomes more straightforward to describe and illustrate the work of qualitative data analysis in greater detail, which in turn enhances the robustness of the claims that arise from the research (Odena, 2013). This straightforwardness can be seen as being equally applicable to different members of a research team, and to the end users of a research project. It is also applicable to research participants. Although respondent validation, as a discrete topic, is under-represented in CAQDAS literature it can be argued that CAQDAS, as a tool that helps make the processes of data management and analysis more visible, would thereby enhance respondent validation through making the analytical steps taken by the researcher more straightforward for the researched to scrutinise.

[iii] rigour and reliability

Greater visibility and transparency can in turn be seen to lead to greater rigour and reliability in qualitative research. More detailed accounts of qualitative data analysis become more straightforward to produce, as a response to the persistent claims that too much qualitative research pays insufficient attention to methodology and theory (Sin, 2008; Trowler, 2012). It becomes easier (assuming that permissions have been agreed) for users to view and reflect upon research materials (Davidson and di Gregorio, 2011). Specific aspects of the research process such as coding can be audited (Gerson, 1984; Roberts, Breen and Symes, 2013). But it is important to note that just as CAQDAS cannot compensate for poor research design, so CAQDAS cannot in and of itself generate greater reliability and rigour during the research process. Rather, it is through the use of CAQDAS that tools are made available to the researcher that will make it simple for her to demonstrate and describe her research work in such a way that the claims or warrants that are being made for the research are robust.

[iv] qualitative/quantitative blurring

The extent to which the boundaries that are said to exist between and hence to distinguish and define qualitative and quantitative research paradigms lie outside the scope of this chapter. They have been extensively and convincingly discussed elsewhere in terms of issues ranging from the extent to which one paradigm or the other is more or less positioned on an inductive or deductive mode of inference or the extent to which one is more objective and the other is more subjective, to whether or not the research in question primarily uses numbers rather than words or the nature and scope of the sample size used in the research. Put simply, paradigmatic debates such as these rest on both methodology and method (Bryman, 2008).

The emergence and development of increasingly sophisticated forms of CAQDAS can be seen as contributing in some way to the broader debate as to the applicability and desirability of the maintenance of

such paradigmatic distinctions. In part this is a consequence of the construction of some of the earliest CAQDAS programs such as 'text-retrievers' (Weitzman, 2000), which provided researchers with the tools to, for example, compile frequency counts which might then be exported to statistical software packages such as Excel (Weber, 1984; Willis and Jost, 1999). As such, one of the characteristic features of CAQDAS can be seen as providing tools for the construction and analysis of numerical as well as textual data. Typical examples of the kinds of numerical data that might be derived from a predominantly text-based data set include frequency counts of a particular word or phrase (for example, the use of a key word or phrase within a policy document or across a number of documents if a comparison is being sought) as well as frequency counts of code usage (for example, in order to compare the prevalence of one code or theme in relation to another, within and/or across different documents). A second characteristic feature of CAQDAS, albeit one that is a more recent phenomenon consequent to ongoing technological innovation, is the capacity for CAQDAS to manage increasingly large and more diverse sets of data drawn from correspondingly larger and more diverse sample populations (Friese, 2012; Seale, 2005; Smith and Hesse-Biber, 1996). In this way it can be argued that the use of CAQDAS contributes to wider debates around the qualitative/quantitative divide through encouraging the very kind of movement both among and across so-called research paradigms that have been entrenched by the politics of method (Cooper et al., 2012: 8). I shall return to this theme later.

Using software for qualitative data analysis: *Higher Education in a Digital Economy*

Higher Education in a Digital Economy (HEDE) is a three-year institutional ethnography, funded by the Social Sciences and Humanities Research Council of Canada (SSHRC). Institutional ethnography is a framework for qualitative inquiry derived from the work of Dorothy Smith (2005) that focuses on everyday activity – work – as a way of investigating the organisation of social life, with a particular focus on the ways in which work is mediated or ordered through text-based artefacts (Tummons, 2010). The broad aims of the project are to explore issues that surround the implementation of a new medical education curriculum that is enacted simultaneously across two locations in Canada (New Brunswick and Nova Scotia) that are approximately 300 miles or 480 kilometres apart. This new Distributed Medical Education (DME) curriculum has been designed to rest on information and communication technologies (ICTs) 'from the ground up': that is to say, the use of technology (digital video, digital learning platforms, e-learning devices and such like) functions as a means to enact synchronously a curriculum across two distinct locations, as distinct from the use of technology as an 'additional' feature within a curriculum that could still be delivered were the technology not present. Thus, instead of simply designating a curriculum as being an example of 'blended learning' through the post- hoc provision of e-learning resources alongside or on top of an existing 'real world' curriculum delivery model, this new medical education curriculum can be understood as only being possible through the affordances offered by ICTs. Without ICTs, this curriculum could not have been written and enacted in the ways that it has been.

The project is approaching the halfway stage, having been active for about eighteen months. Much has been accomplished in this period. The literature relating to distributed medical education has been reviewed. The

theoretical tenets of institutional ethnography and actor-network theory (the two significant theoretical foundations for the project) have been debated, critiqued and sometimes disagreed with during the research team's online meetings (which have ranged in style from informal discussions to formal presentations by individual members of the team). Policies and protocols for the analysis of paper-based and online textual documents, including photographs and videos (one of the major sources of primary research data for the project) have been discussed, piloted and then rolled out across the research team. Thus far, 60 different texts ranging from institutional policy documents to YouTube videos have been analysed by 11 different members of the research team. The first tranche of semi-structured observations has also been carried out. At the time of writing, 108 observations of lectures, seminars, and staff meetings have been conducted by 5 members of the team across the two research sites and a framework for analysis based on Spradley (1980) has been discussed, piloted and then operationalised. Data from the observations is, at the time of writing, being analysed as part of the preparation of two distinct papers being written by different members of the research team. Protocols relating to access and use of the data developed by the team are close to being finalised: an 'open access' approach has been established in order to allow shared access to and ownership of the data across the research team (and which in itself is the focus of a third paper being prepared by some of the other team members). And finally, and perhaps most importantly, the members of the research team have got to know each other, to talk, joke and share frustrations with each other as we discuss issues such as data access, the analysis of online as opposed to paper-based texts, or the desirability or otherwise of anonymity in research.

Mindful of the distributed nature of this new curriculum, it is perhaps appropriate that the research team that is exploring this new curriculum – both its adoption and the ongoing experiences of the staff, students and faculty who are enrolled within it – should be similarly distributed, and thus similarly reliant on ICTs for their work together. The research team consists of eighteen people: the majority of the team are in Canada (distributed across three provinces), and two are in the UK. The research team uses a number of different technologies in order to facilitate working together. Project documentation is stored online using *Mindmeister*, an online mind mapping tool which can also be used for data management and storage. Through the Mindmeister portal, any member of the research team can access any of the materials that have been collected or generated thus far, ranging from the initial grant documents to working papers written by members of the research team, from PowerPoint slides generated by team members for conference presentations to copies of the minutes from research team meetings. It should also be noted that access to the Mindmeister portal is open, in keeping with the open access policy of the project, although some sections (such as those containing raw project data) are password protected. Team meetings are facilitated online using GoToMeeting, a video-conferencing and web meeting tool. As well as allowing for virtual face-to-face communication through webcams and headsets, GoToMeeting facilitates the sharing of documents across the team, in a manner akin to the tabling of hard copy documents at a 'real world' meeting. Thus, PowerPoint slides or PDF files can be 'tabled' and discussed during the meeting. A pop-up screen allows users to toggle their microphones and cameras on and off, and also contains a messaging function that allows the user to send a text-based message to one or more of the other attendees. Finally, entire meetings can be saved and then stored online for future reference.

It can clearly be seen that the new DME curriculum and the research team that is exploring it are accomplished through and because of ICTs. Both the curriculum and the research team rely on technology and are mediated through technology: spoken words and written texts (lectures, curriculum documents, teaching resources, faculty team meetings, students' assignments, slides, prospectuses and so forth) are distributed or stored online. As such, it is perhaps unsurprising that the research team has chosen to use CAQDAS.

Choosing and using software for qualitative data analysis: the affordances of technology

After some discussion, the HEDE team decided to use Atlas.Ti version 7. Whilst the majority of the team is quite comfortable with using ICTs more generally, the choice of CAQDAS required a little more thought. In part this was because only a small number of the team had used CAQDAS previously (both Atlas-Ti and Nvivo/NUD*IST), and in part because some of the team had not conducted ethnographic research before and this relative lack of methodological experience (it should be pointed out that they were are experienced in other forms of research work) resulted in a lack of familiarity with qualitative data analysis more generally rather than the facilities offered for analysis by software. It was decided to proceed with Atlas-Ti primarily because it was felt by the research team that this CAQDAS offered superior and more straightforward tools for the sharing of research work (data, codes, memos and so forth) across the team. The future possibilities of both native mac (at the moment the research team are using Atlas-Ti with windows emulators on apple mac laptops) and ipad versions of the software as also attractive to the team, reflecting once again the importance of personal preferences when using CAQDAS (King, 2010).

A reflexive account of the work of the HEDE team would find it difficult to unpack, exactly, the relationship between the research team, the research project being undertaken, and the role of CAQDAS in the project. At one level it seems right to acknowledge that the specific nature of the field being studied – a field that consists of both the physical and the virtual – is ideally suited to research facilitated by CAQDAS. Many (though by no means all) of the interactions, social practices and artefacts used by the staff and students who are enrolled within the curriculum exist in virtual rather than physical spaces, and often across both. The experience of a plenary lecture, delivered in one location (it could be either Nova Scotia or New Brunswick, although the majority take place in the former, which is also the larger of the two faculties) but simultaneously streamed in the second, provides a good example. The lecturer is required to stand or walk only in a very narrow space, marked out on the floor, so that her or his image and speech can be reliably captured. Whilst attending to the students who are present in the same physical lecture room, the lecturer also has to be mindful of the students who are present in the remote location: s/he has to observe the remote students who only appear on screen whilst also paying attention to her/his teaching materials that will appear on a different screen. Something as simple as walking around the lecture room in order to gain students' attention or to emphasise a point becomes impossible. In order for students at both sites to take part in question-and-answer sessions, a push-button system has been introduced. Students – at both sites – push a button next to an adjacent microphone when they wish to raise a point, and the lecturer pushes a

button at the podium in order to 'activate' the next question in the queue. Questions are answered in the order that they are 'asked', not in any order of relevance or logical progression to a preceding point or theme.

A variety of technologies (webcams, laptops, software, web browsers, ipads) change how the students and the lecturers talk, look, act and even write. How they behave, the kinds of artefacts that they use, how they talk and how they make meaning – all of these social practices are mediated to some degree by technologies, by the virtual. It makes sense, therefore, to use CAQDAS to capture these online as well as physical practices. Field notes can be transcribed and loaded into the software, where they can sit alongside the pdf files or PowerPoint slides that were used by lecturers, which in turn can sit alongside both audio recordings and transcriptions of interviews with staff and students. All of these different modes of data can be gathered in a single repository, which can in turn be easily distributed across the research team (Sin, 2008).

At the same time, it could be argued that doing research within this field is only made possible in the first place by the affordances of technology, without which some of the research goals of the project would be impractical at best and impossible at worst (Mangabeira, Lee and Fielding, 2004). By this I mean to stress that it is only because Atlas-Ti has the functionality that it has, that the HEDE research project can seek to explore the questions that it is seeking to explore. Indeed, it is doubtful whether an earlier iteration of the software would have been able to operationalise the HEDE research in a similarly comprehensive manner. The combination of the physical and virtual that resides within the research team is matched by the field being researched, and the many physical as well as virtual artefacts that the research team are working with could be explored 'off-line' only with extreme difficulty. The presence of two transatlantic members of the research team would certainly be impossible without ICTs, including CAQDAS; and the cooperation between the two faculty sites in Canada would be rendered impracticable at best. It might just be possible to distribute copies of the different primary documents across physical as opposed to virtual spaces, but it would require a considerable amount of printing, copying, and posting. Opportunities for offline collaborative coding would be so difficult as to be impossible without significant amounts of time and resources to facilitate travel and accommodation so that the team could meet and talk: it is hardly surprising that opportunities for facilitating research within teams is highlighted as one of the benefits of using CAQDAS (Friese, 2012; Lewins and Silver, 2007).

There are several distinct, though overlapping, themes at work here, therefore, when considering how the HEDE project might be and is being accomplished or operationalised. Both the researchers and the researched are distributed across physical and virtual boundaries, and both the field of research and the artefacts that enrich that field are similarly distributed across and reified within both physical and virtual forms. So how do we theorise the position of CAQDAS within this complex research field? As an enabler that affords us, as researchers, the tools to conduct our research? Has the shape of our research project been driven in some way by the facilities that Atlas-Ti provides for us? IF we had used Nvivo, would the final shape of the research be different? Could this research be done without Atlas-Ti? There are no simple answers to these questions because of course they are hypothetical. Instead, as a crucial component of the reflexivity that ought to accompany any ethnography (and not just an ethnography that is saturated with

technology), we need to be aware of the ways in which our choices – including our choice to conduct observations in some locations and not others, or our choice to conduct content analysis on some documents and not others, or our choice to use Atlas-Ti and not Nvivo – has shaped our research as a whole.

Choosing and using software for qualitative data analysis: reliability, generalisability and the politics of method

Whilst the exact relationship between our research questions, our research field and our research tools remains to some extent problematic, the possible consequences for research reliability and generalisability of our decision to use Atlas-Ti are relatively more straightforward. By this I do not mean that they lack complexity or difficulty; rather I mean to counter the notion of the *problematic* as it is understood within institutional ethnography, as describing problems or questions that may not yet have been posed but which are nevertheless latent in the experiences of a social actor (Campbell and Gregor, 2004; Smith, 2005). That is to say, the relationship between our research questions, or field and our tools is problematic: it is latent in our work as researchers but is only gradually emerging as a theme for discussion and analysis within the research team. But the impact of those tools – to be precise, of one tool in particular, namely Atlas-Ti – can be theorised with more certainty. There are two elements to this theorisation: reliability and generalisability, and the politics of method (and as before, these should be understood as complementary, not competing).

[i] reliability and generalisability

I have already established that CAQDAS allows a researcher or a team of researchers to work to a high degree of consistency and accuracy. With CAQDAS, document searches are likely to be comprehensive as well as fast, and this speed facilitates searching and re-searching, coding and re-coding. Memos can be coded in turn and all of the data sets – video, audio, text, image, web – are searched, codes and memoed in exactly the same way. Code families, memos or even entire research projects can be easily shared, compared, updated and merged. And CAQDAS allows all of the steps taken by us, as researchers, to be clearly documented for the scrutiny of other research users. The data management and analysis processes are transparent, consistent, accurate and rigorous – and these four qualities are, arguably, all ameliorated by the use of software. I have also already established that CAQDAS allows a team of researchers to draw on data sets of significantly greater size as well as modality than would normally be possible with a research team who chose not to use software, because these same functions – of speed, of capacity of storage, of sharing – make managing large data sets more practicable. The first aspect of using CAQDAS that I wish to posit here, therefore, is that using CAQDAS allows for more consistent and more thorough analysis of larger data sets, meaning that larger research samples can be explored. These two elements – consistent analysis and cross-analysis by a team of researchers who are working with qualitative data sets of significant size –

combine to refute one of the claims that is often made against qualitative research, namely that qualitative research is partial, excessively subjective and lacking in robust generalisability. CAQDAS allows teams of researcher to test and retest their ideas (or hypotheses?) in ways that can be traced and then demonstrated to other research users, across large-scale data sets that can equally easily generate textual and/or numerical (mindful of my earlier discussion relating to 'text-retrievers') outputs.

[ii] the politics of method

I subscribe to the arguments that are both reiterated and expanded by Cooper et al. (2012), namely that many of the distinctions that are drawn between quantitative and qualitative research paradigms are both artificial and unhelpful. I also subscribe to the arguments made by MacMillan and Koenig (2004), Sin (2008) and Odena (2013), amongst others, namely that the impact of qualitative research – the robustness of the claims made by qualitative research – needs to be persuasive and credible. The use of CAQDAS, as explored in literature and as illustrated by the HEDE project – reinforces both of these lines of argument. Whilst it is not an *a priori* necessity to use CAQDAS for any research project, to do so allows the researcher or researchers to draw on significantly large sample sizes and a variety of modes of data, to analyse and re-analyse this data quickly, efficiently and reliably, to open up the data to entirely new patterns of analysis and to revise existing patterns, and to share the data, code families, memos and so on conveniently and comprehensively. All of these operations can be done with an enhanced transparency, ameliorating the visibility of the research process, in turn aiding credibility and rigour. CAQDAS software does not discriminate between numerical or textual outputs: it can generate reports of both kinds. As such, I suggest that a reflexive consideration of qualitative research that uses CAQDAS embodies the need to take a different attitude towards the so-called qualitative/quantitative divide, instead providing a focus on the research questions to be answered and the credibility of those answers, rather than sustaining a focus on spurious distinctions between, or characteristics of, qualitative and/or quantitative research.

Conclusions: research outside paradigmatic boundaries

Attitudes towards CAQDAS have, of course, changed over time. Whilst it may still be a matter of debate as to whether the use of CAQDAS is now "routine" (Seale and Rivas, 2012: 432) or "contentious" (King, 2010: 6), or 'critical' or 'instrumental' (Mangabeira, Lee and Fielding, 2004), it cannot be denied that CAQDAS has changed how qualitative research might be done. However, these changes are rather different to those envisaged by early commentators. CAQDAS has not generated homogeneity amongst researchers, with grounded theory crowding out other forms of analysis. Nor has it 'quantified' qualitative research by promoting positivist research at the expense of the interpretivist or constructionist traditions (although these are in themselves troublesome concepts), or generated distance between researchers and their data. But what CAQDAS has done, and continues to do, is to facilitate ways of doing research work that are fast, transparent, and auditable, capable of encompassing ever larger numbers of respondents and artefacts

derived from both physical and virtual spaces, in formats that are convenient, easy to share and straightforward to work with, revise, re-code and overlay with new and different analytical frameworks. Within CAQDAS, “there are *not* two quite distinct quantitative and qualitative ways of thinking” (Cooper et al., 2012: 8, original emphasis): what there are, are ways of doing research that are thorough, robust, and trustworthy, capable of generating conclusions – and perhaps theories – that will stand up to rigorous scrutiny, even if the conclusions remain as points of disagreement.

References

Bathmaker, A-M. (2004) Using NUD*IST. In Opie, C. (ed.) *Doing Educational Research*. London: Sage.

Brent, E. (1984) Qualitative computing: approaches and issues. *Qualitative Sociology* 7(1-2) 34-60.

Bryman, A. (2008) The end of the paradigm wars? In Alasuutari, A., Bickman, L. and Brannen, J. (eds.) *The Sage Handbook of Social Research Methods*. London: Sage.

Campbell, M. and Gregor, F. (2004) *Mapping Social Relations: a primer in doing institutional ethnography*. Lanham: AltaMira.

Coffey, A., Holbrook, B. and Atkinson, P. (1996) Qualitative data analysis: technologies and representations. *Sociological Research Online* 1(1) http://www.socresonline.org.uk/1/1/4.html [date accessed: 4 May 2014].

Cooper, B., Glaesser, J., Gomm, R. and Hammersley, M. (2012) *Challenging the Quantitative-Qualitative Divide: explorations in case-focused causal analysis*. London: Continuum.

Darmody, M. and Byrne, D. (2006) An introduction to computerised analysis of qualitative data. *Irish Educational Studies* 25(1) 121-133.

Davidson, J. and di Gregorio, S. (2011) Qualitative research and technology: in the midst of a revolution. In Denzin, M. and Lincoln, Y. (eds.) *The Sage Handbook of Qualitative Research*. Fourth edition. London: Sage.

Friese, S. (2012) *Qualitative Data Analysis with Atlas-Ti.* London: Sage.

García-Horta, J. and Guerra-Ramos, M. (2009) The use of CAQDAS in educational research: some advantages, limitations and potential risks. *International Journal of Research & Method in Education* 32 (2) 151-165.

Gerson, E. (1984) Qualitative research and the computer. *Qualitative Sociology* 7(1-2) 61-74.

Gibbs, G. (2007) *Analysing Qualitative Data*. London: Sage.

Hammersley, M. (2008) *Questioning Qualitative Inquiry*. London: Sage.

Kelle, U. (1997) Theory building in qualitative research and computer programs for the management of textual data. *Sociological Research Online* 2(2). http://www.socresonline.org.uk/2/2/1.html [date accessed: 4 May 2014].

King, A. (2010) 'Membership matters': applying Membership Categorisation Analysis (MCA) to qualitative data using Computer-Assisted Qualitative Data Analysis (CAQDAS) software. *International Journal of Social Research Methodology* 13(1) 1-16.

Lewins, A. and Silver, C. (2007) *Using Software in Qualitative Research: a step-by-step guide*. London: Sage.

MacMillan, K. and Koenig, T. (2004) The wow factor: preconceptions and expectations for data analysis software in qualitative research. *Social Science Computer Review* 22(2) 179-186.

Mangabeira, W., Lee, R. and Fielding, N. (2004) Computers and qualitative research: adoption, use and representation. *Social Science Computer Review* 22(2) 167-178.

Odena, O. (2013) Using software to tell a trustworthy, convincing and useful story. *International Journal of Social Research Methodology* 16(5) 355-372.

Richards, T. and Richards, L. (1994) Using computers in qualitative research. In Denzin, M. and Lincoln, Y. (eds.) *Handbook of Qualitative Research*. London: Sage.

Roberts, L., Breen, L. and Symes, M. (2013) Teaching computer-assisted qualitative data analysis to a large cohort of undergraduate students. *International Journal of Research and Method in Education* 36(3) 279-294.

Seale, C. (2005) Using computers to analyse qualitative data. In Silverman, D. *Doing Qualitative Research*. Second edition. London: Sage.

Seale, C. and Rivas, C. (2012) Using software to analyse qualitative interviews. In Gubrium, J., Holstein, J., Marvasti, A. and McKinney, K. (eds.) *The Sage Handbook of Interview Research.* Second edition. London: Sage.

Seidel, J. and Kelle, U. (1995) Different functions of coding in the analysis of textual data. In Kelle, U. (ed.) *Computer-Aided Qualitative Data Analysis: theory, methods and practice*. London: Sage.

Sin, C.H. (2008) Teamwork involving qualitative data analysis software: striking a balance between research ideals and pragmatics. *Social Science Computer Review* 26(3) 350-358.

Smith, D. (2005) *Institutional Ethnography: a sociology for people.* Lanham: Altamira Press.

Smith, B. and Hesse-Biber, S. (1996) Users' experiences with qualitative data analysis software: neither Frankenstein's monster nor muse. *Social Science Computer Review* 14(4) 423-432.

Spradley, J. (1980) *Participant Observation*. London: Holt, Rinehart and Winston.

Tesch, R. (1988) The qualitative researcher and the computer. *Qualitative Studies in Education* 1(2) 179-183.

Thomas, G. and James, D. (2006) Reinventing grounded theory: some questions about theory, ground and discovery. *British Educational Research Journal* 32(6) 767-795.

Tight, M. (2004) Research into higher education: an a-theoretical community of practice? *Higher Education Research and Development* 23(4) 395-411.

Trowler, P. (2012) Wicked issues in situating theory on close-up research. *Higher Education Research and Development* 31(3) 273-284.

Tummons, J. (2012) Theoretical trajectories within communities of practice in higher education research. *Higher Education Research and Development* 31(3) 299-310.

Tummons, J. (2010) Institutional ethnography and actor-network theory: a framework for researching the assessment of trainee teachers. *Ethnography and Education* 5(3) 345-357.

Van Hoven, B. and Poelman, A. (2003) Using computers for qualitative data analysis: an example using NUD.IST. *Journal of Geography in Higher Education* 27(1) 113-120.

Weber, R. (1984) Computer-aided content analysis: a short primer. *Qualitative Sociology* 7(1-2) 126-147.

Weitzman, E. (2000) Software and qualitative research. In Denzin, M. and Lincoln, Y. (eds.) *Handbook of Qualitative Research.* Second edition. London: Sage.

Willis, J. and Jost, M. (1999) Computers and qualitative research. *Computers in the schools* 15(3/4) 21-52.