Count: Developing STEM skills in qualitative research methods teaching and learning

Graham R. Gibbs

HEA Social Sciences strategic project 2012 -13

Teaching research methods in the Social Sciences

In June 2012 HEA Social Sciences held its first learning and teaching summit, which focused on teaching research methods in the Social Sciences (Further details of this summit, including papers and presentations, can be accessed via: http://blogs.heacademy.ac.uk/social-sciences/2012/09/10/teaching-research-methods/

In December 2012 we commissioned 11 projects that were designed to explore further the issues identified at the summit. All the outputs from these projects are available via: http://bit.ly/1jZe0Ft.

The role of assessment in research methods teaching: A literature review

Anesa Hosein (University of Surrey) and Namrata Rao (Liverpool Hope University)

Count: Developing STEM skills in qualitative research methods teaching and learning

Graham R. Gibbs (University of Huddersfield)

Creative research methods in a college-based higher education setting

Alex Kendal (Birmingham City University) and Helen Perkins (Solihull College)

Developing applied research skills through collaboration in extra-academic contexts

Andrew Kirton, Peter Campbell, Louise Hardwick (University of Liverpool)

Developing innovative support structures for students undertaking small-scale research projects in work settings

Paula Hamilton, Peter Gossman and Karen Southern (Glyndŵr University)

Developing peer assessment in postgraduate research methods training

Hilary Burgess, Joan Smith and Phil Wood, assisted by Maria Scalise (University of Leicester)

Engaging students in quantitative research methods: An evaluation of assessment for learning strategies on an undergraduate social research methods module

Ciaran Acton and Bernadette McCreight (University of Ulster)

Innovation in the assessment of social science research methods

Luke Sloan (Cardiff University)

LSE100: An innovative, multi-disciplinary approach to assessing research methods learning

Jonathan Leape (London School of Economics)

Mapping, understanding and supporting research teaching within college-based higher education (HE) networks

Claire Gray, Rebecca Turner, Carolyn Petersen, Carole Sutton and Julie Swain (Plymouth University)

Northern Ireland by numbers: new open educational resources for teaching quantitative methods

Emma Calvert and Paula Devine (Queen's University Belfast)

Contents

Contact details

For further information about this project contact Graham Gibbs (g.r.gibbs@hud.ac.uk)

1. Overview

In natural science this is the age of big data. In medicine, cosmology and sub-atomic physics the availability of enormous computing power means that new understanding can be extracted from extremely large data sets. The same is true in the social sciences. The internet means that there are web pages, web sites and a range of social media such as Facebook and Twitter and support and interest groups where data sets are both large and openly accessible. The rapid digitisation of media also often adds to this with areas like radio, TV, printed media, gaming, YouTube and other video often openly available on the internet. The open availability of these data makes their use in social research relatively easy and means that they can be used in teaching without many of the ethical issues involved in getting data from other sources. However, although digitised, many of these data are not numeric, rather textual, aural or visual, and need to be analysed with qualitative data analysis techniques. Such approaches are usually detailed, intensive and laborious and not easy or efficient to apply to large data sets. However, two recent developments in social data analysis offer some support for such approaches. They are the development and adoption of Computer Assisted Qualitative Data AnalysiS (CAQDAS)[1] programs and numeric approaches to analysis that can be summarised under the general title 'data mining' and that are now often included in CAQDAS programs. So there are clear opportunities for making such data sets accessible to students and for teaching the use of software and data mining techniques with which they can analyse the data.

Despite the potential, there is little current reference to software or numeric techniques in both the use of qualitative methods in social research and, especially, the teaching of qualitative methods, particularly at undergraduate level. The one exception to this is the discussion of the use of computer-assisted qualitative data analysis (CAQDAS) programs in several textbooks aimed at undergraduate users (e.g. Bryman 2008; Gibbs 2007). However, despite the opportunities there is no strong evidence that this translates into widespread use of the programs in undergraduate teaching in the UK. Indeed in a web-based survey of teachers of qualitative methods that I carried out in 2011 for a previous Higher Education Academy-funded project (REQUALLO), I found only 6% of departments that replied to the survey used CAQDAS at undergraduate level, although much higher proportions used it at postgraduate level. At the moment, at undergraduate level, there is a distinct imbalance in software use between qualitative and quantitative research methods teaching. Almost all undergraduate courses on quantitative methods will, at some point, cover the use of statistics and will require students to use software (such as SPSS).

There are two factors that might begin to challenge this comparatively low level of software usage in qualitative methods: the increasing use of CAQDAS in research and the growth of interest in mixed methods approaches. In the case of the former, there is evidence for a strong growth in the use of CAQDAS in social research. A recent bibliographic search I undertook of all social science journal papers using any of a range of qualitative methods and using CAQDAS (Figure 1) shows an increasing number of papers mentioning CAQDAS from 1983 to 2011 (Gibbs 2013).

There is also a growth in the popularity of mixed methods designs, not only in the number of published papers using such an approach, but also in the number of PhD students undertaking such projects. To support this, in the last few years, many of the major publishers of CAQDAS programs have included functionality in their software that integrates quantitative data with the qualitative analysis. This includes a range of data mining, keyword in context searches and cluster analysis techniques as well as the ability to import quantitative case attribute data that can be used in the analysis. All this is integrated with the basic thematic coding system the software supports.

[1] See http://onlineqda.hud.ac.uk/ for further details of CAQDAS

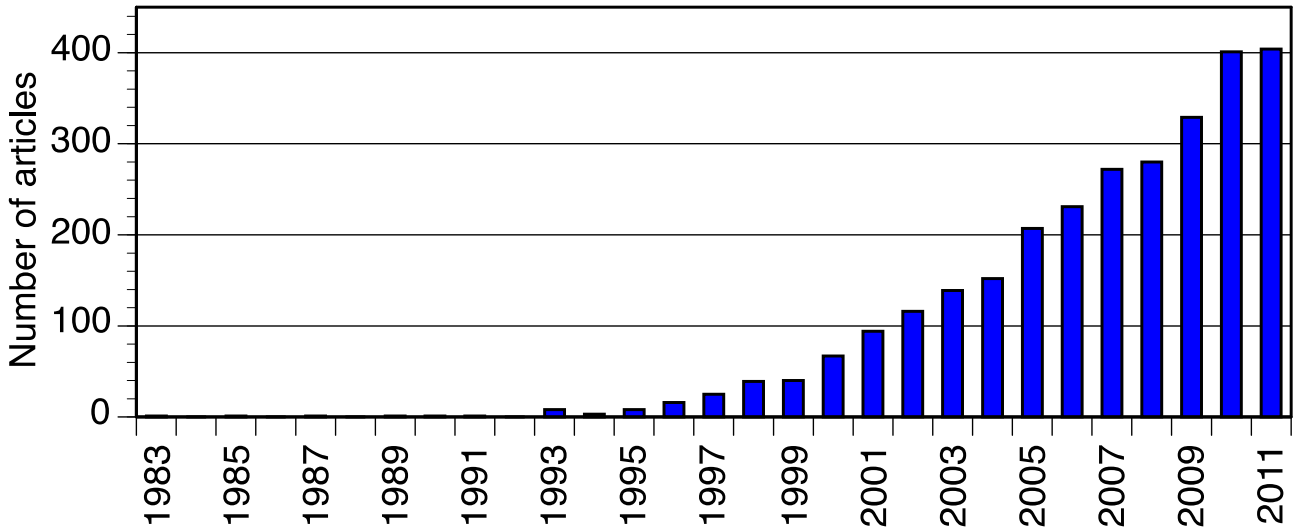

Figure 1: The number of refereed papers published using qualitative methods that used CAQDAS, 1983–2011 (Original to the author)

There are, however, significant barriers to the increased use of number and software in undergraduate qualitative methods teaching. In my survey of teachers carried out for this project I asked respondents why the software was not used at undergraduate level. Common reasons given included no time (17%), lack of teachers' skills in the software (13%), no funding for site licences or local support for software (12%) and, most commonly, that it would take too long to teach (24%). However, significantly, given the antagonism of some qualitative researchers towards software use, only 3% said that the software was not used in teaching because it does not support the methodology or the theoretical approach they used.

This situation is changing. Many UK universities now hold site licences for the software (just as they do for statistical software). About a year ago, QSR, the publisher of the market-leading NVivo software, told me that 74 UK universities held site licences and there are several who hold similar licences for competing programs.

Nevertheless several barriers remain:

- there is a lack of skills in both CAQDAS use and mixed methods among qualitative research methods teachers at undergraduate level;
- there are few good exemplars, usable at undergraduate level, of how software and numeric skills can enhance qualitative analysis;
- lack of open educational resources (OERs) suitable for undergraduate use that tackle this issue.

2. Aims and objectives

This project therefore proposed:

1. to identify further barriers to inclusion of such skills topics in undergraduate qualitative methods teaching;
2. to discover and summarise good practice and to develop guidelines for such teaching content that can be used at undergraduate level;
3. to discover and publicise, where possible, good OERs and data sets that can be used to teach the use of quantitative data in qualitative analysis and the use of CAQDAS.

3. Activities

To address these objectives, two main activities were carried out: a survey of UK teachers of qualitative research methods; and interviews with selected respondents as well as some other key educators (such as book authors) and researchers.

A survey of current undergraduate teaching activity to identify good practice was carried out using Bristol Online Survey between 15 April and 12 May 2013. A major challenge here was to identify the target population of qualitative methods teachers. A variety of approaches were taken. Higher Education Academy (HEA) email distribution lists were used to send invitations to complete the questionnaire. These included a wide range of disciplines where qualitative research methods might be taught at undergraduate level, such as the broad range of the social sciences, Psychology, Education, and Management and Business. This was complemented by sending invitations to the UK CAQDAS Networking Project contacts and to a number of UK-focused JiscMail and other email lists for qualitative researchers.

By the close of the survey there were 115 completed responses of which 90% British, 4% other EU. There were just two responses from the USA. The focus on UK teachers seems to have been successful. While 115 is not a large number it seems to be in line with the numbers achieved by other, recent surveys of UK teaching staff. As is the case with online surveys, there is no easy way to calculate response rate. Indeed, one clear problem here for the calculation of response rate for any survey of this target population is that there is no sampling frame, that is to say, there is no definitive list of UK academics who teach qualitative research methods. However, one survey question asked respondents for their discipline. The largest group of respondents was made up of sociologists – 19 respondents. Assuming there are about 160 institutions most of whom have departments containing sociologists and assuming two qualitative sociology teachers per institution means a response rate of about 6% was achieved.

Another key limitation of online surveys is that it is hard to estimate the degree of bias in the sample obtained. Surveys with a small response rate tend, *ceteris paribus*, to be more biased than those with large response rates. In the case of this survey one might also infer a possible bias towards the views and practice of the topic of the survey. The invitation to complete the questionnaire made reference to the use of software and to data mining techniques, so those replying would have been those interested in these approaches and, more significantly, not ideologically opposed to these approaches. Some caution therefore needs to be exercised in interpreting the survey results as typical of all teachers of qualitative methods. However, the results can still be used to indicate the existence of certain practices and views, although it is not possible to say accurately how frequent these are among all teachers.

The questionnaire used asked respondents about their teaching of qualitative methods to their students, their (and their colleagues') use of CAQDAS, mixed methods and data mining techniques at both undergraduate and postgraduate level and their views about possible barriers to such teaching or the wider use of such teaching. Respondents were also asked whether they had any relevant resources they used in their teaching which they were willing to share and whether they were willing to be interviewed in depth about their teaching.

A small, selected number of survey respondents were interviewed in depth along with a number of experts in the software and data mining techniques and book authors dealing with such approaches. All the interviewed survey respondents were from the UK, but the experts included some from the UK as well as some from the USA, Canada and Australia. Most were interviewed using Skype and the interview was recorded. A smaller number were interviewed by telephone. No physical, face-to-face interviews were carried out. The focus of the interviews was not to build a picture of typical practice, but rather to identify the range of teaching practice along with some key examples and to identify the overall potential for software and data mining techniques at undergraduate level.

A major difficulty experienced was that very few of the survey respondents were using CAQDAS at undergraduate level and even fewer used data mining techniques or had resources they were willing to share. So while examples of good practice were discovered, very few examples of good educational resources including data sets were identified.

4. Outcomes

4.1 Survey results

As expected, teachers of qualitative research methods were found in a range of disciplines. The principal ones are shown in Table 1. There were also responses from a range of other disciplines, from Computing to Sports Studies.

Discipline	%
Business	11
Education	15
Health related (e.g. Medicine, Nursing, Pharmacology)	16
Management	9
Psychology	13
Sociology	17

Table 1: Respondent's discipline – main disciplines

Qualitative research is characterised by a wide range of different and somewhat competing approaches. More surprising was how many of these were being taught at undergraduate level. Respondents mentioned 42 different methods and most mentioned they were teaching or had taught several of them at undergraduate level. The most common were interviews and case studies, which over two thirds of respondents mentioned. Other commonly taught methods, mentioned by over half the respondents, were mixed methods, participant observation, grounded theory and ethnography. A substantial minority (from 35% to 11%) mentioned narrative, action research, thematic analysis, discourse analysis, document use, comparative analysis, life history, biographical approaches, participatory research, phenomenology, feminist, and use of video and conversation analysis.

There are also differences by discipline; see Figure 2. Case study methods were most popular in Business, Management and Criminology. Ethnography was most commonly taught in Sociology, health-related areas and Criminology. Feminist methods were rarely mentioned except in Sociology. Grounded theory was most commonly taught in health-related areas. Participant observation was rare in Business Studies but very commonly taught in Sociology. Similarly, phenomenology was commonly taught in health-related areas but rare in other disciplines. This adds to the overall picture of diversity. There were no approaches taught by all respondents and there were very few that were taught by all respondents from the same discipline.

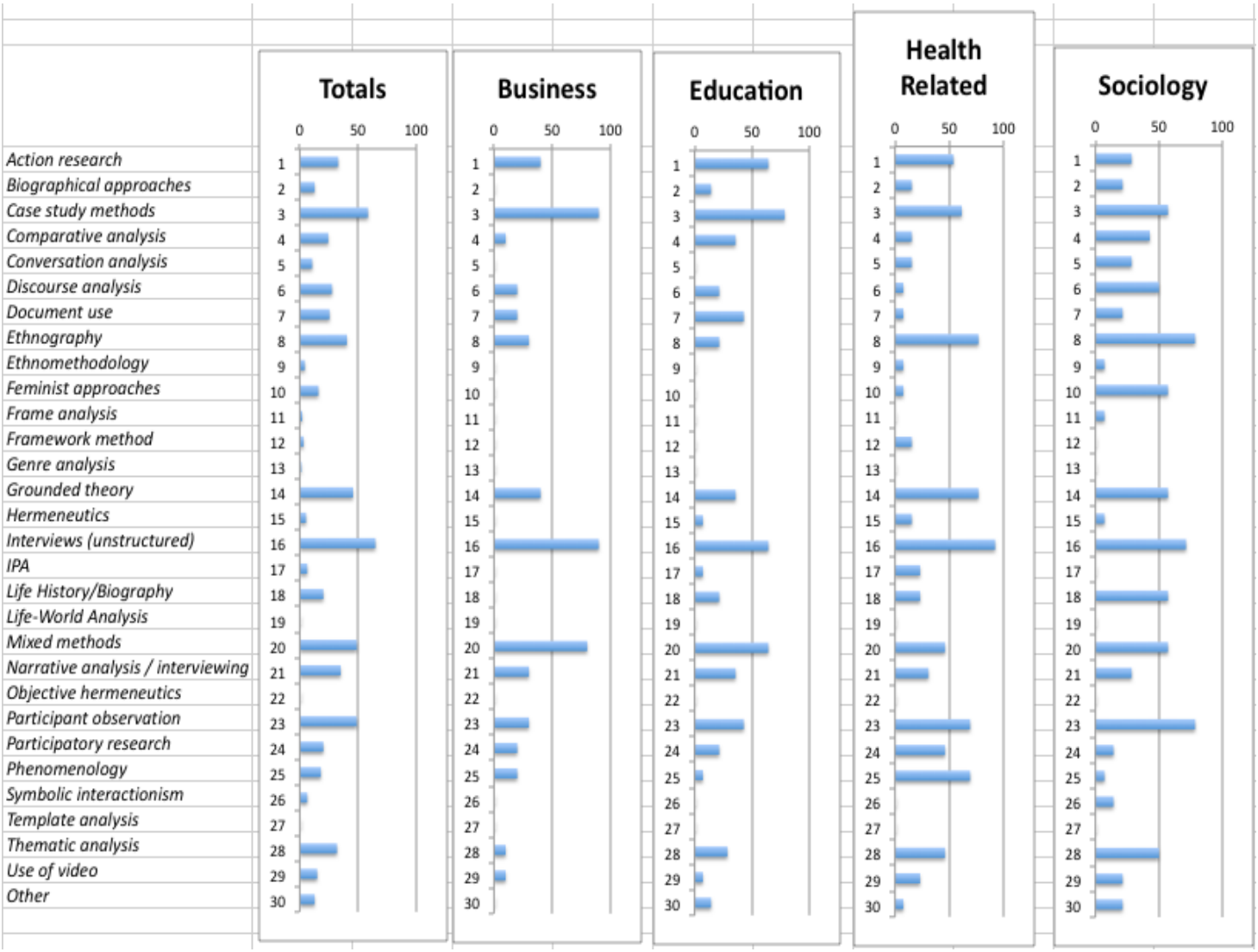

Figure 2: Percentages of respondents teaching different approaches and by discipline for Business, Education, health related and Sociology

Respondents were asked at which level (or year) they, or their colleagues in their discipline, taught qualitative research. See Table 2. Relatively few started teaching qualitative methods to first-year undergraduates, but nearly three-quarters taught it to undergraduates in the middle year(s) of their degree. In around half of cases this was complemented by further teaching in the final year, in addition to qualitative methods being used in the final-year dissertation. The use of software in undergraduate teaching was much less common. Only 13% of respondents reported its teaching in the middle years of degrees and 12% in the final year. 60% reported that they did not teach CAQDAS to undergraduates (there were some non-responses too).

	Qualitative research % per year	CAQDAS %
Year 1	22	3
Year 2 (and year 3 in Scotland)	72	13
Final year	48	12
Undergraduate dissertation	42	-
Other	13	-
Not taught to undergraduates	-	60

Table 2: Year or level at which qualitative research or software use was taught

Of those respondents who reported any CAQDAS teaching at all, the overwhelming majority were using NVivo at both undergraduate and postgraduate levels. Table 3 shows the actual numbers reporting different software use in their qualitative teaching. Sixty-three respondents reported that their institution had a site licence for specific programs.

	Program	n
Undergraduate use	NVivo	21
	Atlas.ti	2
	HyperResearch	1
Postgraduate use	NVivo	46
	Atlas.ti	9
	MAXQDA	2
	Wordsmith	1
	EndNote	1
	HyperResearch	1
	SPSS	3
Site licence	NVivo	63
	Atlas.ti	7
	MAXQDA	2
	Wordsmith	1

Table 3: Number of respondents reporting the use of specific software in qualitative methods teaching at undergraduate and postgraduate levels

In all, 67 of the 115 respondents reported that software was not taught as part of undergraduate qualitative methods teaching in their departments. They were asked to give reasons why the software was not taught in their department. Table 4 shows the percentage suggesting different reasons.

The most common reasons were time and the space it would take in the teaching programme. Comments made by respondents about this included:

- 'Hardly any time to spend on qual. in syllabus as it is, so core teaching focuses on qual. fundamentals.'
- 'Time constraints do not allow attention to statistical analyses.'
- 'Too little time to cover qualitative methods in general – there is a five-week lab and that's it.'

But significantly, 40% suggested there was no expertise to support software teaching. Comments made included:

- 'Lack of staff expertise and confidence.'
- 'Limited number of staff have used mixed methods in large projects so limited experience of other than content analysis techniques using basic frequency counts.'
- 'A lack of experienced tutors to support the teaching.'

As became clear from the interviews, the constraint here is not simply that there is not the expertise to run a course, but that all the additional teaching support (such a co-tutors and lab tutors) needed to teach the large classes typically found in qualitative methods was not available.

Reasons for not teaching software	%
No time to use software	49
Would take too long to teach	52
No teaching expertise in using software	40
No access to software	34
Data sets used are too small to warrant software use	34
No local support for software use	25
Software does not support methodologies/theoretical approach used	10
Software not relevant or not needed for the methodologies/theoretical approach used	19
I was not aware such software existed	10

Table 4: Reasons for not teaching software to undergraduates as part of qualitative methods; percentages of the 67 respondents who reported software was not used in their departments

Given the range of qualitative methods and in some cases the vehement ideological differences between those supporting particular approaches it is not surprising that some rejected software use because it was not seen as necessary or relevant to the methods they were teaching. Some expressed this point in their comments:

- 'I see these as significantly different methods. I want my undergrads to understand the ontological differences, before we support them in considering mixed methods.'
- 'Some people object to quantitizing qualitative data.'

Respondents were also asked what they thought were the main barriers, in general, at their institution or at other institutions to the inclusion (or greater inclusion) of CAQDAS and text analysis in the teaching of qualitative methods. Table 5 shows the responses.

Barrier	%
Too much additional learning for undergraduates	50
Lack of space in the timetable	50
Lack of qualified teachers	42
Lack of experienced tutors to support students	40
Lack of sufficient PC labs with the software	38
The approaches can be taught well enough without software	29
Lack of good learning resources	18
Not relevant to the analytic approaches that are taught	10
Insufficient good data sets available	9

Table 5: Main barriers to the inclusion (or greater inclusion) of CAQDAS and text analysis in the teaching of qualitative methods (at your institution or elsewhere)? (Percentage of all respondents).

Respondents were asked about the specific use of numeric techniques in or alongside qualitative methods and significant numbers (69%) had themselves used the approaches in their own work. See Table 6.

Approach used	%
Basic frequency counts of code use	44
Word frequency counts	35
Keyword in context	23
Co-occurrence analysis	7
Producing scales or typologies from qualitative data	14
Mixed methods approaches	32

Table 6: Percentage of all respondents using quantitative approaches to assist with the qualitative analysis of data or with reporting its results in their own work

Slightly smaller percentages but with a similar range of approaches taught numeric techniques in or alongside qualitative methods to undergraduates. See Table 7.

Approach used	%
Basic frequency counts of code use	36
Word frequency counts	29
Keyword in context	16
Co-occurrence analysis	4
Producing scales or typologies from qualitative data	5
Mixed methods approaches	28

Table 7: Percentage of all respondents teaching undergraduates quantitative approaches to assist with the qualitative analysis of data or with reporting its results

However, respondents also commented about the student fear of numbers:

- 'Generally speaking students don't like language of numbers :-)'

and about the lack of time to take on software teaching as well as qualitative analysis:

- 'They already get three years of quantitative! The qualitative is usually crammed into one or two lectures, so they need to be dedicated purely to qualitative.'

Also, as became clear from the interviews, the mentioning of these approaches in teaching was often not covered in depth and often they were accompanied by warnings about the dangers of using frequencies and counts where samples were not representative.

In the expectation that a number of teaching resources would be discovered, respondents were also asked what kinds of material or media they used in their teaching and where they got these media from. Their answers are shown in Table 8 and Table 9 respectively. The most common media are as expected, but it is interesting that 72% reported using film, video or animation materials in their teaching and that 50% reported getting materials from YouTube.

Material/media	%
PowerPoint slides	100
Recommended texts	98
Reading lists	86
Prepared lecture notes	85
Required reading	73
Film/video/animation	**72**
Case studies/role plays	64
Tutorial/problem sheets	63
Worked examples sheets	48
In-class quizzes/tests	45
Artifacts (as products, models, drawings/designs)	23
Computer-aided learning software/learning technology	21
Task specific software	12
Other ICT	11

Table 8: What type(s) of teaching materials/media do you bring into your teaching with students? (Percentage of all respondents)

Resource	%
YouTube	50
Your libraries' digital resources (such as e-Books)	44
Other courses on your institution's VLE (such as Blackboard)	32
Professional body website	24
HEA website	19
Discipline specific website (such as OnlineQDA.hud.ac.uk)	16
Corporate website	14
Another institution's website/VLE	11
National educational repository (such as JORUM)	8
Open access repository (such as OpenLearn)	8
iTunesU	8
Box of Broadcasts	8
Flickr	4
Other (incl. own developed resources)	3
BUFVC	1
MOOC/opencourseware (such as edShare)	0

Table 9: Where did third-party electronic/online resources used in your teaching come from? (Percentage of all respondents)

In conclusion, the picture drawn by the online survey results is of a group of teachers offering a very wide range of qualitative research methods approaches, though with some clear differences in what was offered by different disciplines. Use of software (CAQDAS or text mining) in teaching was not uncommon but was found much more at postgraduate level than at undergraduate. The main reasons why it was not used more often at undergraduate level were thought to be time and space in the teaching programme. Also some suggested there was a lack of expertise with the software and that in their department there was a lack of support for or availability of the software. Nevertheless, at least 55% thought their institution had a site licence for a CAQDAS program – in the vast majority of cases for NVivo.

4.2 Interviews

The results of the interviews (and other investigations) are reported here as three contrasting case studies and a discussion of the issues that need to be considered in teaching with software and teaching data mining approaches.

The cases studies are:

1. Teaching qualitative methods to second-year Sociology undergraduates at the University of Wolverhampton;
2. Teaching qualitative methods to second-year Marketing undergraduates at the University of Bournemouth;
3. Teaching qualitative methods to third-year Psychology undergraduates at Curtain University, Australia.

In all three cases, students were taught the use of CAQDAS software (NVivo) as part of learning about qualitative data analysis. In one case students were encouraged to use data mining techniques, in another they were discouraged or at least warned of the problems of interpretation.

4.2.1 Case study 1: University of Wolverhampton

A second-year research methods (qualitative and quantitative) course given to 70 students in Sociology. It followed on from a first year course delivered to about 120 students of Criminology and Sociology. The course introduced students to the use of NVivo and by combining qualitative and quantitative elements it touched on mixed methods approaches.

Timetable

The timetable allowed for six weeks of quantitative work followed by the presentation of a report then six weeks of qualitative work with a separate qualitative report. Two of the six weeks were devoted to teaching NVivo. The tutors found that there was too little time to cover what was necessary. Classes were scheduled for two hours a week, but in addition extra workshops, one catch-up hour before the sessions and one lunch time hour after the sessions, were found necessary. Students attended these workshop sessions on demand and attendance increased as term progressed – and the assessment became closer.

Lab work

The two-hour scheduled sessions were held in a large laboratory with about 50 workstations and some students brought their own laptops. Wolverhampton has a site licence for NVivo but it was not installed in every lab on campus. Students were able to put the software on their own computers (and most did), though some experienced problems downloading the software and needed staff with them to learn how to use the software. On the other hand, many were happy, able and confident enough to work on their own.

Two staff covered all the teaching in both the additional workshops and the large lab sessions. Having multiple staff was found to be vital for good teaching in the sessions. The two-hour labs were run as hybrid talk/lecture and hands-on work. Staff would start with a short lecture, explanation or demonstration and then students would follow with some hands-on exercises and this pattern was

repeated several times in a session. As students experienced problems, staff would respond and advise or explain what was needed. In some cases several students would experience the same problem. The lab had three display screens that could be used to show a subgroup of students how to address the problem. However, this proved somewhat disruptive as students had to move to gather around the display screen. This required some pedagogic micromanagement. Teachers had to decide whether to stop the whole class to show a point that many had a problem with or to bring together in a group around a workstation or a projector screen just those needing help.

Interviews

The teaching programme covered data collection via interviews, transcription and analysis using NVivo. Students created their own data by carrying out interviews with a common theme: student experience of social class. They interviewed each other using the research question 'What are second-year students' experience and understanding of social class or what are their perceptions of class?' Students added their own sub-questions to this using the same issue they investigated in the preceding quants section. Students were given the freedom to choose a sub-question to give them ownership and reduce temptation to plagiarise. The teachers made an attempt to group together students with similar sub-questions (e.g. health inequalities and class). Within groups, students were randomly assigned another student to interview and they did this in threes, each taking turns to interview and be interviewed with the third observing the interview and providing some feedback about issues such as gaze/paralanguage to the interviewer. But some did not receive effective feedback from the third person and felt penalised.

They were told they would be assessed not in terms of the quality of the interview or recording, but on what they made of the data they had collected. Students preferred to do interviews in class time. Most did interviews in the computer lab. But this had a high noise level so students were encouraged to book separate rooms (e.g. in the library). Between 10 and 15 of these were available. Half used the computer lab and half the library rooms. Interviews were about 10 minutes long (at most 15 minutes). A trial run with students produced very short interviews. Students had difficulties operationalising the concepts, e.g. talking about issues without mentioning the word 'class'.

Transcription

Students used their mobile phones to record the interviews and then did their own transcription. They had to share these transcriptions with others in their groups so there was communal pressure on them to complete the transcription before the next session. Student participants checked their own transcript as interviewee. But there was limited accuracy-checking, due to a lack of time. Transcripts included backchannel/paralanguage information. They were used in the data analysis with NVivo. The teachers felt that getting students to create and analyse their own data created a better engagement with the data and better motivation to analyse it. Using a topic that students were familiar with from their other modules and using their own data meant that they became rapidly familiar with it and quickly were able to develop a deeper understanding of the data – something that the teaching staff were keen to encourage.

NVivo use

A general session on qualitative data analysis and coding preceded the NVivo sessions. In the previous year a different teacher had used a content analysis approach. This year students were encouraged to go beyond the face content and code in three stages:

1. code people, things and events;
2. code what people say about those things;
3. synthesise and match coding with theory.

This worked reasonably well but students were still finding it hard to escape from simple descriptive codes, so the teachers are considering next year starting with paper-based grounded theory coding, before using NVivo. Students only used a small subset of the functions in NVivo: setting up codes, coding

and code retrieval. For example, this year students did not use code trees (in order not to overload them). The teachers are considering covering this next year. NVivo coding was done in class. The class was kept quiet and coding done individually. Write-ups included screenshots and evaluation of what they had to do. This year, they compared paper coding with NVivo coding. Some students felt NVivo was not a good use of their time.

There was an interesting contrast in where students experienced the most difficulties. Technically, most students were very good at IT and did not experience difficulties with learning how to use the software. They were good at learning the software and good at working out different ways of doing the same thing. The real challenge for the students is not 'which buttons to press' but 'understanding the data' and 'how to categorise the data'. They found difficult that there is no right or wrong way to code. One student created 55 codes at one point. Students struggled to move away from descriptive coding. Many students used the interview questions as their coding scheme. The teachers did not want to provide *a priori* coding because they wanted to encourage students to think about their coding and analysis.

Integration of quantitative skills with qualitative was limited. Word frequency and tag clouds in NVivo were not used because of time limitations. Mixed methods issues were only touched on. The quantitative report was on the same specialist topic as the qualitative and although this clearly would give students overall more insight into the issues of class they were investigating, the reports were written up as two separate stages and there was little combination of analysis. Students were given a very prescriptive report format, including headings, word length and so on.

Students were told NVivo would be useful in their future employment and to mention it in interviews and CVs.

As students created their own data set, no data set for analysis was required. For the sessions on NVivo, in addition to class demonstrations, some videos from the QSR website (QSR publishes NVivo) were used, though some students thought they were tedious. Textbooks were also used to support students, but the teachers thought that many are too dense and detailed for students. The teacher needed to specify to the students which sections or pages to read.

In addition to the two staff teaching the module there are others in the department who know how to use NVivo. So there is sufficient staffing to cover support for students using NVivo in their final-year project.

4.2.2 Case study 2: University of Bournemouth

The course was a second-year module in research methods for students of marketing and used Malhotra (2012) Marketing research: An applied approach, published by Pearson, as the course text. This was the first time NVivo had been taught to undergraduates and there was a new member of staff who knew the software. All the teaching was done by this member of staff. The teacher stated that they felt others might not 'get excited and show passion'. Very few others in the department can use NVivo, so this is a limitation if students choose to use the software in analysing data from their final-year project.

The module ran over two terms. The first term was a series of lectures on analysis that followed Malhorta's text. The second term was given over to seminars and lab sessions with hands-on experience and the use of NVivo. There were no lectures in the second term. Sixty students were taught in three groups of 20 with one teacher.

Interviews and transcription

Students collected their own data. They all conducted two interviews using a convenience sample of other students on the topic of 'ethical consumption'. They were given seven key questions that had to be addressed by the interviews. So interviews were semi-structured. Students recorded the interviews on their own mobile phones or a few borrowed university digital recorders. Interviews lasted about 25 minutes and students also transcribed their own interviews. Students were required to submit the

interview recordings and the transcriptions on CD. According to the tutor, students complained of the amount of hard work involved (and that there were too many assignments) but at the end of the year they realised how much they had learnt and awarded the tutor a 'You're brilliant' certificate. The tutor took the view that after university, out in the market place, students have to compete: 'You have to be challenged not pampered.'

However, students analysed only two interviews each. How could the use of software be justified with such a small data set? The tutor suggested three justifications:

1. Students are learning for later use in their dissertation, at which point they would have six or seven lengthy interviews and this produces a large amount of data. The organisational advantages of the software are clear.
2. Familiarity with NVivo looks good on job applications.
3. NVivo is used as a data management tool. 'Just like a big library, if you don't put your books in order you can't find them.'

Learning NVivo

Bournemouth has a site licence for NVivo so all students could put the software on their own PCs. NVivo teaching was done in the lab sessions. There were two seminars/labs on NVivo use and students were shown how to use the NVivo help system. The teacher went through the software on screen and then the students practised in the lab. For illustration and practice the teacher used her own unpublished research data based on interviews with residents of a slum in Mumbai. Students were given a class exercise based on this data set. They used these data to look for positive and negative things a visitor would find from a visit to the slum. With these ideas, the students then used NVivo to identify some keywords and did a text search using these keywords. Having learned this technique, students could then use it on their own data.

Figure 3: A Wordle tag cloud

Data mining techniques

All students used tag clouds, a simple data mining tool, as a way of analysing and representing their data. Tag clouds are effectively a visual representation of word frequency. Although NVivo can produce tag clouds this function was not used. Rather students used a separate program, the freely available Google online program Wordle (see Figure 3). Some used tag clouds on the whole interview, others on just the few paragraphs of answer to one of the seven key questions. Interestingly, a recent paper (Uprichard 2013) has shown the use of work clouds in a research context.

Coding

Students used NVivo to help organise their data and used free nodes, which are the simplest coding structure, and not tree nodes, which are more complex. Some created their own codes but many just coded using the seven pre-set questions. They used the text search in NVivo to look for key terms, e.g. 'animal cruelty' and 'child labour', to help them analyse the data and to set up new nodes.

Learning difficulties

In this case study too, students found the software easy to use. Students were happy to use the program's help system and were very confident in their work. They did not need a great deal of technical support, and very few had technical problems. Computers were not seen by the students as challenging. The tutor called them the 'Xbox generation'. In contrast they found interviewing difficult. They lacked confidence in knowing what to do and had to be provided with key questions to guide the interview.

4.2.3 Case study 3: Integrating NVivo into qualitative methods teaching at Curtain University, Australia

The module focused on qualitative methods and was offered in the third year of a four-year undergraduate degree in advanced Psychological Science. The module consisted of ten lecture sessions followed by seven lab sessions on qualitative research plus time spent on project work.
Of the seven labs, three were on NVivo use and the software use was fully integrated into the module and the data analysis the students undertook.

Approximately 150 students were taught by several academic staff who were all experienced with NVivo. A range of standardised teaching materials was prepared and piloted with fourth-year student volunteers. Students undertook their own interviews and then analysed these data with the support of NVivo. In using the software the students moved beyond merely coding the data to the development of categories and the use of conceptual maps.

Schedule

The seven weeks on qualitative analysis were as follows:

1. developing a research plan;
2. interview schedules and information sheets;
3. interviewing;
4. qualitative data analysis;
5. thematic analysis, initial coding (NVivo);
6. thematic analysis, category development (NVivo);
7. thematic analysis, conceptual maps and writing up (NVivo).

Outcomes

Psychology is a discipline dominated by the experimental paradigm where qualitative research is often seen as less legitimate or less valid. The interviewee felt that the teaching of software as part of qualitative analysis had the effect of increasing the perceived legitimacy of qualitative research. It also served to provide an initial training for students who went on to use qualitative methods in their final-year projects and wished to use CAQDAS software. It was also perceived that knowing how to use the software was a marketable skill for the students.

The teachers found that it was very important that those teaching NVivo should also be teaching qualitative research methods. It was not appropriate to use teachers who only knew how to use the software. Proper use of the software required proper understanding of what qualitative analysis involved.

It was also important that students had access to books, manuals and the software outside the scheduled sessions. Having the software on their own PCs was key.

5. Issues and good practice

A number of issues relating to the teaching of qualitative data analysis using software arise out of these case studies and the other interviews carried out. There are examples of good practice that can address these issues.

1. Should students be taught analysis first on paper and then using CAQDAS or should teachers integrate CAQDAS and QDA? Both approaches were found in the interviews. Some teach first on paper then transfer to computer. Others just use software to teach analysis. The key point here is that students should get a proper understanding of the nature of qualitative data analysis. It is possible to learn to use the software and use it in an analysis that is very superficial and where the code development is very limited. Facility with the software is perfectly acceptable, but the actual data analysis is very limited. Paper-based teaching of qualitative data analysis (QDA) is often undertaken to ensure that students get a proper understanding of analysis. But this can be done using just the software. This problem is often manifested as students remaining with very mechanical or descriptive codes and not moving on to develop more analytic or theoretical codes. One way to encourage this transition is to use stages in the research, with appropriate guidance as to what students must do at each stage. For example, students may be asked explicitly to develop a certain number of analytic codes or to identify and replace some descriptive codes with analytic or theoretical ones.

2. Should students be asked to use common/shared coding, i.e. a priori coding, or should they develop their own coding? The former helps students get started and helps them understand the nature of coding. The latter gives students better motivation and better understanding but involves a heavier learning load and commitment on the part of the students. Time is an important issue here. Providing an a priori coding system may be quicker for students to do (and probably easier for teachers to manage). But having used such a coding system, students should be encouraged to create at least some of their own codes.

3. Whether to include use of code trees/hierarchy and organisational/theoretical development. Some do, some don't. Again these more advanced analytic stages come with a heavy learning load for students and need bigger data sets and more work on the students' part. These aspects of analysis can probably be left to more advanced classes in qualitative data analysis and software use.

4. What data should students use in their analysis? A shared, existing data set or should students collect their own data? There is strong view (almost a consensus) that for students to use their own data is better. Students are more engaged and have better contextual understanding of data they have collected themselves. However, such data takes time to collect and the consequence might be that students end up working on very small data sets or very short interviews. Pre-existing, provided data sets are acceptable if they relate to issues that students understand, are familiar with and reflect the material they are learning elsewhere on their course. This can provide engagement and understanding. An advantage of pre-existing data sets is that they can be larger than one student can produce on their own, can be partially analysed or coded and can incorporate interesting and significant material from the point of view of analysis. One compromise here that was not found in the interviews or survey is the use of hybrid data, where there is an existing robust and extensive data set but students are encouraged to add cases to it through additional data collection. This would combine the advantages of pre-existing data sets with the engagement that comes from students collecting their own data.

5 A related question is whether students should be allowed to generate their own research questions or whether they should all have the same research questions and interview schedule (easier to organise). The practice found in the Wolverhampton case study provides a compromise solution to this dilemma. Students addressed a common core of research questions but were encouraged to develop at least one distinct research avenue of their own. With large classes and a limited number of tutors, a more structured and specified approach is probably necessary.

6 Who does the teaching? Often there is just one person or two who do all the teaching and marking and extra support. Although this is a heavy load on staff, it can work well for specific modules. However, it is very limiting for the wider use of software in QDA. For example, if students want to use the software in their final-year project, wider support is needed. This may need to be tackled by staff development to improve teaching-staff skills.

7 Students need to have the software installed on their computers. The simplest way to facilitate this is to get a site licence and to develop simple procedures for transferring the software to students' PCs. One interesting and useful development here is that some software is now available in a free PC version or a free iPad version. A limited but functioning version of QDA Miner is now available and there is a free, but somewhat limited version, of Atlas.ti for iPad. These can be used to start the process of instruction in QDA and software use, with perhaps a limited number of students moving on to the full version for later, more advanced work.

8 It is good for students to interview each other in groups of three (interview pairwise with feedback from third). Students can learn by watching others. This may be harder to organise and require more class time, but it does overcome the major ethical and practical issues of students interviewing strangers off-campus. However, it is possible to separate data collection from data analysis. If students have already learnt how to undertake interviews, there is plenty of data on the internet they may use for analysis.

9 Textbook support for research methods can be intimidating to students just starting. Students who are just starting to learn about research methods need closer and more specific guidance on what to read.

10 Should students code individually or in pairs? Students find coding particularly hard to do at the start. Coding in pairs is a longer process, but it shares the mental task between two and allows or promotes discussion about what they are doing. If time allows, this is a good approach. If not, then a more structured approach to coding is preferable.

11 Some interviewees reported that students were sceptical about the value of learning the use of the software and resented the time it took. One way to address this is to stress the use of software for the analysis of final-year project data sets. But this requires that there is adequate support for final-year project analysis and it requires a certain level and spread of staff expertise.

12 Employability and the use of statements of skills in software use at job interviews and on CVs. For some academics this is a plus (especially if the employment sought is close to the discipline area so the usefulness of software skills is apparent.) But for others it adds no benefit because employers aren't interested in or don't know about the software. Nevertheless, employers do know about statistics software and SPSS and even if the employer hasn't heard of NVivo or other CAQDAS programs, students may at least claim expertise in learning another piece of software. Linked to this is the idea of badging software use, that is, providing students who successfully complete a module that includes software use with some kind of certificate or badge so that they may include it in their CV and show it to potential employers.

13 Using word frequency or tag clouds. Students like it but there is a danger that it promotes superficial analysis and becomes a way of avoiding the more difficult process of qualitative analytic development. Interviewees in some disciplines pointed out that in their area students had a natural bias towards quantitative research methods. It was necessary to ensure that they did not use quantitative approaches as shortcuts to inadequate, descriptive qualitative analyses. Frequency tables and tag clouds are probably best introduced as an initial stage of analysis, used to identify themes or text for closer scrutiny.

14 How to run labs: just hands-on practice or mixed lecture/demonstration and hands on? How to help those with problems: should teachers deal with the whole group or identify subgroups and re-arrange the lab or provide one-to-one assistance? The common experience of all those interviewed is that dealing with individual problems arising during hands-on practice requires several teachers in the lab. More teachers means a one-to-one approach is more feasible. If sufficient tutors are not available one possible solution is for students to work in pairs (helping each other). They are more likely to be able to work out what actions are needed to address their problem or to work out how they can find out the answer to their problem (e.g. in the help files or in the class instructions).

15 Software developers should develop data sets that show partly coded data so students can add their own and then fully coded data to support further analysis. These data sets need to be carefully designed so that students may get interesting results from them.

6. Barriers to the use of software and the inclusion of quantitative approaches in qualitative data analysis

From the survey and the following interviews we can identify the following barriers to software use in qualitative data analysis.

6.2 Lack of time or space in existing courses

This was probably the most frequently expressed reason why software and data mining techniques were not included in modules. Teachers clearly felt that these techniques would impose an overhead on an already pressurised teaching programme. However, the case studies show clearly that using the software is not something that students find especially challenging. For that reason it probably imposes very little additional load on the course. To avoid issues of space and time it is probably best to integrate the software use fully into the module. Data mining techniques such as word frequency and tag clouds are popular with students and can be used as ways of getting students to think about a deeper qualitative analysis.

6.3 Lack of teaching expertise

This still seems to be a barrier to development. There were examples in the interviews of departments with just one person with skills to teach CAQDAS. This remains an issue, however over time there will be more new staff who have used CAQDAS in their own research and are able to teach its use. A separate dimension of this is where there are staff who can teach the software but there are not sufficient teachers to support a number of lab classes or to support software use in final-year projects.

6.4 Lack of support for the software

It is clear that we are moving towards a majority of institutions having site licences for CAQDAS software with NVivo as the clear market leader. This means staff and students can get access to the software and that all accessible labs can have the software installed. This particular barrier seems to be fast coming to an end.

6.5 Software not relevant or not needed for the methodologies/theoretical approach used

In the near future this is likely to continue, but the sheer range of qualitative research methods taught in departments suggests that there are commonly taught approaches in most departments that are supported by CAQDAS software.

6.6 Lack of good learning resources/data sets

Only a few mentioned this as a barrier and in the case of data sets, many of those interviewed preferred students to collect their own data as this improved engagement with the analysis and ensured that students had a basic contextual understanding of the data they were examining. However, this does not rule out the benefits to be gained from the use of hybrid data sets where students analyse data from a pre-existing data set (possibly with some coding already done) but add some data they have collected themselves. The advantage is that a bigger and more developed data set means that a deeper and more refined analysis is possible, while the student still has an investment in the data they have collected, to improve engagement.

7. Available data sets and other teaching materials

7.1 Datasets

The UK Data Archive (formerly known as Qualidata) (http://data-archive.ac.uk/) is a key source of existing qualitative data sets. They are well documented and generally available for teaching use. (Teachers may need to keep a record of the names of students who have used the data.) However, data used needs to be appropriate to the student's discipline, not only to promote their interest and engagement, but also so that they are able to apply their discipline's theoretical ideas in their analysis. Many of these data sets are large and teachers will have to reduce their size for undergraduate use.

The UK Data Service is developing further teaching resources and these can be accessed via http://ukdataservice.ac.uk/use-data/teaching.aspx.

The software companies usually include data sets with their software. In some cases the unanalysed versions can be obtained separately from the software and can be used in other contexts. One disadvantage of these data sets is that they are designed to show what the software can do rather than facilitate interesting and meaningful analysis. The most popular software, NVivo, in its latest version comes with a large, analysed data set on the topic of the development of coastal communities in Australia. At least one interviewee thought this was a relevant data set for her discipline of Environmental Studies. A demo version of NVivo, including the data set, can be obtained from the QSR website: http://www.qsrinternational.com/products.aspx.

7.2 Learning resources

Software companies also produce a range of resources to help students learn to use their software. In the case of NVivo these include a number of videos and documents:

- http://www.qsrinternational.com/support_getting-started.aspx
- http://explore.qsrinternational.com/nvivo-how-to-videos

The videos can also be found on YouTube as can videos made by others. For example, YouTube videos by Deborah Rowe

- http://www.youtube.com/user/DWRowe1/videos?flow=grid&view=0

and Ben Meeham videos on NVivo 9:

- http://onlineqda.hud.ac.uk/Step_by_step_software/NVivo/NVivo9/BenMeeham_video.php

QSR also produces resources to assist teachers: http://explore.qsrinternational.com/nvivo-in-the-classroom. You will need to register to get access to these resources.

Other CAQDAS software publishers have similar resources. I have given details of NVivo resources as this is the by far most popular program in the UK

Wordle, the tag cloud software, can be found here: http://www.wordle.net/

8. Impact

This project was always very speculative. I hoped to find a number of teachers already using CAQDAS and data mining software with educational resources and data sets they were willing to share. In the event, I found very few teachers using these approaches and few of these had resources they were able or willing to share. In most cases the teachers were getting students to collect their own data for use in analysis so no pre-existing data sets were being used.

There are developments in the research field that will eventually filter down to teaching practice. These include the work by Uprichard (2013) already mentioned as well as a chapter 'Keyword analysis' by Seale and Charteris-Black (2010). I believe, however, that a key push for change will come from student demand, when they discover the kinds of data already openly available for social science analysis and want access to the tools that will enable them to analyse them.

I still believe that there is a great potential for both CAQDAS software use in teaching and the use of quantitative data mining techniques, but teaching practice is at an earlier stage than I imagined. I hope that this report will support the development of this important area of teaching research methods in the social sciences.

References

Bryman, A. (2008) *Social research methods* (3rd edn). Oxford: Oxford University Press.

Gibbs, G. R. (2002) *Qualitative data analysis: Explorations with NVivo*. Buckingham: Open University Press.

Gibbs, G. R. (2007) *Analyzing qualitative data*. London: Sage Publications.

Gibbs, G. R. (2012a) Grounded theory, coding and computer-assisted analysis. In S. Becker, A. Bryman and H. Ferguson (eds) *Understanding research for social policy and social work: Themes, methods and approaches* (2nd edn). Bristol: Policy Press, 337-43.

Gibbs, G. R. (2012b) Software and qualitative data analysis. In J. Arthur *et al.* (eds) *Research methods and methodologies in education*. London: SAGE Publications Ltd, 251-8.

Gibbs, G. R. (2013) Using CAQDAS programs. In U. Flick (ed.) *SAGE handbook of qualitative data analysis*. London: SAGE Publications Ltd.

Seale, C. and Charteris-Black, J. (2010) Keyword analysis: A new tool for qualitative research. In Ivy Bourgeault, Roger Dingwall and Ray de Vries (eds) *The SAGE handbook of qualitative methods in health research*. London: SAGE Publications Ltd, 536-55.

Uprichard, E. (2013) Describing description (and keeping causality): The case of academic articles on food and eating. *Sociology*. **47** (2), 368-82.

Contact us

The Higher Education Academy
Innovation Way
York Science Park
Heslington
York
YO10 5BR

+44 (0)1904 717500
enquiries@heacademy.ac.uk

The Higher Education Academy (HEA) is a national body for learning and teaching in higher education. We work with universities and other higher education providers to bring about change in learning and teaching. We do this to improve the experience that students have while they are studying, and to support and develop those who teach them. Our activities focus on rewarding and recognising excellence in teaching, bringing together people and resources to research and share best practice, and by helping to influence, shape and implement policy - locally, nationally, and internationally. The HEA supports staff in higher education throughout their careers, from those who are new to teaching through to senior management. We offer services at a generic learning and teaching level as well as in 28 different disciplines. Through our partnership managers we work directly with HE providers to understand individual circumstances and priorities, and bring together resources to meet them. The HEA has knowledge, experience and expertise in higher education. Our service and product range is broader than any other competitor.

www.heacademy.ac.uk | www.twitter.com/heacademy

To request copies of this report in large print or in a different format, please contact the communications office at the Higher Education Academy: 01904 717500 or pressoffice@heacademy.ac.uk