Introduction to the KWALON Experiment: Discussions on Qualitative Data Analysis Software by Developers and Users

Jeanine C. Evers, Christina Silver, Katja Mruck & Bart Peeters

Abstract: In this introduction to the KWALON Experiment and related conference, we describe the motivations of the collaborating European networks in organising this joint endeavour. The KWALON Experiment consisted of five developers of Qualitative Data Analysis (QDA) software analysing a dataset regarding the financial crisis in the time period 2008-2009, provided by the conference organisers. Besides this experiment, researchers were invited to present their reflective papers on the use of QDA software. This introduction gives a description of the experiment, the "rules", research questions and reflective points, as well as a full description of the dataset and search rules used, and our reflection on the lessons learned. The related conference is described, as are the papers which are included in this FQS issue.

Table of Contents

1. A Collaborative Exercise: Four European Networks United in a Shared Interest
2. The Ideas behind the KWALON Experiment
3. The KWALON Experiment: Project Design and the Rules of the "Game"
   3.1 The substantive topic
   3.2 The data set
   3.3 The research questions
   3.4 Presenting the results
   3.5 Users on QDA software
4. What Have we Learned from the Experiment?
Appendix 1
Appendix 2
References
Authors
Citation

1. A Collaborative Exercise: Four European Networks United in a Shared Interest

The current issue of FQS is devoted to what we have named: The KWALON Experiment; a collaborative effort of four European networks that focus on qualitative research: KWALON1, based in the Netherlands; The CAQDAS Networking Project2, based in the UK; FQS, an international journal, based in

---

1 KWALON, The Netherlands Association for Qualitative Research, was founded in 1995. Its mission is to promote the use of qualitative research, and to discuss, develop and enhance qualitative methods. KWALON issues a Dutch-language journal with the same name three times a year, provides training in qualitative research methods and QDA software, and organises a yearly national conference. Members are mostly affiliated to the several Dutch (Applied) Universities and research organizations.

2 The CAQDAS Networking Project was established in 1994 to provide information, advice, training and ongoing support in the use of a range of software packages designed to facilitate...
Germany, and Kwalitatief Sterk\textsuperscript{3}, based in Belgium. The experiment involved asking developers of several Qualitative Data Analysis (QDA) software packages to analyse the same dataset in order to establish the extent to which their products are comparable. Section 2 of this introduction describes the ideas, experiences and questions that guided the design of the experiment. The experiment was organised within the framework of a broader conference, which took place on April 22 and 23, 2010 at the University for Humanistics in The Netherlands. At this conference, entitled: "Is QDA software really comparable?", the preliminary results of the experiment were presented by developers or their representatives, together with papers presented by users of QDA software about the role and use of software in their projects. Section 3 outlines the design and "rules" of the experiment and Section 4 discusses lessons that can be learned from this first endeavour. [1]

Let us start by introducing ourselves. The idea started in The Netherlands, with Jeanine EVERS from KWALON, hence the "KWALON experiment", who was interested in systematically uncovering the implications of differences between CAQDAS packages. She had some informal discussions with delegates at the 2009 Berliner Methodentreffen\textsuperscript{4} at which she received some encouraging reactions from developers. She then approached Christina SILVER from the CAQDAS Networking Project, University of Surrey in the UK, Katja MRUCK from FQS, Freie Universität Berlin, Germany, who is also co-organiser of the Berliner Methodentreffen, and Bart PEETERS from Kwalitatief Sterk in Flanders, Belgium, to help design the experiment and organise the conference. The results of our joint efforts are presented in this issue of FQS. [2]

From the KWALON point of view, the experiment originated from a growing uneasiness with some questions asked by potential QDA software users at training events concerning the comparability of software, the extent to which these are addressed during conferences dealing with the use of QDA software, and the role of developers as actors in the debate. These aspects are dealt with in Section 2 of this introduction. As all of the national networks approached have a common goal, i.e., promoting qualitative research in their respective countries, and wanting to discuss and expand its methods, a collaborative effort would not only suit our national goals, but might foster a European effort to this regard, as well as expand knowledge concerning each others' national efforts and work. As such, we could be inspired by one another. Another reason to seek collaboration was the specialist and international nature of the topic. QDA software is developed in various countries, as is the expertise surrounding it. It was therefore

---

\textsuperscript{3} Kwalitatief Sterk, The Flemish/Belgian Association for Qualitative Research, was founded in 2004. It is a cooperation between different Flemish institutes for higher education and serves as a platform for the promotion of activities aimed at the enhancement of knowledge on qualitative methods. Kwalitatief Sterk organises a yearly conference on qualitative research.

\textsuperscript{4} The Berliner Methodentreffen Qualitative Forschung [Berlin Meeting on Qualitative Research Methods] is the most important event for researchers from German speaking countries and takes place annually at the Freie Universität Berlin since 2005. See http://www.berliner-methodentreffen.de/ for further information; at http://www.qualitative-forschung.de/methodentreffen/archiv/video/ videos from different sessions are available.
important to reflect this in the organisation and amongst those attending the conference. Last but not least, there was a simple pragmatic reason for organising the conference jointly: with more hands on the job, the job became feasible. Jeanine EVERS organised and chaired the Experiment and conference on behalf of KWALON. [3]

The CAQDAS Networking Project participated in this endeavour because they were keen to explore differences between software packages from the point of view of the developers themselves. Christina SILVER and her colleagues spend a lot of time comparing packages and helping researchers make informed choices and were interested in understanding developers' own assessments of the strengths of their products and their analytic utility. [4]

The Flemish network Kwalitatief Sterk decided to participate in the scientific organisation committee since they are regularly confronted with questions regarding the use of software for qualitative research. Bart PEETERS, sociologist and experienced NVivo-trainer, coorganised the conference on behalf of Kwalitatief Sterk. [5]

Katja MRUCK joined the effort, as the topic this issue is dealing with is important for qualitative research(ers) and therefore of an enduring interest for the international and multidisciplinary readership of FQS. She especially appreciated to work and reflect on this topic in a rather systematic way, exploring the links (and gaps) between method(olog)ical decisions, software use, and the results produced. [6]

The organisers of this first experiment found the process to be both valuable in terms of the aims, and inspiring in terms of the collaboration between the respective organisations. We have therefore agreed to explore the possibilities and opportunities for making this a regular event, occurring every few years, with a changing central theme. [7]

2. The Ideas behind the KWALON Experiment

One of the key and unique aspects of the KWALON conference was the experiment undertaken with software developers to investigate the question of whether these packages are really comparable. Representatives from the following products attended the conference and present the results of their participation in the experiment in this issue: ATLAS.ti, Cassandre, MAXqda, NVivo, Transana (see articles by DEMPSTER & WOODS, FRIESE, LEJEUNE, KUCKARTZ & SHARP, and WILTSHIER, respectively, to be found in this issue). [8]

There were several ideas behind conducting the experiment. Firstly was that as teachers of qualitative analysis using software, we frequently receive questions from students and researchers as to which software is the "best" or the most suitable for particular projects, and in which way the use of technology influences the research process and its results. Indeed, this is a long-standing question; one that has been around as long as qualitative software itself (for example, see
KELLE, 1995; FIELDING & LEE, 1998; GIBBS, FRIESE & MANGABEIRA, 2002; LEWINS & SILVER, 2007; DI GREGORIO & DAVIDSON, 2008). In addition, a growing number of paper submissions are made to FQS in which authors simply state that a particular software has been used without discussing in detail the implications of use; in which it is implied that the software is a method of analysis; or in which there is limited discussion of the relationship between theory, method(ology) and software. Several articles in this issue address the topic of software enabling us to enhance the quality of our projects (EVERS, FRIESE, SILVER & PATASHNICK, MÜHLMEYER-MENTZEL), as do some authors (BAZELEY, 2007; DI GREGORIO & DAVIDSON, 2008; KONOPÁSEK, 2008; LEWINS & SILVER, 2007; RICHARDS, 2005) but this is not to say, that software will enhance quality merely through its use, nor that software should be seen as a method. It was therefore of interest to observe how software developers themselves would use their own products to analyse data; what analytic approach would they adopt, how would that correspond to particular software tools, and to what extent would their practices reflect those commonly experienced by researchers? However, addressing questions relating to the appropriateness of software packages for particular types of projects and approaches is not as straightforward as those who ask them may like. Similarly, the relationship between methodology and technology is complex. A whole range of factors need to be considered in choosing between software, from the research design; type and amount of data; the approach to data analysis; the dynamics of project (e.g. individual, team); personal style of working; to time and financial resources; etc. There are now many qualitative software packages available to choose between; from the well-known "market leaders" through freely available options and those developed to support specific analytic approaches or methodologies. Each have their advantages and disadvantages, and for these reasons there can be no one best software for analysing qualitative research. It is therefore understandable that researchers often struggle to decide which package to invest the time and effort into learning. In addition, this raises the question of whether the specificities of the range of packages available renders direct comparison between them possible, or even useful. [9]

Secondly, in organising an international conference on qualitative software use we wanted to ensure it would be as useful as possible in assisting researchers faced with these dilemmas to find valuable answers. We were aware that it is common at such conferences for researchers working in quite different disciplines, substantive areas and analytic approaches to present their work on projects using a wide variety of software packages. This was certainly the case at the last such academically organised international conference, the CAQDAS 07 conference5, at which 22 different software packages, designed to facilitate some aspect of qualitative research process, were discussed6. We felt that this could be confusing for those seeking advice on the appropriateness of software for their own needs and that, as a result, many conference delegates may leave with more questions, than they had arrived with. Finding relevant software advice from

5 London, 18-20 April 2007, organised by Christina SILVER of the CAQDAS Networking Project.
6 Besides that, there are the yearly Computer Assisted Qualitative Research Conferences, organised by the Merlien Institute.
research peers may therefore be an accident of coincidence at such conferences, rather than design. For example, stumbling upon a paper in which a similar topic, design and analytic approach was being employed. Where this happens a fairly reliable picture of how to use a particular software effectively may well be obtained. However, the chance involved in this happening was something, as conference organisers, we wanted to address. [10]

We also wanted to invite software developers of both commercial and non-commercial and/or open source software to attend the conference, as this provokes both interesting debate and provides relatively rare opportunities for researchers to meet developers and discuss particular needs, as part of decision making processes. [11]

As a result of these issues, the idea of the "developer experiment" began to grow: to invite developers to analyse the same dataset using different QDA software as a means to make more direct comparisons of the capabilities of software, and therefore be more explicitly useful for researchers trying to choose between products. [12]

3. The KWALON Experiment: Project Design and the Rules of the "Game"

In order to make the experiment directly comparable we needed to keep as many elements as constant as possible. However, we also needed to be sensitive to the variety in the software packages taking part. Asking developers to participate meant that we had to provide each with the ability to showcase their particular strengths. The aspects that follow were at the forefront of our minds as we designed the project, and underlie the specific questions we asked developers to address. [13]

3.1 The substantive topic

The topic needed to be substantively accessible to those analysing the data, the conference delegates and readers of FQS, to whom the results would be presented. We therefore decided to choose a topic of broad and topical interest, deciding upon the financial crisis, then topic of the day, which we felt had relevance across disciplines, methodological approaches and national boundaries, and could provide us with enough material to be found on the Internet. [14]

3.2 The data set

Giving the developers an opportunity to demonstrate their products and to take the role of a user had a particular impact upon compiling the dataset. For example, some packages focus on the analysis of audio and video data (e.g., Transana and DRS) whereas others focus on the analysis of text (e.g., Cassandre, and at the time of designing the experiment, MAXqda), or on inter coder agreement in textual coding (CAT-QDAP). Yet others focus on both (e.g.,
ATLAS.ti, Hyperreasearch or NVivo), or on a combination of qualitative and more quantitatively oriented content-analysis (QDA Miner with WordStat or Veyon®), or on a very specific type of analysis of organisations called GABEK, using the software WinRelan. We therefore compiled a data set of freely available data, collected from the Internet; of newspaper articles, websites and weblogs, video and audio files. We made sure that in all of the files collected, there was a broad variety of geographical origin, actors involved and file formats (see Appendix 1). We did not have the resources to generate any interview or focus-group data, but considered newspaper articles to be an adequate form of textual data for the purposes of the experiment. As a further way in which developers could be able to work to the strengths of their software, they were allowed to collect up to ten additional data files of one type, to bolster the data set. Although we encouraged them to use all of the data set we had provided, they could choose not to if this was not practically or technically possible (see Appendix 2 for an overview of files used by each developer). [15]

In terms of timescale, we aimed to collect data of each type published in June 2008 and June 2009. However, this proved particularly challenging for the audio and video data and therefore not all of this data conformed. We did not alert the developers to this because we were interested in the extent to which those analysing the data set would notice this inconsistency and how they would handle it within their analysis. [16]

3.3 The research questions

Although developers were free to take any analytic approach to the data set, we needed them to attend to the same set of research questions to ensure conference delegates and FQS readers would be able to assess the extent to, and manner in which particular software packages facilitated their answering. The following research questions were posed:

1. Did views or arguments about the causes and prognosis of the economic crisis change between 2008 and 2009?
2. How do different actors identify the causes of the economic crisis?
3. How do different actors predict the future consequences of the economic crisis?
4. And how do they justify their positions? [17]

3.4 Presenting the results

We also gave developers some guidance about the content of the presentations we hoped they would provide at the conference, as another way of maximising the extent to which the results would be comparable. They were asked to attend to the following questions in their presentations:

• How did you use the software? What analysis strategies did you use, for which reasons, and how did the tools in the software enable that?
• How did the software enable you in keeping track of the way your data analysis was influenced by your background, personality, and/or theoretical framework?
• With which type of data did you work best and why?
• How did the software help you in your project? What were down sides, what were good aspects of the software?
• Did the software alter your way of analysing and if so, how?
• How much time did you spend on the analysis? [18]

Unfortunately not everybody who was willing to participate in the experiment, managed to be present at the Conference. In this respect we should mention CAT-QDAP, Digital Replay Systems, Framework, WinRelan and Veyor®. In addition, not everyone approached was willing to participate in the experiment. Developers who participated are represented in the current issue, as are selected papers from the conference. [19]

3.5 Users on QDA software

As well as inviting developers to participate in our experiment, we wanted to give users a forum to present their critical appraisal of a wide range of packages. Submitted abstracts could deal with different themes related to the use of QDA software, in which we wanted them to reflect on how the software (tools) helped or hindered them with their analysis, and how their own preferences, theoretical, professional, and cultural background, experiences and personality shaped the analysis. We asked users similar questions to the ones we asked the developers above, to guide their presentations, in order to ensure the focus of all the papers presented at the conference and in this FQS issue would be somewhat similar. By inviting independent users and asking them questions very similar to the ones we asked the developers, we wanted to make sure that conference participants would get a critical (albeit incomplete due to there only being five developers partaking in the experiment) overview of available software. [20]

After the review process nineteen presenters were invited to the conference of which fourteen were, despite an Icelandic volcano which locked the European airspace, able to present their findings in Utrecht or via a video conference. Four out of five developers managed to get to the conference in time, and they presented their results of the Experiment in the morning sessions. Afternoon sessions were preceded by a reserved timeslot for developers to be present at their software stands in separate rooms, thus enabling users to speak with developers of their choice, and for developers to attend the whole conference as participants. The afternoon sessions at which users presented papers were organised according to the following themes: Future developments, Users and their use of software, Combining qualitative and quantitative methods, Epistemological issues in using software, Different analysis strategies using software, Users comparing software, and Teaching and learning software. The afternoon sessions were closed with a plenary session. In this issue, users contributions are presented in the section "The Experiment: Users' Perspectives"
Two delegates (Di Gregorio and Schumann, this issue) were asked beforehand to attend the conference as "mystery guests" and write a column for this issue. We were interested in how the experiment and users' papers would be conceived by delegates, related to their experience with QDA software use and qualitative research in general. The two columns provide an insight into how the experiment and conference were perceived by delegates with differing experience with qualitative software. [22]

4. What Have we Learned from the Experiment?

As this is the first time ever that such an experiment was organised, with developers participating by free choice in an experiment explicitly designed to compare software products, we are quite pleased with the outcome. The conference and the papers presented in this issue contribute to the satisfaction of our first goal; to illustrate that there is no "best" software and the decision for a software is a multi-facetted one. The next goal; having developers actively contribute to the debate, was successful as well, and we sincerely hope they enjoyed having the time to participate and having a timeslot available to talk with interested (potential) users. [23]

There has, however, been some criticism; points for us to take into the future, to our next endeavour. Regarding the dataset, one criticism from developers concerned the amount of data that was provided. It was felt that we should have provided fewer data files. A related point was the freedom we left developers to pick and choose from the dataset and generate more of their own data. Indeed, with hindsight, reflecting on the data set, less data would not have affected the sample selection criteria, mentioned in Appendix 1. However, ensuring developers had the opportunity to showcase their product nevertheless made a large number of different data files necessary. [24]

Criticism was also levied at the format of the data; specifically the absence of RTF and DOC files and the difficulty in accessing some of the (video) files (as some of the data had been created on an OS-X based computer). It was important to include several data formats as this resembles the reality of a researcher, not necessarily being a computer expert and needing to convert files to make them readable by particular software. It was also important to preserve the original layout, as this might have added to the analysis. Accessing and converting data is part of the daily struggle of a researcher, who needs to do so without necessarily being a technical expert or having sufficient support in order to do so. The more expertise one has on the technical aspects of computing and QDA software, the less one remembers how difficult, time-consuming and annoying those things can be. [25]
Some felt the subject matter, the financial crisis, being one that nobody was specifically knowledgeable about, to be a disadvantage. We do not think so, as our main point was to have something of broad interest and it was a gain, we think, that none of the developers was an expert in this topic. Thus everybody was more or less on the same level in this respect. [26]

On another foot was the point mentioned about the different backgrounds of persons in the experiment; not only qua discipline, but qua methodological viewpoint as well. This of course creates differences in approach, but then again, is part of the reality in the world of social sciences and specifically in the world of qualitative research. Indeed, this was one of the aspects we wanted to uncover. As there is no fixed way of interpreting qualitative data, there will always be differences in results. Even within the same discipline, people can have different viewpoints, depending on theoretical preferences, that they will apply to their data. Papers presented at the conference and published in this issue vary in the detail with which the analysis process using QDA software is described and underpinned, and this represents the trouble some researchers have in dealing with the relationship between software, theory and method. This then, is a point for continued debate. [27]

As such, a "real" experiment is non-feasible. We think we managed to keep constant as much as was possible. So one might ask: was it an experiment then? Not in the strict scientific sense, for sure. JANSEN (2010) refers to it as a "social experiment", and maybe that indeed is a better description. However, we very much enjoyed designing and organising it and feel the process and outcome was of value. We would therefore like to express our deepest appreciation to developers attending, for their willingness to put a lot of time and effort in this "experiment" and take on the role of a user. The world of qualitative data analysis software development is commercialising rapidly, which puts a lot of pressure on the time available for this kind of knowledge-building. Again, many thanks! We hope to continue the dialog and debate between users and developers regarding the use of QDA software in qualitative research, thus continuing to enhance our understanding, methods and methodology into the future. [28]

Appendix 1: Dataset for the KWALON Experiment Specified

Appendix2: Overview of Data Used by Developers

References


**Authors**

**Jeanine C. EVERS** is an active member of KWALON, the Dutch Network for Qualitative Research and founder and manager of its course division, which provides training in qualitative research methods and QDA software in The Netherlands from 2004 onwards. KWALON is the Dutch Network for Qualitative Research, which promotes the use of qualitative research and strives to enhance the quality of qualitative research in The Netherlands. She has been an active member of KWALON from 1995 onwards. Besides that, she is a lecturer in qualitative research methods at the University for Humanistics in Utrecht, and she has her own company, Evers Research & Training, which provides training and consultancy services for groups and individual researchers.

**Contact:**
Jeanine C.Evers
University for Humanistics
P.O. Box 797
3500 AT Utrecht
The Netherlands
E-mail: j.evers@uvh.nl or jcevers@eversresearch.nl
URL: [http://www.eversresearch.nl/](http://www.eversresearch.nl/)

**Christina SILVER** is the manager of the CAQDAS Networking Project, for which she has worked since 1998. The project provides information, advice and training in a range of software packages designed to facilitate the analysis of qualitative data. She leads the training and capacity building activities of the project and is involved in research conducted under the National Centre for Research Methods (NCRM) funded project Qualitative Innovations in CAQDAS (QUIC). Christina's particular interests relate to the relationship between technology and methodology and the integration and analysis of non textual data, using software. She is also the co-founder of Qualitative Data Analysis Services (QDAS), which provides customised training and consultancy services for individuals and groups engaged in qualitative analysis.

**Contact:**
Dr Christina Silver
Qualitative Innovations in CAQDAS
CAQDAS Networking Project, Manager
University of Surrey
Department of Sociology
GUILDFORD, GU2 7XH, UK
Tel: +44 (0)1483 689455
E-mail: c.silver@surrey.ac.uk
URL: [http://caqdas.soc.surrey.ac.uk](http://caqdas.soc.surrey.ac.uk)

**Katja MRUCK** is *FQS* main editor and Head of the Institute for Qualitative Research, Berlin, Germany.

**Contact:**
Dr. Katja Mruck
Freie Universitaet Berlin
Ihnestr. 24
Berlin, Germany
E-mail: katja.mruck@fu-berlin.de

Bart PEETERS is a member of Kwalitatief Sterk's executive committee since 2005. As a medical sociologist he has been performing qualitative research on diabetes, adherence, ethnicity & health and poverty & health.

Contact:
Bart Peeters
Ghent University
Faculty of Pharmaceutical Sciences
Pharmaceutical Care Unit
Harelbekestraat 72
9000 Gent
Belgium
E-mail: b.peeters@ugent.be

Citation