



Odena, O. (2011) To use or not to use software? An exploration of the use of software for qualitative data analysis. In: British Educational Research Association (BERA) Annual Conference, 6-8 Sep 2011, London, UK.

Copyright © 2011 The Author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

Content must not be changed in any way or reproduced in any format or medium without the formal permission of the copyright holder(s)

When referring to this work, full bibliographic details must be given

<http://eprints.gla.ac.uk/84657>

Deposited on: 21 August 2013

To use or not to use software? An exploration of the use of software for qualitative data analysis.

Paper delivered at the British Educational Research Association Annual Conference 2011,
Institute of Education, London, September 6-8

Oscar Odena, University of Glasgow, UK

This document was prepared for oral delivery. For literal quotations please refer to the full paper subsequently developed and published as:
Odena, O. (2013) Using software to tell a trustworthy, convincing and useful story. *International Journal of Social Research Methodology*, Volume 16, Issue 5, pages 355-372, available at <http://dx.doi.org/10.1080/13645579.2012.706019>

Abstract:

This paper discusses the potential of specialist software to enhance qualitative data analysis and to substantiate the researchers' conclusions. An example from an enquiry on using music education as an inclusion tool in a post-conflict context is considered. A number of suggestions on how to support the researchers' claims are made. It is argued that the use of specialist software can enhance knowledge generation and, ultimately, if analyses processes are fully disclosed, improve the perception of educational research.

Introduction

This paper aims to discuss the potential of software to assist in category construction and to substantiate the researchers' conclusions. The use of specialist software for qualitative data analysis is a recurrent theme in research conferences, handbooks and special journal issues (for a comprehensive review see Odena, 2013). Over the last two decades competing claims on the value of computers for qualitative analysis have been put forward by advocates and sceptics. An undesired effect of software usage include that researchers may be misled to focus on frequency counts in transcripts rather than meaning, whether frequent or not. It has been argued that software packages may come to define the analysis processes they should merely support, de-contextualising the dataset (Lu & Shulman, 2008). Contrastingly, advocates describe the numerous advantages of using software, such as keeping track of developing ideas and an increased power for querying the datasets and for making links between their parts (Davidson & di Gregorio, 2011; Konopásek, 2008).

Advocates or not, software users sometimes mention the package used while not fully disclosing the particular analysis processes. This may perhaps be due to the word limit in articles as well as to the unstated assumption that readers will be familiar with such analysis processes. It appears that in some instances computers may be employed in a superficial way, to facilitate data management without making full use of the software possibilities, which may affect the way processes such as category construction is undertaken and subsequently explained to readers.

The purpose of this paper is twofold: to examine some of the possibilities of using software for qualitative data analysis and to discuss how its uses may best be reported to substantiate the researchers' claims. After reviewing a selection of relevant literature the paper considers an example from a recent study on the potential of music education as a tool for inclusion in a post-conflict context (Odena, 2010). In the conclusion it is suggested that disclosing the analysis processes can assist researchers to substantiate their claims for end-users.

Using computers in qualitative data analysis

The use of computers for qualitative data analysis has been a feature in social research since the 1980s, and today its use is part of many research methods courses (Davidson & Jacobs, 2008). When analysing text, as with any type of qualitative data analysis, there are several ways (and steps) in carrying out analysis processes that may be assisted by a software package. Apart from assisting with the managing and retrieving of different types of data across a number of datasets, the software may be employed in the process of category construction. This process may be located in a continuum depending on the degree of openness/closeness of the themes to be explored as well as the inductive/deductive methodological approach. At one end of the continuum and with little preconceived expectations, we would find 'grounded theory' (Birks & Mills, 2010; Glaser & Strauss, 1967). In grounded theory the categories would emerge through a process of inductive reasoning, rather than the data being allocated to predetermined categories. Ideally, the researchers would start without any defined ideas on what the findings may be. The analyses would then be undertaken following 'a constant comparative method', which would include: (1) Immersion: producing detailed transcriptions from diaries, interviews, observations, etc; (2) Categorisation: assigning categories; (3) Reduction: grouping categories in 'themes'; (4) Triangulation: checking themes against all transcripts, preferably with other people; and (5) Interpretation: making sense of data with new model or established theory.

At the other end of the analysis continuum we would find studies in which researchers have to identify predetermined categories using a deductive process and making use of, for instance, Boolean operators and set theory. In Qualitative Comparative Analysis (QCA) the approach requires the data to be manipulated as variables in order to maximise the number of comparisons that are made across a number of cases (Ragin, 1987; Rihoux, 2006). Somewhere in the middle of the inductive/deductive continuum we would find enquiries in which closely defined themes have to be explored from the outset but which do not require data manipulation (for instance in programme evaluation, e.g. Miller, Connolly, Odena & Styles, 2009; Odena, Miller & Kehoe, 2009). Regardless of the degree of inductive/deductive processes, qualitative data analysis, with and without the assistance of software, would always need to go through a process of reading, categorising, testing and refining, which is repeated by the researchers until all categories are compared against all the participants' responses, and the analysis validated with other individuals. The same process has previously been labelled recursive comparative analysis and thematic/content analysis (Odena & Welch, 2009, 2012).

The use of computers for qualitative data analysis, also known as CAQDAS (Fielding & Lee, 1998), appears to have a number of practical advantages in comparison to more traditional methods such as cutting and sorting quotations into boxes. Sorting text by hand is viable with tens of pages, but with hundreds of pages the researchers' memory may be aided by the software, as the number of categories and the relations between them is likely to develop with each additional reading of the transcripts. A number of software packages are currently available in the market, and although all

have particular features that are constantly being developed by their manufacturers, their baseline capabilities are similar. For instance, researchers can identify relevant quotations on the computer screen and code them using virtual coloured stripes. As the emerging ideas become clearer, whole categories can be easily merged or renamed. Most packages now have auto coding features and in some cases (e.g. *Qualrus*) these are designed to relieve the researcher of much of the assignment work. Auto coding does make some qualitative researchers uncomfortable, as it has been argued that although auto coding allows for fast exploration of all the answers to a question, it 'might not create a deep understanding of the issues raised' and has the potential of encouraging 'code fetishism' when 'the act of coding becomes an end in itself' (Richards, 2002, p. 269).

Nevertheless, it is the researcher who defines the auto coding parameters, amends the allocation of quotations assigned to categories and derives meaning from them. Some programmes have the option of counting the characters coded within each category, which can then be used to obtain the percentage of transcripts coded. Packages also have the option of writing memos and linking them to transcripts or other data, and of importing numerical results to other programmes (e.g. Lewins & Silver, 2007). Other possibilities include saving interim categorisations - allowing for analysis replication and tracing back/revising thinking paths - and the sharing of coded files (which aids collaborative work).

With all the above capabilities, these packages may ease the time spent managing data and ensure that no relevant quotations are overlooked. Nevertheless, there is some reticence regarding the use of this type of software, especially surrounding the perceived change of the researchers' role. Some researchers think computers can distinguish the relevant information from datasets and develop the ideas, in order to meet the research project's requirements (Crowley, Harre & Tagg, 2002; Lu & Shulman, 2008). In fact, the researchers are still in charge of building up the analysis, having the ideas, engaging with the data, assigning meaning and making all the decisions about the study.

Indeed, a challenge for all researchers is how they might substantiate their claims. In other words, what can researchers say which will enable readers to decide how much confidence they should place in the findings. In the next section some examples of category construction using specialist software are discussed. The case is made that a more detailed explanation of the researchers' analyses processes may better support their claims.

An example of category construction using software: *Music Education as a Tool for Inclusion Project*

This example of using software for qualitative data analysis is from an exploratory study of practitioners' views on the potential of music education as a tool for inclusion in cross-community activities in Northern Ireland (Odena, 2010). The main aim of the study was to explore how to develop music skills while bringing children from both main communities together. Fourteen interviewees were purposefully selected following a maximum variation sampling approach, taking into account their potential as key informants as determined by having extended experience with this type of activities. Interviewees were working or had worked in a wide variety of contexts including school and out-of-school music projects. The interviews were semi-structured and attempted to explore the participants' background, their views on music education in Northern Ireland, and advice on how to increase the effectiveness of cross-community projects. Verbatim transcriptions were analysed using thematic analysis with the assistance of specialist software (*NVivo*). 93.32% of over 216 double spaced pages of text were coded into categories - 253,742

characters out of 271,905. This process consisted of repeated readings of all transcripts, looking for commonalities and themes, which were tested with each new reading and evolved into the final categories. A sample of the categorised text was discussed with two colleagues, giving further reliability to the analysis. Thirteen categories emerged, four of which were most relevant in addressing the aim of the enquiry across all interviewees, including ‘project processes and effectiveness’ and ‘music as a sign of identity’.

The analysis showed how the activities and aims explained by interviewees varied depending on a number of factors, one of the most important being the level of acknowledgment of integration of the educational setting, which appeared to be influenced by the socio-economic environment. It was apparent that cross-community music education projects had been and continued to be an effective means of addressing prejudice amongst young people, although addressing prejudice may have not always been the aim of all projects but a welcomed side effect. The analysis also highlighted barriers for cross-community education and some negative musical stereotypes linked with each community. The potential of using music for cross-community activities was highlighted by all interviewees. Successful activities described included school visits with a musical element, shared after-school music education activities in neutral settings, and collaborative performances between schools across the community divide:

[Music] is a superb tool for encouraging children to work together...they throw themselves into it wholeheartedly and are quite prepared to work with other people in doing that.

[Children] can inspire people like no other group of people can.

The specialist software aided in the process of disclosing the most relevant categories, not just in addressing the research aim but as strong categories emerging across conversations with interviewees in all their different contexts. Table 1 below shows the number of appearances of the four main categories within the transcripts and the number of interviewees that had quotations coded within these categories (with % in brackets):

Table 1. Transcript appearances of the four main categories in the *Music Education as a Tool for Inclusion* study (adapted from Odena, 2010, p. 91)

Category	Number of quotations categorised	Number of interviewees
‘Project processes and effectiveness’	51 (16 in subcategory ‘Barriers for cross-community education’)	14 (100%)
‘Music education potential’	29	14 (100%)
‘Music as a sign of identity’	23	10 (71%)
‘Socio-economic factors’	12	8 (57%)
Total number of quotations for this set of categories	115	

Disclosing the relative weight of the emerging categories was used to substantiate the conclusions of the study; for instance, by showing the degree to which quotations used in written outputs were representative of the participants’ views, and by evidencing that particular categories appeared across interviews that had been carried out in a wide variety of contexts due to the maximum

variation sampling approach. Disclosing frequency data in the analysis allowed for the informed assessment of emerging patterns across datasets and for a consideration of alternative explanations.

A few months after completing the first analysis a second one was carried out, which was aimed at developing theory-practice links between the interview transcripts and Social Psychology theories. In particular, instances of 'optimal conditions' for cross-community contact and developing stages of inter-group relations as described in the literature were mapped out in the participants' explanations (e.g. Kenworthy, Turner & Hewstone, 2005). At the core of organising cross-community activities amongst confronting groups lies the idea that intergroup contact, under certain conditions, can be effective in reducing prejudice and hostility between groups. The optimum conditions for this to happen include: (1) equal status of both groups in the contact situation; (2) ongoing personal interaction between individuals from both groups; (3) working towards a common goal and (4) official social sanction for contact between groups (Hughes, 2007). This theory, also known as *contact hypothesis*, was first proposed by Allport (1954, p. 489) who observed that to maximise programme effectiveness, contact activities would need to 'occur in ordinary purposeful pursuits'. In a subsequent reformulation of the contact theory, Pettigrew (1998) outlined a sequential model to reduce conflict between groups containing three stages:

1. The first stage or initial contact, where anxiety is likely to be more pronounced and where personal identity and interpersonal interaction are emphasised in an effort to 'de-categorise' the individuals.
2. A second stage when contact is well established, which affords a situation with less anxiety in which the old salient categorisation of belonging to a group is disclosed, resulting in weakened prejudices that are generalised beyond the activity.
3. And a third and final stage in which, after extended contact, individuals begin to think of themselves as part of a re-defined new larger group that comprises all communities.

The original analysis of the interviews showed that cross-community music activities were perceived positively by both communities and that such activities worked as an inclusion tool in a variety of contexts. Projects seemed to be well received and were described as benefiting the children, but the first analysis did not illustrate the nature of the activities in relation to any diminishing prejudice. Conversely, after the second analysis aimed at outlining links with Social Psychology theories, it was apparent that the majority of activities described would fall within Pettigrew's 'initial contact', as weakened prejudices were not generalised beyond the often one-off cross-community element. A number of educational implications followed, such as that after decades of violent conflict, projects in highly polarised areas would benefit from focussing on the quality of the children's musical experience, 'leaving positive attitudes towards the other community to develop naturally' (Odena, 2010, p. 99). To move beyond Pettigrew's 1st stage, projects needed to be sustained and had to offer something that enticed all those involved: fun for participating children, attractive educational aims for parents, and a degree of status to attract school leaders. In affluent contexts with less obvious polarisation, cross-community activities such as youth orchestra rehearsals were already happening regularly (some interviewees participated in them while growing up). This supported the suggestion that negative anxiety produced by the anticipation of contact with the other community was likely to be ameliorated by early participation in contact activities under optimal conditions (Kenworthy, Turner & Hewstone, 2005).

It may be argued that similar conclusions could have been arrived at without using software. Nevertheless, the software afforded a second in-depth and rigorous analysis including comparisons

within a layered structure of categories and subcategories, and facilitated the search of particular expressions across all interviews making sure that no stone was left unturned. The ready access to all quotations also eased the retrieval of representative instances to prepare research outputs with different foci for different audiences. Ultimately, the study's educational implications reached beyond the original context of the enquiry, as teachers in Cyprus adapted them for their particular post-conflict environment (Odena, 2009).

Concluding thoughts: substantiating claims when analysing qualitative data with computers

As suggested before, perhaps due to word limitations researchers often mention employing a particular software package while leaving the analysis processes undisclosed. In order to be able to assess the researchers' claims, the processes employed to analyse each dataset would need to be explained, as well as the provenance of the emerging themes (and its relative weight if relevant for the research questions). There are an increasing number of organisations commissioning studies whose research aims are often tailored to the organisation's interests. This is a trend that can be seen across developed democracies: market forces bring increased funding sources and a variety of actors that produce social research, often blurring the boundaries between enquiry, advocacy and business. Beyond the unresolved debates about research 'quality' that draw on what counts as knowledge which would fall outside the scope of this paper, some general suggestions on ways to substantiate the researchers' claims are discussed in this section.

Before the advent of computers the option of disclosing the percentage of data categorised was not readily available to researchers. Current software packages can count the number of characters coded and percentages can be easily calculated. In order to rule out explanations by what in terms of qualitative analysis might be labelled as 'unquoted evidence', the percentage of text categorised may be stated. The number of pages and line spacing of text analysed and the number of pages for each distinctive dataset could also be disclosed to give readers an overall sense of the relative size of the datasets available. Moreover, the type of data analysis, whether it follows an inductive/deductive approach or a degree of both, would need to be explained, as there is a variety of interpretations even for established approaches.

If coding is undertaken by a single researcher reliability procedures would need to be put in place. Sharing the analysis with others is useful to ensure the interpretation is not biased towards the views of a single researcher, and software can assist in sharing interim analyses and developing them collaboratively. If software is used there is the potential for greater transparency but there is also more procedural complexity to be described.

Examples of how analytic processes may be aided with the use of software have been considered. For instance looking for the frequency of categories across a number of datasets to minimise any unconscious bias when reporting emerging themes across a number of documents (e.g. Table 1). Although the meaning and relevance of categories are attributed by the researchers, such disclosure would aid in ruling out counter explanations in a systematic way. Gorard (2002) argues that research claims would need to persuade the sceptical reader rather than play to a gallery of existing converts. Specialist software can assist in eliciting new viewpoints by offering different perspectives into the datasets, which is particularly useful when seeking alternative explanations and when trying to reduce any bias (Webb & Vulliamy, 2007).

The above procedures would provide the readers with an indication of the robustness of the analysis. In studies with a small number of participants this would ensure that representative quotations provided in reports are truly representative of the data collected. In a seminal paper Adelman, Jenkins and Kemmis (1976; p. 143) observed that in case studies, when attention is given to the particular, the meanings presented in the report are likely to become apparent by 'the shock of recognition'. Misrepresenting the case might weaken the researchers' claims as much as over-claiming beyond the scope of the sample. In this type of enquiry the applicability of the claims is often made by the readers, who are able to recognise similar realities which may help them make sense of their own. Ultimately, the development and disclosure of research analysis processes facilitated by software programs can afford readers the opportunity to judge whether the researchers' claims are sound rather than whether they agree with them simply from the perspective of their own particular experience.

Nevertheless, the possibility of producing conclusive claims would depend on the issues under scrutiny and how the research questions are framed, because implications developed from social enquiry are normally context and time bound. Popper (1963) maintained that researchers should not try to predict things by solely amounting instances of evidence. Instead, the way forward to advance knowledge would be to falsify accepted theories. This is often achieved when using a methodological approach focussed on the in-depth study of a small number of participants. With this approach, however, different data analyses processes applied to the same datasets may yield different claims, which is why these processes would need to be disclosed to end-users.

A number of authors have discussed whether it can ever be assumed that the reasonableness of a claim derived from evidence can be beyond doubt (e.g. Hammersley, 2007). Ultimately, the impact of research will be determined not just by quality and relevance but also by luck and rhetorical power (Gardner, 2011). There is a need to be persuasive and credible to facilitate research impact and disclosing the analyses rigorously may improve research reports and their perception by end-users. It is hoped that the suggestions outlined in this paper will contribute ideas for the consideration of researchers engaged in interpreting qualitative data from any settings, regardless of their particular research practices.

Acknowledgement

The above example of category construction comes from a study funded by The Bernard van Leer Foundation and directed by the author.

References

- Adelman, C., Jenkins, D., & Kemmis, S. (1976). Re-thinking case study: Notes from the second Cambridge Conference. *Cambridge Journal of Education*, 6(3), 139-150.
- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Birks, M., & Mills, J. (2010). *Grounded theory: A practical guide*. London: Sage.
- Crowley, C., Harre, R., & Tagg, C. (2002). Qualitative research and computing: Methodological issues and practices in using QSR NVivo and Nud*ist. *International Journal of Social Research Methodology*, 5(3), 193-197.
- Davidson, J., & di Gregorio, S. (2011). Qualitative research and technology: In the midst of a revolution. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (4th ed., pp. 627-644). London: Sage.

- Davidson, J., & Jacobs, C. (2008). The implications of qualitative research software for doctoral work: Considering the individual and institutional contexts. *Qualitative Research Journal*, 8(2), 72-80.
- Fielding, N. G., & Lee, R. M. (1998). *Computer analysis and qualitative research*. London: Sage.
- Gardner, J. (2011). Educational research: What (a) to do about impact? *British Educational Research Journal*, 37(4), 543-561.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine de Gruyter.
- Gorard, S. (2002). Fostering scepticism: The importance of warranting claims. *Evaluation and Research in Education*, 16(3), 136-149.
- Hammersley, M. (2007). The issue of quality in qualitative research. *International Journal of Research and Method in Education*, 30(3), 287-305.
- Hughes, J. (2007). Mediating and moderating effects of inter-group contact: Case studies from bilingual/bi-national schools in Israel. *Journal of Ethnic and Migration Studies*, 33(3), 419-437.
- Kenworthy, J., Turner, R. N., & Hewstone, M. (2005). Intergroup contact: When does it work, and why? In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice* (pp. 278-292). Oxford: Blackwell.
- Konopásek, Z. (2008). Making thinking visible with Atlas.ti: Computer assisted qualitative analysis as textual practices. *Forum: Qualitative Social Research*, 9(2). Retrieved August 1, 2013, from www.qualitative-research.net/index.php/fqs/article/view/420
- Lewins, A., & Silver, Ch. (2007). *Using software in qualitative research: A step-by-step guide*. London: Sage.
- Lu, Ch.-J., & Shulman, S. W. (2008). Rigor and flexibility in computer-based qualitative research: introducing the Coding Analysis Toolkit. *International Journal of Multiple Research Approaches*, 2(1), 105-117.
- Miller, S., Connolly, P., Odena, & Styles, B. (2009). *A Randomised Controlled Trial Evaluation of Business in the Community's Time to Read Pupil Mentoring Programme*. Belfast: Centre for Effective Education, School of Education, Queen's University Belfast.
- Odena, O. (2009). Exploring the potential of music education to facilitate children's cross-community activities in Northern Ireland. Keynote at the international conference 'Intercultural education and peaceful co-existence: The role of the school', Pedagogical Institute of Cyprus, Larnaca, November 19.
- Odena, O. (2010). Practitioners' views on cross-community music education projects in Northern Ireland: alienation, socio-economic factors and educational potential. *British Educational Research Journal*, 36(1), 83-105.
- Odena, O. (2013). Using software to tell a trustworthy, convincing and useful story. *International Journal of Social Research Methodology*, 16(5), 355-372.
- Odena, O., Miller, S., & Kehoe, S. (2009). A qualitative evaluation of a mentoring reading programme for 8-9 year olds in Northern Ireland. *Education-line*. Retrieved August 1, 2013, from www.leeds.ac.uk/educol/documents/184822.pdf
- Odena, O., & Welch, G. (2009). A generative model of teachers' thinking on musical creativity. *Psychology of Music*, 37(4), 416-442.
- Odena, O., & Welch, G. (2012). Teachers' perceptions of creativity. In O. Odena (Ed.), *Musical creativity: Insights from music education research* (pp. 29-48). Farnham: Ashgate.
- Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, 49, 65-85.
- Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge and Kegan Paul.

- Ragin, Ch. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. London: University of California Press.
- Richards, L. (2002). Qualitative computing—a methods revolution? *International Journal of Social Research Methodology*, 5(3), 263-276.
- Rihoux, B. (2006). Qualitative Comparative Analysis (QCA) and related systematic comparative methods: Recent advances and remaining challenges for social science research. *International Sociology*, 21(5), 679-706.
- Webb, R., & Vulliamy, G. (2007). Changing classroom practice at Key Stage 2: The impact of New Labour's national strategies. *Oxford Review of Education*, 33(5), 561-580.