Learning and qualitative data analysis with information technology: the role of exploration

Graham R Gibbs

A commentary submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Huddersfield
September 2015
Abstract
Over the past 20 years there have been rapid developments in IT to create software that supports both learning and qualitative research. This thesis examines the design and use of that software, and argues that the exploratory approach in both learning and analysis produces superior outcomes. As such, the exploratory approach is seen as one that is particularly well supported by the software.

A range of learning software and objects is discussed: Correlation Explorer, coMentor, learning websites, reusable learning objects, open educational resources, and videos. These are successive attempts by myself, and others, to develop software and other objects that support high quality learning. They do this in a variety of ways: by creating learning tools that promote exploration, by encouraging online collaboration and sharing, and by providing materials that can be used in a range of learning contexts. Some of the problems of their use are discussed, such as mistaken conceptions, and finding and adapting learning objects.

In a parallel fashion, this thesis argues that the development of software to assist qualitative data analysis has supported a range of analytic approaches. By their very nature these tend to be exploratory – the thesis argues that the core of qualitative analysis involves exploration of the data. The new analytic tools the software offers afford especially good support to exploratory analyses. These tools include text searching, code hierarchies, code queries, and the use of charts and diagrams.

Keywords
Exploration, learning, qualitative data analysis (QDA), computer assisted qualitative data analysis (CAQDAS), information technology (IT), virtual learning environment (VLE), open educational resources (OERs), IT in learning, relationship of learning and research.

Commentary
The research presented in this thesis draws upon and combines three separate areas of scholarship: teaching and learning the social sciences, the qualitative analysis of data and the use of information technology (IT) in the social sciences. What this synopsis will do, in addition to highlighting my contributions to these areas, is to suggest that there are some underlying themes in my work that bring these apparently disparate areas together and demonstrate some of the deeper similarities and synergies between them. The core of this approach is that learning and qualitative analysis share many cognitive processes. Recognising these commonalities is helpful not only in understanding some of the challenges in promoting high quality learning and undertaking reliable and insightful qualitative analysis, but can also suggest ways in which new technology of both learning and teaching and of computer assisted qualitative data analysis (CAQDAS) can contribute.

The key common theme, that is central in much of my published work, is that of exploration. Exploratory learning is a well-established approach in the literature on learning and teaching, and I also believe that exploration is a fruitful way of approaching qualitative data analysis. Exploration of this kind is both made possible by CAQDAS programs and can produce novel
analytic insights. Indeed, the subtitle of my first book, *Qualitative Data Analysis* was *Explorations with NVivo*.

The publications included in this thesis represent work undertaken over a period of 20 years, and, as one might expect with anything involving IT, the issues and contexts, in both education and in qualitative research, have changed rapidly during this time. When I started work developing learning tools and the coMentor project in the 1990s, although the Internet and email were well established, the World Wide Web and the first visual browsers were quite recent innovations. So, along with others, my work on using the Web for teaching and learning was breaking new ground. The pedagogic approach I adopted saw the Web not as something like a large bank of knowledge to be worked through in a programmed fashion but rather as a way of making available to learners a set of tools with which they could explore the factual and conceptual landscape. The field developed quickly along the lines we investigated and soon Virtual Learning Environments (VLEs) like coMentor became the norm in universities. Educational developments pushed on in parallel and it was recognised that digitisation and the use of IT enabled the construction of learning materials in a flexible way so that reusable learning objects (RLOs) could be created, used and re-used in a range of pedagogic contexts. VLEs and RLOs raised similar questions about the kind of pedagogy they best suited. In my view, the flexibility and openness of these approaches mean that exploratory approaches to learning are both possible and advantageous.

In the last decade RLOs have morphed into Open Educational Resources (OERs) of which videos (now in digital format and commonly found online) are a key subtype. The move to openness has been particularly important. It recognises that educational institutions have a wider responsibility than just addressing the needs of their own students. That also means there is a much wider range of potential students and educational needs that can be catered for, such as, adult learners and life-long learners. Of course exploratory learning may not suit all such students. Those new to a topic may require a lot of guidance or scaffolding to help them overcome the conceptual hurdles they face. But many learners already have a basic understanding and have established their own learning approaches. The success of one form of open education, the MOOC (Massive Open Online Course), suggests that exploratory learning is useful in open education. The majority of students taking MOOCs are established and experienced learners, and the major learning approach used in MOOCs, constructivism by means of collaboration, encourages learners to take an exploratory approach.

In my work I have argued that the developments of educational technology, and especially IT and the Internet, have made it possible to adopt different pedagogies. This can be seen most visibly in contemporary distance learning. In the past, distance learning tended to follow a strict, programmed approach to learning which was felt necessary in the absence of any face-to-face interaction and student collaboration. Now, with the technology of VLEs, online videos, and online collaboration, much more constructivist and exploratory approaches to learning are both possible and productive.

The developments of CAQDAS programs have been just as rapid over the same period. Undertaking a qualitative analysis without software was, and still is, a complex, onerous and painstaking affair. Whilst the software does not remove the need for deep and insightful thinking about the data, it does provide valuable assistance with many of the more routine tasks qualitative analysts have to undertake. As in the case of educational technology, the use of IT in qualitative analysis was once left to specialists and those willing to tackle what were then somewhat unfriendly software interfaces. I soon realised that the kind of exploratory
approaches to learning that I had been promoting in learning technology could also be applied to qualitative analysis. Not only was it the case that good qualitative analysis required the analyst to avoid immediate and superficial interpretations, but the very software tools that were now being offered by the CAQDAS programs afforded support for more exploratory types of analysis. Traditional forms of pen and paper analysis followed rather programmatic forms of thematic analysis and relied on the analyst’s deep familiarity with the data to do any kind of analysis beyond a summary of the major themes. The software followed this thematic coding approach to begin with but its swift development meant two things. A wider range of tools was added that automated and rapidly undertook analytic activities that would have taken humans hundreds of hours to do. Second, the speed at which the software did this meant that a try-and-see, exploratory approach was now much easier to undertake. In my books on analysis I have demonstrated how the software is able to support this exploratory approach. Although the centrality of coding in CAQDAS programs has often led to accusations that it is only suitable for thematic analysis, I showed in these books that the development of additional tools and functions in the software meant that a range of different analytic approaches could be supported by CAQDAS programs.

**Exploration**

Traditionally the explorer is trying to find or discover new things or at least to offer some insight and understanding of things with which we are unfamiliar. The explorer sets off to fill in the blanks in the map, or at least the blanks on his or her map\(^1\). The explorer does not know what to expect; they have to ‘wander around’, physically or metaphorically, to find the new. Exploratory research does the same. In this case the researcher may have little idea what to expect and must be open to and look for a variety of phenomena. There are precious few directions for the explorer to follow – or for the researcher few tools, questions, scales, step-by-step instructions or proformas. So the explorer must look in all places and maintain an open mind about what is to be described, only then can they come back with knowledge that is ‘new’ or at least new to them and to their audience who they can tell about these new things.

The explorer is often an outsider with respect to what is explored. This has two implications, help from locals is often key and the new knowledge has to be presented in a way that is acceptable to the audience. Past explorers of distant lands and, for that matter, social anthropologists of more recent centuries commonly relied on local assistants or informants – key insiders – to help them reach the right places and people and ask the right questions. But these informants were not necessarily experts, or at least not in the way the explorers needed in order to construct the new knowledge needed by their audience. That required a combining together of an account of the new phenomena with a set of theories and questions arising from the culture the explorer shared with his or her audience. In this sense, the new knowledge involved a fusing of horizons – to use Gadamer’s term (Gadamer, 1975). The pre-existing views of the explorer and their culture may be a problem for the explorers if that means they have preconceptions of what they might find, what is there, what they might see or even what they might look for. Hence the need for them to stay open to the new and for

---

\(^1\) Of course the notion of exploration brings with it all kinds of gendered and imperialistic connotations. Those peoples who were explored were often those about to be exploited by imperialist powers. Whilst those individuals explored (often in the guise of scientific or medical progress) were often the weak or different or female and the explorers were often high status men. My use of the term here sidesteps these historical and social issues and simply uses the idea of exploration in a logical sense without discussing its wide historical and sociological setting.
local assistance. Thus exploration in research might enable us to understand the limits and boundaries of our current thinking or help us use current thinking in different contexts or to bring existing frameworks into comparison with the new.

To summarise, exploration involves

- Filling in the blanks. We do not know what we will find
- Being an outsider to the phenomena we are exploring
- Creating new knowledge or understanding
- Not following step-by-step guides or methods
- Remaining open to the new
- Guarding against preconceptions
- Creating an account of the new that fuses the new with existing understanding.

Exploratory research is well established and is a key justification for much qualitative research. Typically it is used when we do not have enough prior understanding to use more formal, empiricist or pre-existing tools. A key outcome is the discovery of new phenomena or new theory. But it is my view, and I will return to this later in this discussion, that the exploratory approach is a very valuable method in both qualitative data analysis as well as in the acquisition of that data. Remaining open, guarding against preconceptions, being an outsider (at least metaphorically) and not following step-by-step guides or methods are key aspects of good quality data analysis. Much advice about how to do qualitative data analysis is concerned with how to maintain openness and guard against preconceptions.

Exploratory approaches to learning are just as well established (Freitas & Neumann, 2009; Ifenthaler et al., 2012; Kolb, 1984). Exploratory learning can be defined as an approach to teaching and learning that encourages learners to examine and investigate new material in order to discover relationships between existing background knowledge and unfamiliar content and concepts. Instead of working through precisely sequenced training materials, the learner investigates material on his or her own initiative and without a lot of guidance from a tutor, often in pursuit of a task, real or artificial. Exploratory learning can be seen as a variety of constructivist pedagogy, where the learner does not simply ‘import’ or ‘take in’ the knowledge (the transmission model) but needs to interact with the tool or resource – to use or explore it and in this way construct new understanding (Rogoff et al., 1996).

Ideas of exploratory learning were taken up soon after the development of interactive computing in the 1970s and 80s by researchers who recognised the popularity and engaging aspects of computer games and suggested that the exploratory nature of game playing could be used in designing software. Thus, learning how to use new software should be intrinsically motivating. Its functions and its content should be revealed incrementally and the system should be at least capable of being used to some degree without formal training (Malone, 1982; Carroll, 1982; Schneiderman, 1983).

**The role of exploration and practice in learning**

My early work focussed on the use of IT in learning and specifically on learning social research methods and methodology. **Work 1** reports on the evaluation of a software tool to assist students in learning about correlation. I wrote the software on a Macintosh and with a colleague, we tested it out with a number of students, using verbal protocols to record how they were using it. The tool was designed to help students understand what a scatterplot
showed and how changes in the number of data points and their values had an impact on the value of Pearson’s r and its statistical significance. This was done by allowing them to interact, in real time, with a scatterplot and change the number of data points showing and their position (and hence their values) so they could observe the effect of this on the statistics.

Three main conclusions were drawn from this work:

i) There was a strong difference in approach shown by a sequential analysis of the subjects' activities. This could be interpreted in terms of Marton’s distinction of Holistic and Atomistic learning (Marton, 1988).

ii) There was clear evidence that subjects were learning by thinking as well as by doing despite the absence of a specific task and the exploratory nature of the program (Carroll and Mack, 1984). Typically the verbal protocols showed that a serendipitous “doing” was followed by significant examples of “learning by thinking”.

iii) On the basis of their verbal protocols, subjects were learning significantly about the nature of correlation and scatterplots.

However, we also noticed one of the downsides of exploratory learning, namely the ease with which mistaken conceptions could be maintained. We found some learners who had mistaken ideas about correlation but who seemed to be using the software in reasonable ways. Exploratory learning may be seen as just ‘playing with a thing to get quick results’. For example, on getting a new camera a user may just play with it by pressing all the buttons to see what happens. This may be helpful as a way of learning to use the camera but only because the camera designers have designed the menu system in a systematic and easy-to-use way and because the user (learner) already has a good idea of the functionality of the camera and of what it can be used for (its affordances) (Gibson, 1986). On the other hand, if these things cannot be assumed – particularly the latter, approaching things this way may limit learning. Pre-existing ideas about the object’s functionality may be wrong or partial. Learning is then restricted to new applications of what is already understood. Moreover, areas that require the development and use of a new conceptual framework will not be explored. Even if the area is explored, learning does not take place because the conceptual framework needed is absent. What is needed in this situation is either a way to provide the conceptual framework (and do this in a way that is accessible and not intimidating) or a way of encouraging the kind of exploration that will support the spontaneous generation of a new conceptual scheme.

Collaboration
One way to address this need is to promote learning in a collaborative setting, where other students may bring different and perhaps correct conceptions into the learning process. I explored this possibility in my work with the coMentor system as reported in Work 10 and Work 2.

coMentor, developed by the team I led at Huddersfield (a research assistant and a computer programmer), funded by a Jisc grant, was an early iteration of what is now known as a virtual learning environment. Its origins lay in an existing multi-user dungeon (MUD) gaming system. In Work 10 this development, the design philosophy and the opportunities it offered are discussed. coMentor involved the reprogramming of the MUD to serve up web pages, but at the same time it preserved two key aspects of the gaming system, the ability to converse and collaborate (via text messaging) and the metaphors of place and objects. This meant that
 unlike many web-based systems, coMentor could promote certain kinds of playful and exploratory activities that would promote learning.

A key issue was support for collaboration. The software supported both synchronous and asynchronous, text-based discussions along with the depositing and use of a range of (mainly text-based) resources. The software was initially evaluated with a group of students studying methodology and philosophy of the social sciences where debating and interacting would be particularly appropriate. The outcomes of this evaluation are discussed in detail in Work 2. The topic they were learning about is one that students, notoriously, find difficult and challenging. There is plenty of scope for mistaken conceptions and even complete lack of understanding of concepts. Students were helped by seeing others struggling with the same issues they found difficult and even sometimes found erroneous, partial or garbled discussions of the concepts helpful. Such discussions acted as a kind of scaffolding enabling them to acquire a more accurate and logical understanding of the issues. But, as I discovered, the scope for learning was not limited to acquiring accurate conceptions or ideas about methodology and philosophy. The evaluation reported in Work 2 included the use of the ASSIST Learning Styles Inventory. What this showed was that students not only learned the content and concepts of methodology and philosophy, but they also learned how to learn. That is to say, they learned that a deep learning style, encompassing a rounded and abstract understanding of the concepts was the most appropriate for this subject matter and that, in contrast, surface styles were both inadequate for mastering the ideas and failed to challenge the mistaken concepts that underlay many of the difficulties students encountered in their learning. Crucially, students were not taught this explicitly; rather they discovered it through their use of the system and the set of activities that were required to undertake in their collaborative learning.

Learning tools

In most cases, students do not simply discover and then internalise new ideas on their own. As Kolb (1984) has pointed out in his discussion of the experiential learning cycle, after exploration and the discovery of some new ideas the learner needs to reflect on the experience so that what they think they know can be deconstructed and ambiguities and mistakes addressed and more general applications can be recognised. In the case of exploration, learning is often done alone, so there is no one to prompt this reflection. As discussed above, one way to address this need for reflection is through collaboration. But even then the reflection might be partial. There remains a role for the teacher either to prompt this reflection directly or to organise the learning activities such that reflection has to occur. A colleague and I discussed this in Work 9. Our argument was that computer assisted learning should not attempt to replace the teacher, but rather should be designed to work alongside the teacher in ways that can enhance student learning. Following Chute (1995), we suggested that what was needed was software tools for learning that could be used alongside the contribution of teachers and fellow learners. Since we wrote this chapter, several ideas have been developed that support the approach, including ideas of blended learning and reusable learning objects. We discussed three examples of this kind of software, one being the Correlation Explorer program discussed above. However, another was Inspiration, a concept-mapping program. The software has no educational content, the task of the learner is to use the concept map to explore and express their understanding of the ideas they are learning. This program and others like it have been and still are widely used by students, but interestingly for the later discussion here, Inspiration has also been used to support the qualitative analysis of data (Weitzman and Miles, 1995).
Reusable Learning Objects

I brought together my interests in learning, software use, and qualitative methods in a project, funded by the ESRC, to support the needs of those learning to use computer assisted qualitative analysis (CAQDAS) programs. A major output of this project was the development of a website containing a range of learning materials about CAQDAS: OnlineQDA.hud.ac.uk. To guide the design of this website and to ensure that the learning materials that it contained addressed real needs, we undertook a range of investigations. One aspect of this was a set of qualitative interviews and focus groups with those who were learning or who had recently learnt to use the software. This was reported in Work 12, an interim report delivered to the ESRC. Among other things we found a range of misconceptions about what the software could do and hence how it could be used. The most common issue was that before using the software in earnest, learners thought that, to at least some extent, it would do some of the analysis. In part this was a result of wrongly applying their knowledge of how quantitative software works to the CAQDAS programs. Many were very disappointed when they discovered that the software does not do this. In one case a respondent reported that their supervisor told them they must have missed something. In some cases this mistaken view led learners to overuse or wrongly use some of the functions in the software. One example was the over-reliance on the autocoding facility in NVivo, in some cases without any other kinds of coding being done. This was an example of a more general tendency in such learners to use the software in a very limited way. Most often this meant they took a very straightforward approach to coding which they carried this out in the software. But it also meant that they only used a few of the many functions offered by the software.

A second issue we found was the common fear of learners that they might be misusing the software. Such learners were unwilling to experiment and explore with the software partly because they thought the software might produce inaccurate results and partly because they were unsure about the general requirement of qualitative analysis. Actually, ironically, the learners who were most confident about experimenting with the software were those with prior experience with quantitative software, and they were often learners who had the mistaken and over-confident views about what the software could do.

An important result of this ‘needs analysis’ was the recognition that many people who were learning to use CAQDAS software had very limited knowledge about how to do qualitative analysis. They were often learning to do both at the same time, and in many cases using a very attenuated qualitative analytic approach. There was a need for good and widely available materials on the details of how to undertake qualitative analysis, so, after being awarded my National Teaching Fellowship, I successfully bid for a project aimed at creating a range of reusable learning objects (RLOs) on the topic of qualitative data analysis (REQUALLO – a play on the idea of reusable qualitative learning objects). The intention of the project was to create a range of multimedia materials resulting from work with a small number of qualitative researchers to show in detail how they undertook their analysis. These materials constituted the RLOs. The idea of RLOs was that they could be used by themselves by students to learn about the topic they addressed or used by teachers as part of a wider course on the topic. The term reusable was used to indicate that they were not attached to only one teaching context or one curriculum.
In the end five PhD students from the social sciences (two in social psychology, two in business and one in education) were included and a range of audio, video and textual materials were created to show their analysis. They were all near the end of their thesis work (and in two cases were interviewed just after they had gained their PhD). These were all made available in the website, onlineqda.hud.ac.uk that we had established for the ESRC funded project just discussed. The report on this project, Work 14, gave evidence of just how helpful researchers found not only the textual material, but especially the video material. During this project I had started a YouTube channel and all the videos were loaded onto that website as it gives viewers a much more satisfactory access from all kinds of devices than we could offer from the OnlineQDA website. This meant that we had information not only about the website usage, but also about the usage of the videos.

It was the intention from the start of the project to make these reusable learning materials open to all educational users. While the project was being worked on there was a rapid growth in educational interest in the idea of openness and particularly the development of open educational resources (OERs). OERs cover a wide range of resources, from whole courses and curricula to small-scale topics that might constitute just one session in a course or even just one skill or aspect of knowledge in a session. The RLOs we had produced on this project came somewhere in the middle of this spectrum and it was a fairly easy process to make them OERs. They were, of course, already educational resources and we simply indicated their openness by releasing them under one of the Creative Common licences.

The growth in the production of OERs brought with it another set of issues. From the REQUALALLO project we had good data on how the resources were used by learners and teachers, but the wide availability across the Internet of OERs raised the issue of how relevant resources could be found and selected by users. This question was addressed by a Jisc/HEA funded project I undertook with colleagues that looked at this question in the case of OERs about social research methods. We reported on this in Work 6. We found a mismatch between the level of detail about the resources that those creating them attached to the resources and the information about provenance, quality and relevance that those searching for resources needed. Crucially our focus groups and test groups showed us that teachers looking for materials wanted to be able to judge quite rapidly whether what they had found was relevant and usable. Their model for this was the ability to rapidly find appropriate materials in other contexts that Google gave them. Many of the repositories of OERs, such as the UK repository, Jorum were found wanting at the time of the research (2011).

There have been several attempts to try to address these issues along the lines we suggested in the paper. So, for example, the Jorum repository undertook a substantial and well-designed revision of its tools for finding resources, though sadly, the development of Jorum is no longer being supported by its funder, Jisc. There have also been a few attempts to create websites where users of the resources can review and comment on them so that those looking for resources can make much more focussed searches. But these have had little success. However, in the area of video at least, one approach to making OERs available and easily found stands out. The YouTube website (owned by Google) has proved not only to be an excellent repository of educational resources (and where many of those creating OERs deposit their videos) but it has produced an excellent search engine that enables students and teachers to find appropriate resources reasonably quickly. The availability of tools like

---

2 I was the academic lead and worked with a research assistant, a member of the LTSN Centre for Learning & Teaching in Sociology, Anthropology & Politics (C-SAP) and a programmer.
Google and the YouTube search engine along with a multitude of OERs on the world wide web has meant that exploratory learning and independent learning are more feasible strategies for both those already following a course of study and for the occasional or informal learner. Alongside the material produced for REQUALLO I have close to 100 other videos on social research methods on my YouTube channel, and feedback from users indicates, both how useful viewers find the videos and that in many cases these are students already taking formal courses who are taking exploratory approaches to find other material to help them learn. One reason why YouTube has been so successful as a repository for video OERs is that the materials are immediately available to the learners and are produced in engaging and authoritative ways. In contrast, repositories for textual materials, websites, online learning tools, lesson plans, curricula etc. have been less successful and are less often used by learners. The popularity of my channel (over 1 million views since 2010) also suggests that I am making the material accessible and useful for learners who are extending their studies in exploratory and independent ways. Though I also know that many of the videos are being used in formal courses.

**Exploration and QDA**

The exploratory approach in qualitative research is well established. In fact in many cases qualitative research is seen, perhaps unfairly and somewhat dismissively, as merely exploratory. However, I think there is much to be gained from explicitly encouraging the exploratory frame of mind in qualitative research and especially in qualitative analysis. The comparison with exploratory learning suggests that the processing and analysis of qualitative data is similar to the way in which learners try to explore and learn about new topics. This is nothing new. In ethnographic research there has long been the acceptance that the researcher is like an explorer in a novel society. He/she needs to learn about that society and how they do things and how they see the world. One possible measure that this has been done successfully is being able to pass off as a member of that society. The researcher knows how to act appropriately in the variety of social situations that present themselves. In some ways this is rather like the way that Wittgenstein discusses the criterion of successfully learning any new concept: knowing how to go on (Wittgenstein, 1953, no. 154). But the similarities between research and learning go even further when we look at how exploratory learning takes place.

**Research and learning**

In exploratory learning insight comes from reflective practice. Just experiencing a new practical situation is not enough. Learning from practice requires the subsequent reflection on the experience so that the learner can gain from it. This is because the experience is not transparent. Lots of things happen when the learner is exploring the practice and the learner may not be clear about what happened and what made it happen. So reflection is needed so that the experience is examined to look for causality, critical interventions and generalizability to other situations. These are tantamount to a set of research questions about the experience. To gain proper learning from the experience, the learner needs to ask these questions. In this sense, learning means having to ask research questions.

So the parallel with qualitative research is strong. In analysing qualitative data we have to ask similar questions and in so doing, explore the data. However, a key difference is that there can be no teacher or textbook to guide the reflection and the questions about the data. This
makes the importance of exploratory approaches to research even greater. Exploration can be done without a textbook to guide or without a set of set stages to follow. Much qualitative research, especially, is exploration because there is no guide, no theory and hypothesis to test. The researcher has to construct, *ab initio*, the causes, the critical interventions and the general understanding of the data.

The usefulness of stressing the role of exploration in qualitative analysis can also be seen in the contrast of observation and exploration. Of course, all research must involve some form of observation, but in most cases you can only observe what you already know about or at least what you already understand. Observation means looking at what you know is there, what you are familiar with. It involves recording, counting, measuring etc. but cannot involve any new categories, concepts. If new categories or concepts are encountered, they appear as mysteries, things that cannot be properly described. By contrast, exploration means going and observing where one is a stranger, where one has no prior experience. It involves taking new looks, seeing in new ways. Seeing what has not been seen before. Even exploring the familiar means treating it as unfamiliar and seeing things in ways that have not been seen before. Therefore exploration is creative. New understandings and new categories and concepts are discovered or created. This is exactly what is demanded of essentially inductive qualitative research.

But this demand to see the new and the novel takes effort. When examining qualitative data it is all too easy to remain at a level of understanding and interpretation that is already familiar and that reflects the existing conceptual view of the data. One way of seeing this is to consider the kinds of biases that qualitative analysts are prey to. In the first edition of his book *Real World Research*, Robson suggests that the human analyst suffers several deficiencies (Robson, 1993, pp. 374-5). These reflect the kinds of biases that people experience when trying to interpret the world and trying to make decisions about it. So, for example, the analyst is overly influenced by first impressions. Data they interpret first may come to dominate their overall understanding of the material. To make this worse, there is a tendency to over-emphasise positive instances. Information which confirms any hypothesis already held by the researcher is emphasised and counter instances are played down. Many of these biases are familiar to psychologists studying decision-making. Kahneman, who has been central in much of the work about decision-making suggests that it takes real effort to overcome these biases (Kahneman, 2011). It takes effort to be creative and effort to overcome the straightjacket of prior concepts and understandings. Offering interpretations and explanations that avoid making this effort is a real temptation for the qualitative researcher. For example, qualitative analysis commonly means trying to understand what people have said; to understand their words. In one sense, this is an easy task. We speak the language (usually) and we can easily understand what they have said. But we understand in a way that is familiar and commonplace. This is because the application of prior concepts and heuristics is almost automatic and takes little effort. What the qualitative analyst needs to do is to see new things in the data, to see things in a different, novel way. To do this takes effort and it takes effort to avoid the danger of regressing back to existing views and concepts; the way things have always been seen. Kahneman suggests this can be seen as the difference between thinking fast and thinking slow. For the analyst, thinking fast would be the immediate understanding of the text. Thinking slow would be the more reflective, nuanced, insightful and novel understanding that comes from much more intense and focussed analysis.
What is needed is a frame of mind that is open to the new and that is willing to expend the effort needed to reveal novel interpretations of the data. In my view, exploring, doing the things that explorers do and adopting the explorer’s outlook is one such frame of mind. This is the approach to qualitative analysis that I took in my book on the topic (Works 7a, 7b, 7c, 7d, 7e, 7f). In fact, I went a stage further in suggesting that this kind of exploratory approach to analysis could be especially supported by the use of CAQDAS software.

In Work 7a (Chapter 5) I discuss the basics of coding in qualitative analysis. This is a common activity in QDA and a key feature of CAQDAS programs. A key point that I discuss is the distinction between *a priori* and inductive coding. Many writers on QDA and particularly many discussing the grounded theory method (GTM) take the view that inductive coding, that is developing novel thematic ideas about the data expressed in codes is central to QDA (Glaser & Strauss, 1967; Corbin and Strauss, 2015; Birks & Mills, 2015). As I have suggested in the discussion above, it is, however, very hard to avoid the pull of existing frameworks and as I suggest in this work, social scientists are particularly well qualified to draw on lots of existing, *a priori* theories and explanations. So far, these issues are common to both manual analysis and CAQDAS. But what I try to emphasise in this chapter and in later ones is the ways in which the software can support a wide range of analytics traditions and also how its use can promote a flexible and exploratory approach to analysis. The software makes it easy to modify the analytic framework (represented principally by the coding) that one builds up. So, for example, I explain how to use the node hierarchy (nodes are the term used in NVivo for codes) to develop the analytic scheme, initially by grouping nodes and clarifying concepts, and then in later chapters (Works 7c and 7d) I explain how the affordances of the software can be used to discover new ideas and record them in the software. An important activity early in the analysis is keeping memos about ideas, hunches and other analytic insights which can be recorded in memos. Of course this is standard practice in GTM, but in Work 7b I explain how the software allows not only the recording of the memos, but also their linking in meaningful ways to other parts of the data set and the developing analysis. This allows a flexible analysis and also supports the use of memos as first drafts of final write-ups.

**The development of new analytic techniques using CAQDAS programs**

An important tool that NVivo and other CAQDAS programs include is a text search tool. As I explain in Work 7c this tool is much more powerful than simply searching a document for a word. The fact that someone has used certain words does not necessarily tell the analyst exactly what they meant, and of course, the absence of certain words does not mean that certain ideas were not expressed – respondents may have expressed themselves in different ways. But I suggest here that there is a variety of ways in which word searching can be used in both confirmatory and exploratory ways. I extended this argument in a later book (Work 8a). Not only does the text search tool allow searching for words with the same root and words with similar meaning but it also finds all the occurrences in the selected texts. That means it can be used in two ways, either to support the quality of the analysis by providing evidence that the coding is not missing important issues, or as a way of exploring the data to find new themes and to establish new codes. Works 7c and 8a explain how this can be done with the software. This means that the analytic approach possible when using software can be quite different from that which relies solely on the manual approach. I developed this argument further in Works 7d and 7f. In Work 7d I demonstrate a range of software tools such as building and modifying the node hierarchy and the range of different node queries that enable analytic strategies which are hard to do manually or well nigh impossible to
achieve without the software. Yet, as I argue in Work 7e, this does not mean that traditional approaches to QDA cannot be used with the software. To the contrary, I look at three approaches, the structured, GTM and the narrative, which are well supported by the software. It has been argued that the software designs and functions have been significantly influenced by the kind of analysis promoted by GTM and with the centrality of coding and memo writing in most CAQDAS programs this is clear. But in this work I also demonstrate how key aspects of other approaches such as the structured and narrative approaches can be supported in the software. I developed the suggestions about how the kinds of functions supported by the software could assist narrative analysis in Work 8b. The view I take in this work is that the kinds of thematic and coding approaches that the software supports well can be used alongside the discursive and narrative analyses that are commonly found in the analysis of stories and accounts. In the least controversial sense this is because thematic and coding approaches can be used as a kind of pre-processing of the data to select subsets upon which a more constructivist method can then be used. But I also suggest some ways in which thematic and constructivist approaches can be used together in complementary ways.

Another way in which I believe CAQDAS software is developing the methods of QDA is discussed in Work 7f. Here I looked at the way that the software supported visualisation of the data and I suggested a range of ways in which tabular and diagrammatic tools could be used to develop the analysis. The use of tables is not new. It was written about by Miles and Huberman (1984) and similar tabular approaches were developed by the National Centre for Social Research in their framework approach (Ritchie & Lewis, 2003). But what I demonstrate in this chapter is how the software tools can be used to implement the range of table uses that have been suggested with the additional benefit that using the software means that there is always an easy and rapid linkage from the tables constructed back to the original data so that interpretations the analyst has made can be properly evidenced. As for the visual tools available in NVivo I suggested a much more exploratory role for them. I presented examples of the ways that the charting tools in the software could be used during the analysis to express developing models about data. But I saw these charts being useful not only to capture the developing analytic ideas of the researcher but also for use in the final presentation of the analysis in reports and publications. As in Work 7e, I argued both that the software offered new ways in which researchers could develop their analysis of the data and that it allowed the implementation of many techniques derived from manual analyses. This brings the great advantage that the new techniques and the traditional can be used to complement each other both in exploring the data and in supporting the quality of the analysis.

I have written about the more general use and development of CAQDAS in two other works: Work 3 and Work 11. The earlier of these, Work 3, written with two colleagues³, accompanied a special edition of the journal FQS which we edited. This article, much cited, gave an overview of the available CAQDAS programs and discussed their use. We also made some suggestions about how the programs and their use might develop in the future. This focussed on the range of data types that could then be used or were about to be introduced in the programs and indeed since we wrote the paper all the leading programs now handle, images, sound, pdfs, video and survey data as well as text and some have facilities for working with social media data, and GIS data.

However, in that paper we suggested that some forms of analysis (such as discourse analysis,

³ Susanne Friese & Wilma C. Mangabeira
were less well supported by the software. That was true at the time and it certainly reflected research practice then, as well as one or two key papers that rejected the use of CAQDAS in DA (MacMillan, 2005; Ten Have, 2007). However, as I have argued above, although CAQDAS is not routinely used for DA there are now many researchers doing DA who do find it useful to use CAQDAS in their work. This is a point I argue in Work 11, the second state of the art review I have written.

The rapid expansion of the use of social media and the development of tools in CAQDAS programs for dealing with such data was something that we underestimated in Work 3. In fact the growth of interest in what has become known as Big Data and the expanding use and interest in mixed methods is something that the CAQDAS program writers are addressing. There are some basic tools for dealing with large data sets using quantitative tools in many of the programs (such as the text searching tools I discussed in Work 8a.). In Work 15, a report I wrote for the Higher Education Academy, I examined whether such quantitative techniques within qualitative methods were being taught at undergraduate and postgraduate levels and if so what good practice could be found in teachers’ practice. Although there is interest in such approaches in research papers and there are some tools in the CAQDAS programs, I found relatively little use and teaching of these techniques at undergraduate level. It seems to me that social science teachers are missing a trick here. The analysis of social media data (as part of Big Data) is proving very attractive to government, public bodies and commercial enterprises, and social scientists, and especially qualitative researchers, have the skills needed to interrogate and interpret this data in useful, reliable and interesting ways. Social media data are also a source of unobtrusive, real world, natural data that it is relatively easy for students to get hold of and free of many of the ethical and risk constraints they may experience collecting other kinds of data. The reasons I found that teachers did not include such approaches in their teaching were varied but centred around lack of skills amongst the teachers in doing this kind of research and a rather conservative division of methods teaching into quantitative and qualitative elements that allowed little time for new techniques.

The development of research and learning resources

The availability of a growing range of qualitative data sets on the Internet and the development of functions in the CAQDAS programs to support their use is something I have investigated in the Victoria Climbié Inquiry Data Corpus (VCIDC) project. This was a project I worked on with colleagues at the University of Huddersfield and was concerned with promoting the use of a large data set produced by the Public Inquiry into the death of Victoria Climbié (whose death was a result of a child protection failure) which was put online by the UK government (Parton, 2004; Hall, 2003).

From the start it was our intention to process this very large data set, the VCIDC in a way that would make it more accessible to other users for research and for teaching. It was for this reason that we undertook a Delphi exercise which took us further than just coding the data in response to researcher’s categories and brought in a wide range of expertise to establish the key themes which could be used to organise the data. (Work 4)

On the face of it, this work takes a very different approach from the more direct analytic exploration. Using the Delphi exercise to establish a coding frame for use on the data meant that we adopted a strict, a priori coding of the corpus. Whilst this was clearly not an exploratory approach to analysis, it did have several advantages. First, it meant that the coding of the very large data set could be done by the single assistant we had in the time.
available, and could be easily quality checked by other team members. Second, it meant that the major themes used called on the multidisciplinary expertise of stakeholders in the child protection services and reflected issues of interest to social work colleagues. As we envisaged the processed data set being used by teachers and researchers in social work, this was clearly an advantage. Third, and above all, though, it has to be recognised that we were not undertaking a pure analysis of the data (where an inductive and exploratory approach might have revealed new insights) but rather attempting to organise the data, via a system of coding or tagging, in a way that would make it more accessible to other researchers and teachers. This is a key point, which I will return to later. Not every form of coding data (or tagging it) implies a thematic analysis of the data. Our tagging of the data did not represent a novel analysis of the data but rather it facilitated the development of further, inductive analyses. We certainly expected additional analyses in researchers who used the VCIDC. They might use our coding for initial selection to identify a subset of material on which to carry out further work, using both inductive coding and other techniques, such as discourse analysis.

In a later paper, Work 5, we argued that there were a growing number of such online and openly available data sets being made accessible, usually as a result of government inquiries across different countries and we outlined some suggestions of the kinds of ways they might be analysed. These included the use of some of the functions of CAQDAS programs discussed above and in Works 7c and 8a, such as lexical searching. We discussed how this could be done alongside thematic coding, narrative analysis and discourse analysis. In doing this we addressed some of the reservations that some discourse analysts have about the use of CAQDAS programs, as discussed above and in Works 3 and 11.

The nature of coding

The use of CAQDAS programs and of online data sets raises the issue of the nature of coding. In a traditional sense, when researchers were still analysing texts using pen and paper, coding meant thematic coding. The researcher identified passages of text as being relevant to a particular theme. The passage was then marked in some way, i.e. coded with the name of that theme. The development and naming of themes is a key part of the analytic process and inevitably was confused with the act of coding the text. But the development of CAQDAS programs (and the development of social networking sites and software where tagging is a common process) means that the distinction between the logical activity of coding – tagging text with a name – can be more easily separated from the analytic activity of developing themes.

There are two consequences of this distinction. First, the activity of coding, seen as mere tagging, can be used in a variety of different ways. This becomes useful when using software to assist the analysis. I discussed several of these uses in Work 7. For example, a code (node in NVivo) can be used as a placeholder. No text is coded at the code, but it has a logical place, perhaps as a parent code in the hierarchy, or as an example of a type (e.g. as a result of dimensionalising a concept in a GTM analysis) that might exist but where no evidence for it in the data has yet been found. Such simple tagging can also be used to support the thinking process involved in constructing the thematic categorisation or organisation of the data. Thus it might identify data which has raised issues (negative cases, alternative outcomes, critical cases) but for which the researcher does not yet have or does not want to commit to a thematic name. Alternatively coding can be used as a simple highlighting device or an aide memoire to highlight the use of particular words or phrases or to mark the specific way that language has been used. This kind of use can range from tentative identification of hunches
(a reminder to look at something again later) to the marking of specific discursive or narrative moves in the text. In all the cases, the coding of text does not necessarily mean that the marked passage is representative of an analytic theme. The consequence is that the software allows a much wider range of use of the coding process than thematic coding and again this range of use reinforces the point made earlier that CAQDAS use has had an impact on QDA practice. A simple example of this would be the need when coding by hand to code substantial chunks of text so that when the similarly coded text was retrieved there was enough text to convey the meaning. In contrast, when doing CAQDAS it is practical to code single words either because it is a simple matter to retrieve all the coding whilst continuing to see the context or because retrievals can be done so that adjacent text is retrieved too.

The second result of distinguishing the logic of coding from the activity of thematic analysis is that we can differentiate coding as part of analysis from coding of texts for the purposes of reanalysis. The latter is what we were trying to do in the Climbie project. Actually the two approaches to coding do overlap in application, but they differ in their motivation. Coding done for analysis stresses interpretation, the development of novel ideas and understanding and going beyond the participant’s views of what is happening and why. But the coding we undertook in developing the VCIDC did not need to do this. We could focus on categorising by biographical data (of the witnesses) and codes that reflected both already known theory and the descriptive practice of those professionals involved in child protection.

I discussed another of the differences in the types of coding in Work 13. This was a commentary on a paper that failed to make a clear distinction between cross-case or case-by-case coding and the kind of flexible coding undertaken in QDA. Case-by-case coding is commonly found in the coding of open-ended questions in surveys. The coding and categorisation of responses is done with an identical coding scheme across all cases and each case is only coded once for each category. Of course thematic coding in QDA does much more than this, allowing us, for example, to code many passages in the same case and to ask questions about how these coded passages related to passages coded other ways in the same case or in other cases. Moreover, as I have argued above, qualitative coding, especially using CAQDAS can be done in an even wider range of ways and for a wider range of purposes than just thematic coding. I have found the confusion between case-based coding and QDA coding very common, in experienced researchers as well as in novices. Perhaps there is some justification for adopting a different term to describe the different kind of coding used in QDA and ceteris paribus in CAQDAS. A variety of terms are used currently: tagging, coding, categorizing, thematic coding, indexing, classifying, and creating a typology or taxonomy. For some writers these mean slightly different things, but they all seem to be versions of coding in the logical sense. To distinguish the activity done in QDA from that done in quantitative analysis with survey data we need a different term. A candidate I favour is ‘tagging’. That seems to be the most neutral term, logically, and suggests just the idea of attaching a label or tag to some data, without necessarily categorising or classifying it.

**Contribution to knowledge**

*IT in learning*

A lot of my work has looked at the creation of learning materials for resource-based learning. One of my earlier activities was the development and testing of software which encouraged exploration in learning: Correlation Explorer. This research showed that most students could learn from such exploratory, computer-based activity and it formed part of the wide range of investigations of how students can be supported in their learning by the use of computers.
(initially running on a PC, but increasingly now online). With a colleague, I argued that such an approach and such use of software should act as a learning tool and be used alongside tutor guided lessons and activities. Such an approach has now become common under the name of blended learning and the philosophy of the creation of such tools is encompassed by the idea of reusable learning objects. Of course much of the student engagement with such ‘tools’ is now online rather than with software just on the local computer. A principal contribution to this move was the work I did on various projects which developed coMentor, a fore-runner of modern virtual learning environments (VLEs) such as Moodle and Blackboard. The research work I published on this system and the promotional work I did not only showed how such online environments could be used to support and promote learning, especially in discursive and continually debated subjects such as philosophy and the social sciences, but also demonstrated that exploration and collaboration were key elements in learners’ development of a range of conceptual learning and learning skills. Along with other teams working on similar projects, the work we did was crucial for the development and adoption of VLEs by UK universities.

I then turned my attention to learning about QDA and CAQDAS, again using IT-based and especially online resources. The material the OnlineQDA project produced was carefully designed to reflect the learning challenges and learning needs we found in our analysis of how postgraduates and researchers learned QDA and CAQDAS. A particular problem we identified was that many of those learning to use CAQDAS programs knew little about how to do QDA. We addressed this in the project and I returned to this need in a later project, REQUALLO, designed to create a range of multimedia resources to assist those learning QDA. The importance of addressing this need in those learning CAQDAS programs can be seen in the support materials now produced by the CAQDAS software companies which often address issue of qualitative analysis technique as well as software functionality.

CAQDAS

In my book on NVivo and my book on QDA I was keen not only to demonstrate how exploratory approaches to analysis were a good way of directing research but also to show how they were directly supported by the functions in the software. In so doing I contributed to the understanding of CAQDAS software use in two ways. First, particularly in these two books, but also in other publications, I demonstrated how the use of software could have and was having an effect on the way researchers prosecuted their data analysis. This development is not transformational, but it has allowed researchers to address issues of analytic quality, the discovery of novel patterns in their data and crucially the integration of qualitative and quantitative approaches to analysis. Second, the book on NVivo was one of the first publications to show how standard qualitative approaches, such as GTM and narrative analysis could be carried out using the software. At the time, this was something that the software manuals did not address and which was relatively little discussed in the research literature.

The use of software to assist with analysis opens up other avenues of research and I have investigated these, especially in my research with colleagues on the VCIDC project. This not only identified a rich seam of data relevant to the social sciences but also suggested some ways in which the data sets might be made available to researchers and learners and some ways in which they might be analysed, especially using CAQDAS programs. These ideas, and those concerned with linking qualitative and quantitative data have now become much more relevant with the growing interest in and use of so-called ‘big data’. There is a rich seam of investigation examining how qualitative research approaches can be used to shed
light on the mainly quantitative results produced by the analysis of big data. This is an area I am continuing to research.

**Relationship of learning and research**

One issue which I have argued for in this commentary, but is not explicit in my publications, is the close similarity of certain learning approaches to the activities of much QDA. I have argued above that we can think of the researcher analysing qualitative data doing similar things and encountering similar challenges to the learner exploring a new learning landscape. There is a key consequence of this parallel. In part because of the effort needed to discover, understand and interpret the new there is a tendency for the analyst to rely on existing heuristics to produce a quick, but superficial understanding: one which may miss the novel, the unusual and the original in their interpretation. Just as learners may need a certain degree of scaffolding to help them explore new areas of learning, so the researcher needs some ways of working to force a deeper understanding and interpretation of the data. This I believe explains the importance of the kinds of advice found in many texts on QDA, such as ‘waving the red flag’, looking for the critical case and investigating the missing cases. Each promotes what, to paraphrase Kahneman, we might call ‘slow analysis’.

(10175 words)
Published works.

All single authored except if marked with icon 🧑‍⚕️

Refereed Journal Articles

Work 1
1995 “Learning abstract concepts with computers. The experience of users learning about correlation” Psychology Teaching Review 4 (2) pp. 110-120. ISSN 0965-948X (with David Robinson) 🧑‍⚕️

Work 2

Work 3

Work 4

Work 5

Work 6
2012 “Obstacles to creating and finding OERs: A case study of social science research methods” Journal of Interactive Media in Education, Special Edition on OERs, March 2012 = Gruszczynska, A., Gibbs, G.R. and Brent, I. 🧑‍⚕️
Books

**Work 7a**
Ch 4 Coding.

**Work 7b**
Ch 5 Memos and attributes.

**Work 7c**

**Work 7d**
Ch 7 Developing an analytic scheme.

**Work 7e**
Ch 8 Three analytics styles.

**Work 7f**
Ch 9 Visualising the data.

2007 *Analysing Qualitative Data*. London: Sage. Part of the Qualitative Research Kit, ed. U. Flick. ISBN 9780761949749 (Trans to Polish, Spanish, Portuguese)

**Work 8a**
Ch 9 on Searching…

**Work 8b**

Book chapters

**Work 9**

**Work 10**
2001 ‘Modifying Multi-user Discussion Systemss to Support Text-based Virtual Learning Environments on the Web – the coMentor Experience’, Chapter in Reza Hazemi,

Work 11

Other reports

Work 12

Work 13
2012 Gibbs, G. R. “Different Approaches to Coding”. Sociological Methodology, 42 (1). pp. 82-84. ISSN 0081-1750 (2012)

Work 14

Work 15
References


