Computer Assistance and the Evolution of Qualitative Data Analysis

PEKKA RÄSÄNEN
Researcher
Center for Knowledge and Innovation Research
Helsinki School of Economics and Business Administration
Finland

Abstract

The aim of the paper is to examine the emergence of computer assistance in qualitative research and to describe its impact on research practices. The use of data analysis programs is approached from methodological frameworks of basic and structured qualitative approaches. Computer use is also examined from the viewpoint of concrete operations of data analysis. Particularly the new features, which software packages are claimed to have brought to data analysis, are discussed in the paper.

It is argued that computer assisted techniques provide possibilities for developing the often impressionistic qualitative data analysis procedures into something more integrated, explicit and systematic. However, it is also emphasised that analysis programs themselves do not represent any methodological approaches. By offering a special toolkit for data processing, computer assistance should be understood only as an attempt to enhance analytic practices in qualitative research.

Keywords: computer assistance, methodological transparency, grounded theory, structured qualitative approaches, research validity and reliability

Introduction

During the 1990s computer assisted qualitative data analysis developed into a distinct topic among qualitative researchers. It has been stated that computer assisted data analysis is no longer a peculiarity in social sciences, but merely one of the key areas of qualitative methodology development (e.g. Kelle & Laurie 1998, 27-28; Dey 1993; Luomanen & Räsänen 1999). There is no doubt that computers have become increasingly

common tools in academic research in the last two decades. It is also evident that there have been major advances in computer programs designed to support qualitative data analysis. Software packages have become, for example, noticeably more user-friendly. Most people familiar with PC or Macintosh operating systems are able to learn, in just hours, how to utilise different programs effectively. As a consequence, software use has indeed become an indispensable part of research processes for many qualitative researchers.

It is often argued that computer software can be of great benefit in research procedures. For instance, it is claimed that programs provide more reliable results faster and more easily than manual analysis. Software packages are said to make qualitative data analysis techniques more sophisticated. In addition, computer use is understood to improve the position of qualitative research in general. The evaluation of these arguments, however, tends to be left mid-way in terms of empirical research practices. Textbooks mostly offer only a general discussion on qualitative methodology or a heavily technical evaluation of different software applications (e.g. Weizman & Miles 1995; Tesh 1990; Luomanen & Räsänen 2000). It is therefore necessary to address the actual methodological advantages that software is argued to bring in more detail.

The aim of this paper is to examine the emergence of computer use in qualitative research and to evaluate its impact on research practices. On the one hand, I wish to investigate the current trends of computer assistance. On the other hand, I shall discuss whether computer assistance has influenced the way in which we define good qualitative research. Can software applications have an impact on the issues of validity and reliability of qualitative data analysis, and can they really improve the quality of research?

First, the paper traces the nature of qualitative approaches in social science. The essential criteria of qualitative research are examined in light of qualitative and quantitative research models. After that, more structured qualitative approaches and the use of computers in data analysis are discussed. It is argued that there is a connection between structured qualitative approaches and the development of computer programs. Also, the differences between different methodological approaches and computer-based techniques are scrutinised. This is followed by a discussion on the use of software applications and the general progress of qualitative research. Particularly the features that are said to have improved as a result of software use in data analysis are specified.

Conceptualising quantitative and qualitative approaches

Two mainstream methodological approaches in social and behavioural sciences can be distinguished by referring to a taxonomy of quantitative and qualitative methodology. During the past decades, there have been several debates over the relative merits of these approaches. Qualitative and quantitative approaches are at least partly based on opposing paradigms or belief systems, typically labelled *positivism* and *constructivism* (Raunio 1999, 18-19; Newman & Benz 1998). Reflections regarding the superiority of one or the other of these paradigms can still be seen in current literature (e.g. Tashakkori & Teddlie 1998, 3-4; Alasuutari 1994; Töttö 1999). However, I do not wish to process this debate in greater detail because it is not very useful for the purpose of this paper. Quantitative and qualitative approaches can be broadly outlined and distinguished from each other also in terms of analytical principles.

In general, quantitative research models can be referred to as hypothesis testing research that typically contributes scientific knowledge through theory testing (Newman & Benz 1998, 18). Of course, hypothesis generation is included in most research procedures, and many data-driven techniques have been developed. Despite this, it should be mentioned that the explorative techniques contain certain methodological assumptions which have an impact on the analysis process (see Tashakkori & Teddlie 1998, 112-115). For example, it is typically assumed in cluster analysis (and in other similar techniques) that all single variables are equally relevant to all cases in the study. Because of such presumptions, certain hypotheses are usually formulated before actual data collection and analysis begin.

The aim of quantitative research is to provide generalisability from the sample to the population, that is, from a small group of individuals to other individuals who are similar according to certain properties. Quantitative research design thus includes the controlling of variables, randomisation of subjects included in the study, and the requirement of reliable measures. It provides established rules that will guide the researcher in different phases of research. Sampling theory, for instance, is part of the basis for the whole research process, for the collection and analysis of data.

Data are typically collected by using structured survey or interview questionnaires. The objective is to record and code data according to a certain *a priori* operational and standardised definitions. After that, the selected variables are processed in numerical, or at least in categorical, format. There are also more or less unambiguous rules for hypothesis testing with statistical techniques, as there are for the presentation of tables and figures. Data analysis is based on the measurement of the dependent variables and controlling for the effects of the independent variables. In brief, this means that research hypotheses are accepted or rejected by calculating the required test statistics,

such as *Chi-square* or *F* ratios, and the determination of their significance. The analysis procedures and results are also expected to be reported by referring to the statistical significance of the measures (e.g. Toivonen 1999, 80-84; Alkula et. al. 1994).

Moreover, the guiding rules and methodological solutions of quantitative analysis are in a way universal. The same methods and techniques of analysis can be often utilised in the same way across various data sets and research problems. In principle, only the measurement scales of variables and number of observations restrict the use of different statistical measures. In this way the procedures of quantitative research can be seen as deductive in nature, when tightly controlled situations of true experimental designs become imitated.

Quantitative designs have been the dominant research models in social sciences (Newman & Benz 1998, 10; 15-16; Raunio 1999). Yet the use of qualitative material and qualitative research methods have increased in popularity over the last decades. This can be observed, for example, in the base of books, research articles and academic dissertations of many disciplines. There is a growing recognition among the researchers in all branches of social sciences that surveys and register-based data sets do not provide all the relevant information that can be available by using qualitative techniques (e.g. Bacon 1993, 1; Strauss 1987, 3-4). But what are the generic definitions of qualitative research design and how do they differ from those of quantitative design?

As a general orientation, the qualitative approach fails to present standardised or distinct rules for analytic work. In other words, when compared to quantitative research designs, methodological openness can be seen as one of the essential criteria of qualitative research. Although there is a lot of variation in different designs, many basic aspects of qualitative data analysis can be understood broadly in light of the methodological principles of the Chicago School and social anthropology.[1] All major perspectives emphasise, in one way or another, the phenomenological nature of social life.

This can be seen, for example, in the ways in which the data are generally defined. Qualitative data are characterised as detailed descriptions of examined events, quotations from people about their experiences and attitudes, or different records and case histories. Very diverse materials provide data for qualitative analysis, but usually materials are represented in textual (or linguistic) form. This has many implications for the

[1] It should be stressed here that there certainly is no one kind of qualitative data analysis, like there is no one kind of quantitative data analysis. There is a variety of approaches, related to different perspectives and research purposes. Renata Tesch (1990, 58), for instance, has distinguished over forty types of qualitative approaches. However, distinguishing the different perspectives completely would be a fruitless task because the boundaries between different approaches are far from clear. This question is even more problematic in relation to what researchers do in actual research procedures. This is why I wish to concentrate on a basic core of qualitative research, which can be broadly characterised as intepretivist and ethnographical frameworks to patterns of culture or social settings, and diverse personal interpretations related to them.

nature of analytic procedures. Above all, the understanding of the meanings and meaningful actions becomes the most important aspect of research. It follows that the meanings of the examined phenomena, for certain people or culture as such, are usually elaborated upon in the analysis (Newman & Benz 1998, 16-17).

Also, the place of theory in qualitative analysis can be seen differently from that of quantitative analysis. The starting point of data analysis is not typically a certain theoretical model or a clearly formulated hypothesis, but merely the researcher's own intuition. Research is carried out by entering the fieldwork (or reading task) and describing what happens. So, instead of coming from the conceptual to the empirical level, qualitative analysis begins at the level of data collection and ends more or less at a conceptual level. This requires the construction of unique methodological solutions, which are hard to reproduce or adapt to the analysis of other research materials. Of course, the analysis procedures can not be totally inductive, but theoretical ideas are conceived to emerge and are developed on the basis of the data. It is understood that interpretations are guided primarily by the characteristics of data.

Qualitative research designs do not necessarily require control or reference groups, or even information outside the sample of subjects chosen for the analysis. Researchers have always worked with rather small samples and the very limited accounts of individuals. Only one subject, one case, or one unit may form the focus of a whole investigation. Under such circumstances, generalisability outside the sample is not seen as a problem (e.g. Raunio 1999, 327-328; Alasuutari 1994). There are also various more variable-oriented qualitative research designs, which may include, for example, theoretical sampling and comparison of multiple cases (Tashakkori & Teddlie 1998, 117-118). Still, the aim of research is merely to understand the studied phenomena against their unique features, not to produce a representative view of the population under study.

Traditionally, qualitative methods have referred mostly to the ways in which data have been collected. Analytic considerations have been transmitted to the audience tacitly rather than explicitly (Strauss 1987, 4-5; Töttö 1999, 281). One could claim, for instance, that someone has used structured or in-depth-interviews as his or her "research method". This kind of methodological thinking has not provided clear guidelines for the actual data analysis procedures, or presentation of research results. Data have been analysed freely, without strict technical rules or standardised measurements. In qualitative analysis data are thus coded *a posteriori,* from the subjective interpretations of that data. The examined phenomena are approached holistically, trying to maintain the richness and depth of meanings related to the research problems. It follows that the only principles that could actually guide the researcher's work are speculative and heuristic in nature, involving a vision of "lived reality" (Newman & Benz 1998, 13).

A coarse schema of quantitative and qualitative research approaches appears in Table 1. Approaches are assessed in the table on the basis of methodological principles and research practices. One approach possesses certain common qualities that the other lacks. In general, I think these differences can roughly illustrate the key features of qualitative approaches from the researcher's perspective.

Table 1. Generic definitions of quantitative and qualitative approaches

Quantitative approach	**Qualitative approach**
●Theory testing	●Theory building
●Deductive	●Inductive
●Standardised rules for data analysis	●No explicit rules for data analysis
●Begins with theory	●Ends with theory
●Methodological solutions are universal and easily reproduced	●Methodological solutions are not easily reproduced

Compared to quantitative analysis, qualitative data analysis can be characterised as non-standardised, flexible, and more or less vague. In other words, the researcher's work is based merely on inductive logic, and there are no explicit rules guiding the data analysis. It is hardly surprising that in addition to the limits of generality of information, for example, the validity of qualitative techniques have been questioned from many viewpoints (see Kelle & Laurie 1998, 21-22; Bogue 1993, 7). Qualitative research has been most widely criticised, however, on the basis of its analytic vagueness.

There have been suspicions that, in practice, data analysis involves little more than stringing together quotes extracted from notes or transcripts. In many cases, especially when a lot of research material is used, it is surely impossible to tell how interpretations from the data were made. The findings culled from the data may be based on rare or exceptional cases, which are not generally represented in the data. Some critics have even questioned the sensibility of qualitative research from this basis (e.g. Toivonen 1999, 14-15; 113-14; see also Strauss 1987, 5; 26-27). Of course, these kinds of claims are meant to be mostly provocative. But one can still understand the criticism behind them: If there is no explanation for the practical procedures that were followed in the data analysis, the validity of results cannot be trusted. It is obvious that theoretical concepts do not emerge independent of the person interpreting the data. Data do not develop abstract ideas, only people do. In this way it should be the primary task of all research methodologies to clarify the practical procedures on which the interpretations are based.

Many alternative and more detailed designs for qualitative research have been presented over the years. Especially the approaches that can be labelled structured ap-

proaches yield interesting principles, considering the methodological criticism of qualitative data analysis. How do these approaches differ from the basic perspective on qualitative approaches?

The emergence of structured approaches in qualitative research

From time to time there have been earnest discussions on the criteria that are applied to assess the validity of qualitative research in social sciences. As mentioned above, qualitative researchers have thought that there should not be strict criteria for describing and standardising qualitative research procedures. Qualitative analysis has been understood to be very different from quantitative analysis, and this is why it has not been considered important to develop explicit rules for data analysis techniques. In this way only the impressionistic nature of qualitative data analysis and the role of the researcher as "a creative problem solver" or "an explorer" (e.g. Alasuutari 1994, 44-45; Eskola & Suoranta 1998, 20-21) may be emphasised as prerequisites for good research.

Many qualitative researchers still believe that no explicit methodological guidelines for analytical work are necessary. In Finnish literature, for instance, qualitative research has been considered as an art form in which rather intuitive approaches to data analysis are preferable (e.g. Roos 1989, 143; see also Raunio 1999, 293-294). Nevertheless, many procedures or "rules of thumb" have been presented to systematise qualitative data analysis and make the analysis techniques more standardised and transparent. Over the years, a lively discussion has surrounded *content analysis*, *grounded theory* and *discourse analysis*, as well as other detailed narrative approaches (e.g. Fielding & Lee 1998, 28-32; 45-46; 52-54; Alasuutari 1994). During the last twenty years or so, it is surely possible to see a trend towards more detailed approaches in qualitative research in social sciences.

In general, many alternative approaches to qualitative data analysis can be characterised as more structured (or systematised). While there are differences in the methodological application of different structured approaches, their aim is to provide explicit rules for the collection and/or analysis of research materials as well as to make theoretical interpretations from the data (Newman & Benz 1998, 17; Strauss 1987, 12). Grounded theory represents perhaps the most widely known and best clarified structured qualitative research design or type of data analysis (Fielding & Lee 1998, 28-29; Lonkila 1998, 41; Miles & Huberman 1994). The approach was originally presented by Anselm Strauss and Barney Glaser in the late 1960s (Glaser & Strauss 1967). While there have been many efforts to clarify and modify the original ideas (see Strauss 1987, 5-6; Strauss & Corbin 1990, 112-113), the basic design has not changed remarkably.

If the principles of grounded theory are applied to data analysis, they can be roughly described as follows:

"(T)he researcher starts by reading and carefully analysing a small amount of data. He or she codes (read: analyses) the data (most often text) by following very detailed and complex procedures and "rules of thumb". During the analysis the researcher is continually asking questions about the data and checking them by constantly comparing different instances of data. (—) During the whole research process the researcher writes "memos" on, for example, his or her ideas about codes, their interrelations, new directions for the research, etc. (—) Continuous interaction between the collecting and coding of data and the writing of memos is essential. As the research advances, the researcher develops an increasingly abstract and complex conceptual structure, specifying the connections between the concepts of the emerging theory and regularly returning to the data to check whether this theoretical structure is in fact supported by data." (Lonkila 1998, 42)

The grounded theory approach aims to present a systematic way to interpret qualitative data, and its particular emphasis is on how to develop theoretical ideas through analysis (Strauss 1987, 22-23; Glaser & Strauss 1967). Data analysis is based on a very detailed and explicit coding of texts, which involves the discovery and naming of categories, as represented in the quotation above. Abstraction and theorising processes are very central in qualitative data analysis in general, because researchers usually wish to make some kind of general claim on the basis of their data. This is typically done through coding, which involves creating abstract categories from the data (or text) segments and specifying their properties, dimensions and interrelations. Data chunks with relevant information are assigned (or simplified) to a conceptual category or categories. In other words, coding refers to the act of attaching words that describe certain meanings to text segments.[2]

Perhaps the greatest value of the grounded theory approach lies in the fact that it provides various systematic guidelines for data parsing, and for the qualitative research process as a whole. Hence, it at least sets certain unambiguous requirements for analytic procedures that would otherwise remain without clarification. I think that separate, and clearly explicit, sets of methodological criteria for qualitative research are required. This is why structured guidelines or "rules of thumb" outbid easily the rhetoric that emphasises "lived reality" or "unique, but simultaneously undeniable" interpretations of the data. However, it should not be assumed that the grounded theory approach is able to lay down strict rules for conducting social research. Like qualitative data analysis techniques in general, it by no means can be regarded as a set of fixed rules for

[2] Of course, different formal meanings are given to coding within different approaches of qualitative research. Some approaches define coding as attaching keywords to text segments, and some others more broadly as the process of analysing data (see Glaser & Strauss 1967, 111; Strauss 1987, 55-56; Dey 1993, 114-115). Nevertheless, in one way or another, coding always takes place when the researcher is trying to develop abstract ideas from the data and wishes to classify, interpret or categorise them.

developing data into theory or final results. From the perspective of research procedure, as an interpretative analysis technique of unstructured qualitative data, it is similar to any other data parsing technique.

Grounded theory can only constitute guidelines that can help researchers to keep in mind the constraints and challenges of research settings and research aims. It is not a kind of "laundry list" of suggestions to data analysis; the guidelines it provides simply emphasise and remind that certain operations must be carried out (Strauss 1987, 8). Of course, many other collections of systematic advice apart from grounded theory can be utilised in qualitative data analysis. But without certain principles or guidelines, it can be extremely difficult to establish research validity. And if there is no validity that is established through a set of rules that are applicable across different data sets (or texts), it inevitably becomes very difficult for a researcher to justify his or her conclusions to the reading audiences. In addition, with implicit or non-systematic analysis techniques, it is just as difficult to test the presented arguments again.

The problem of validity in qualitative research is to find a way to draw from the data valid meanings and results that others can rely on (Newman & Benz 1998, 18; Töttö 1999, 289-290; Bogue 1993). I believe that one central difficulty is that the methods of analysis have not been well formulated. Grounded theory and other structured approaches aim to make explicit the research procedure by describing data analysis techniques concretely. Yet it should be stressed that qualitative research is quite unlike quantitative research. The analysis techniques cannot be based on the same (mathematical) models and measurements as statistical techniques. Data are nearly always in the form of words rather than numbers. Hence, the standards for good qualitative data analysis surely must differ from traditional (quantitative) definitions of research validity and reliability (e.g. Raunio 1999, 362; Newman & Benz 1998).

One could criticise the structured approaches by claiming that the true characteristics of qualitative research principles shall be compensated by quantitative ones, if the guidelines of data analysis become too explicit and systematic. This would lead researchers to analyse qualitative data quantitatively (Roos 1989, 142-143). For example, the richness and depth of meaning of studied phenomena may be sacrificed in too-controlled analysis situations. One may find it tempting to relay only on quantitative information and aspects, if the processed data have been coded, say, according to multidimensional hierarchical structures. However, there is no reason to believe that systematic construction of categorisations would necessarily lead to this.

In general, systematic strategies can only guide to browse data in a sensible manner, they certainly do not urge researchers to abandon methodological eclecticism. Of course, accepting an explicit methodological principle may be a problem for those who believe that the requirements for openness and credibility are important principles only for

quantitative research. But one can argue that they should concern all scientific research. When methodological principles and techniques are clarified, it is possible to have at least a presentiment of how preliminary qualitative field notes will actually develop into the final conclusions (Mäkelä 1990, 57-59). Otherwise it is nearly impossible to tell whether the conducted analysis makes sense or not.

So, the emergence of structured approaches should be understood, above all, as an attempt to make qualitative data analysis more transparent and more systematic. Structured approaches present a clear methodological standard for analytic considerations. The use of computer software in qualitative data analysis is related to similar attempts. It has been noted that there has been a kind of second coming of structured qualitative analysis during the 1990's, as software packages have become more frequently utilised by social researchers (e.g. Miles & Huberman 1995, 2; Luomanen & Räsänen 1999). In this way a temporal connection can be pointed out between the two. But in what other ways is computer assisted analysis connected to structured approaches of qualitative research?

Computer-based analysis as structured qualitative analysis

Computer-assisted qualitative data analysis includes the use of software packages especially designed for qualitative data analysis. Basically, computers replace manual tasks in data management and analysis procedures. The researcher records and manages his or her data using a certain analysis program.[3] The analysis programs provide several tools for coding, retrieving, and linking different segments of data. What is usually called a project is created in order to integrate the data and all the operations and modifications related to this data. In short, in computer assisted analysis, the data analysis is carried out within the user interface of an analysis program.

There are at least several dozens of programs that researchers can use to support qualitative data analysis. Since the early 1980s, three to five major program families or genres can be clustered according to program capabilities and limitations (see Weizman & Miles 1995, 16-22; Luomanen & Räsänen 1999, 6-7, for more detailed descriptions). The earliest applications were basically text retrievers that provided tools for locating and counting words from text files and were able to create word lists and simple concordances. Next, in the late 1980s, came programs that could help research-

[3] Several of the tasks of qualitative data analysis can surely be done by computer, even if there is no specific software available. Writing up and editing field notes or interviews, for example, may be covered by a word processor. It is possible to argue that this, or even the writing process of a research report, represents computer-assisted analysis. However, approving these kinds of arguments would make the term 'computer assistance' very inaccurate. This is why I wish to restrict the term to the utilisation of specific software packages.

ers in coding procedures, in other words, in dividing texts into segments, attaching codes to these segments, and finding all the instances of coded segments. Around the same time several programs with developed memos and numerical matrix-building capabilities came out. The latest application family, introduced in early and mid-1990's, provided tools for conceptual or semantic network building, which meant that researchers were able to link coded categories to other categories with specified relationships and values. These kinds of "theorising" capabilities, easy linkages among text and codes, and graphic displays of linkages, have become central features in many new programs.

Programs have been developed for different purposes over the years. This is still the case: the functions of a single software package are best matched for certain approaches. However, the most advanced programs are nowadays suited to versatile analysis purposes — from a simple word frequency count to the development of a conceptual coding structure, and displaying it graphically. At a more detailed level, each program naturally performs the same functions a little differently and has a more or less unique configuration of features. This is why only a few examples, the latest versions of *NUD*IST* and *ATLAS/ti*, are referred to and discussed below.[4] In spite of this, the actual analysis facilities described later are also provided, at least in limited ways, by many other available software packages.

Structured qualitative approaches, grounded theory in particular, have had a strong influence on the development of contemporary qualitative analysis programs. This is apparent in the structure of many programs, in the publications of program developers, and in the software manuals (Lonkila 1998, 46). For example, the user's manual of the NUD*IST Vivo includes numerous references to the analysis techniques of grounded theory, content analysis and discourse analysis. The functions of the program and the ways in which to do particular tasks with it, are very often described by referring to the methodological principles of these approaches (see Richards 1999, 15; 29; 47-48; 137; 165-166). In the case of an another popular software package, ATLAS/ti, an even more explicit connection to structured approaches can be found. Anselm Strauss, the co-developer of grounded theory, has actually written the foreword for the user's manual of ATLAS/ti version 4.1 (see Strauss, 1997).

Many analytical operations can be seen as central to both the grounded theory approach and computer-assisted data analysis. Especially the more or less structured coding of data, linkages between text segments and codes, writing of memos and the verification and checking of ideas developed from the data are key features in many programs as well as in grounded theory procedures (see Lonkila 1998, 46-49). To the

[4] This decision might be argued by referring to the user-friendliness and versatile functions provided by these programs. For example, the latest version of ATLAS/ti and NUD*IST Vivo support, not only the analysis of text materials, but also the on-screen processing of multimedia files. In addition, since the mid-1990's, NUD*IST and ATLAS/ti have been published by a major international distributor, which has increased their visibility and accessibility. Also the recent methodological literature on computer-assisted qualitative research discusses them as a rule.

same extent, a connection to discourse analysis and content analysis techniques is apparent in many programs. Interview data, for example, can be typically organised and coded in the most effective way according to the sequential structures of speech. This is one of the most widely used codification strategies of texts in narrative approaches (Fielding & Lee 1998, 46; Alasuutari 1994).

A connection can also be observed in the structure of search tools, provided by most software packages. The specification of text searches, among other things, is very often based on the idea of "grep", or regular expression of text parsing, which is one of the essential techniques in content analysis. Particularly in enumerative strategies, frequencies of certain words are understood to reflect an assumption of their salience and importance for topics being covered. Because of computer-based tools, these kinds of techniques have become much more effective, and more widely used among qualitative researchers (Weitzman & Miles 1995, 17; Dey 1993).

Regardless of the above, the use of programs does not force the researcher to conduct an analysis in line with grounded theory or discourse analysis methodology. They can be, and mostly are, used for other purposes (Luomanen & Räsänen 1999, 10; Lonkila 1998). Still, it is reasonable to suppose that the programs can be best fitted to the methodological framework of certain structured approaches. This can be assumed simply because particular approaches were seen as functional models by many software developers (Weitzman & Miles 1995, 333; Lonkila 1998).

Of course, many tasks differ from each other when moving between the "rules of thumb" provided by structured approaches, and what can be done, and how, with the tools provided by different software applications. Programs themselves are not committed to specific theoretical traditions in a way that methodological approaches usually are. It follows that software packages do not provide any methodological advice for the researcher. They do not make any suggestions on how operational problems should be solved, or what should be emphasised in the interpretation of data. What they provide is only a certain kind of tool-kit for data analysis purposes. In other words, programs are not methods (or methodological approaches), although there have been misunderstandings about this at least in recent Finnish literature (e.g. Toivonen 1999, 15; 135).

However, I believe that computer-assisted analysis techniques do have an important methodological impact on the research process. These techniques can be said to involve more than just reading, underlining, cutting and pasting of piles of paper. In other words, when operations conducted according to the functions of certain program, this makes possible a more organised data analysis that is easier to describe to other researchers. In this way analysis programs may be effectively used to reveal the guidelines of the analytic processes in research to the reading audiences. Table 2 lists some of the central phases of data analysis, where computer programs can be used to make explicit the analysis procedure.

Table 2. Central phases involved in the qualitative data analysis procedure

- Managing and splitting research materials (text-base management)
- Associating and linking different parts of data (text segment classifications)
- Marking and annotating data (memos)
- Labelling, arranging and categorising data segments (coding)
- Connecting created categories to each other (conceptual networks of created categoies)
- Illustrating and visualising of categories (graphic models)
- Finding support for the created theoretical ideas from the data (search operations)
- Preparing a research report (graphic models, matrix tables)

The program stores each change made to data and records all the created links between the data and developed categories. It follows that the researcher is able to manage data, create indexes from the data and handle the modifications in a way that everything will be stored systematically by the computer. He or she can thus easily estimate what kind of progress there has been in the data analysis. There are also various facilities available in the advanced software applications that assist the researcher in discovering and developing ideas from the data. Different documents can be linked to each other, and memos can be easily created almost for any purpose and attached to documents or created categories. In most contemporary programs the coding system can also be presented in graphic form. Further, the interrelations of created categories can be tested and integrated into the data by conducting versatile search operations. This brings a feeling of "closeness to data", which is highly valued by many researchers.

So, if data are processed using an analysis program, it follows that one can estimate whether data analysis has been detailed enough. For example, when reliable and stable coding is applied, computer software offers many facilities to collect and retrieve relevant information considerably easily. It is also easier to discuss the methodological problems related to data modification, if one can show afterwards what really was done. Print-outs, say, from different stages of the coding process can be utilised in the appendix of the research report. Also, code and word frequencies can be calculated from the data and results can be presented in matrix tables. Software provides various ways of recording memos in order to annotate the present state of theoretical or methodological understanding associated with certain data segments.

In other words, analysis programs offer tools that could make it possible to trace the paths by which particular ideas have emerged. In addition to this, the computer environment can offer significant assistance in following certain methodological principles. One can imagine that it is far easier to process each chunk of data, for instance by following the coding phases provided by grounded theory, if the documents are automatically managed in alphabetical order — and not just sorted with coloured pencils and little yellow labels.

Some of the uses of the computer as an analytic tool in qualitative research were briefly described above in order to distinguish computer-assisted analysis techniques from methodological approaches, and to illustrate how application software might be utilised in the actual data analysis. It could be said that programs provide a choice of possibilities, quite similar to structured approaches, to weaken the stigma of analytic vagueness that is often attached to qualitative approaches. Under the circumstances, computer assistance can, at least in principle, ensure more a transparent qualitative data analysis process. This may increase the scientific status, or face validity, of qualitative research in general. But can computer assistance influence the way in which we understand the nature of qualitative research? Is it possible to argue that there could be methodological progress or evolution in qualitative data analysis as a result of computer use?

Towards a higher state of qualitative data analysis?

It has been thought that computer programs could be misused as purely rhetorical weapons, used to convince readers of the scientific nature of one's research (Lonkila 1998, 46; Fielding & Lee 1998, 72-73; Luomanen & Räsänen 1999). For example, simply mentioning a certain software package in a research report may work as a certificate of, or even as a deterrent to, valid and highly sophisticated analysis to some readers. Similarly, the potential and practical benefits of computer assistance can sometimes become exaggerated. While software users tend to stress the accomplishments of computer-assisted analysis techniques, they may simultaneously fail to see the comparable possibilities of "traditional" manual techniques.

It is true that the use of computer software in data analysis does not itself guarantee that data are actually analysed properly. The researcher can, as easily as in manual analysis, make poor interpretations from the data. In many cases the justifications for using computer software in research practices are tied to the character of the research materials and research questions. The possible restrictions and limitations are typically related to questions of modifying and managing data in text format, or utilising a certain program in particular variable- or case-oriented methodologies. While being aware of problems related to these questions, I think it is possible to see several, more or less general, advantages to computer use.[5]

When using an analysis program, it is far easier, when desired, to illustrate what has been done with the data during the research process. As mentioned above, using a

[5] Yet this is not a simple question between clear and unambiguous advantages and disadvantages. Benefits and drawbacks of software use might be experienced differently by the same user in the context of different research questions. Further, how the actual methodological advantages software packages are registered is often far beyond one's theoretical framework or research materials. Some researchers simply like, while some dislike, using them.

program encourages the researcher to produce an analysis that is explicit, systematic and well-documented. It is also easier to handle text materials, in extensive or small amounts, as digital files rather than as paper piles. It follows that accessibility to the data increases, and the analytic process is more likely to be laid bare to critical inspection. In the same way, computer use assists the management of larger samples of data. Seeing and browsing the data, codes and categories simultaneously on the computer screen can practically force the researcher to think about the data analysis as an integral part of the research process. The researcher is able to monitor and check each analysis operation and estimate whether he or she is following certain methodological principles systematically or in enough detail. I consider this one of the main advantages of computer software use in relation to manual techniques of data parsing.

In a way then, the use of computers could already influence the ways in which qualitative researchers should handle their data. The detailed documentation of the analysis procedures, which is central in computer assisted techniques, makes it possible to address the requirement of analytic reproducibility from a new angle. As is known, qualitative data analysis might be seen as a mysterious, unique and inevitably a non-reproducible process that focuses primarily on improving the technique of data collection. But now, detailed clarification and replication of data analysis procedures become real possibilities.

Computer use requires users to be explicit and clear about what they are doing. Those who use computers are generally more used to logging the trails of analytic procedures, for example, by continually writing memos, using document annotations and writing up relations between different data segments. This can make it easier for another researcher to replicate an existing analysis (e.g. Fielding & Lee 1998, 57-58; Kelle & Laurie 1998). For example, if the used data sets are available, the second researcher should be able to inspect whether the claimed results are actually entitled. He or she can also analyse his or her own data following the same coding phases and search techniques by using the same software. By doing this, other researchers are able to estimate the methodological sensibility and validity of the chosen analysis techniques in each other's research projects.

Secondly, as discussed above as well, computer use has brought qualitative data analysis techniques closer to structured analysis. It is far easier to process data sets systematically when the recording of codes, memos and annotations are linked to documents or their (text) segments electronically by computer. Hence, the created categories can be retrieved separately or together with data segments in a way that makes their interrelations visible throughout the analysis. This is to say that computer assistance increases the reliability of qualitative findings because software facilities ensure that the hypotheses and ideas developed are really grounded in the data (e.g. Kelle & Laurie 1998, 27; Weitzman & Miles 1995). In other words, researchers are able to show that results are

not based on single and atypical incidents in the data. These features, too, indicate expanded possibilities for the reproducibility of data analysis procedures, and reflect a trend toward methodological sophistication in terms of research validity and reliability.

In addition, because computer-based analysis is typically quite systematic in nature, there are also advanced possibilities for using the analysis procedure as a gateway to statistical analysis. Traditionally there have been no effective tools available for the quantitative analysis of small and purposefully selected samples. Quantitative content analysis has represented perhaps the only way of analysing sets of textual data quantitatively. However, since the development of computer software for qualitative data analysis, some new tools have been made available to reconcile the different methodological approaches.

Alone, compatibility with the statistical packages may be seen as a promising and useful option. Most popular qualitative data analysis programs support the tab-separated format of *SPSS* and *Microsoft Excel* in their numerical displays of search results, document reports or coding structures. In this way numerical information on the created categories, for example, can be exported to a statistical package for further analysis. Some simple facilities for code and word frequency counting are also enabled in most advanced programs. For multi-methodologically oriented researchers, these functional features provide feasible tools for combining qualitative and qualitative methods across different phases of the research process.

Frequency distributions of code patterns might be found very helpful in research projects in which data have been collected from various sources, or that include comparisons between many individuals, groups or cases. Significant correlation could be found, for instance, between the number of certain codes relevant to the research questions and the research groups' background information. These kinds of findings usually help researchers in asking new questions and framing the further analysis. In some cases it is possible to view quantitative findings as results themselves. By using software, researchers may be able to analyse the kinds of dimensions from their data that might otherwise not be taken into consideration at all.

Many examples of quantitising the qualitative data have been presented in literature (see Tashakkori & Teddlie 1998, 128-129; Miles & Huberman 1994). Critics of computer-assisted qualitative research have claimed that the use of a computer may encourage researchers towards analytic madness, to analyse their data, above all, quantitatively (Seidel 1991, 109-110; see also Fielding & Lee 1998, 78-80). Similar criticism has been focused on structured approaches more generally, as discussed earlier in this paper. However, the examination of quantitative dimensions of data should be understood only as a possibility, not as a standard phase of research procedure. There seems to be no real evidence that software use would lead researchers to this, at least without

their notice. On the contrary, it may be said that the differences between the two aspects of phenomena being studied, quantified and textual, become clearly recognised and reflected when working in a computer environment. For example, search operations are typically designed to be used differently if search results are displayed in numeric instead of textual form (Luomanen & Räsänen 2000, 144-148; 155-157). Therefore, the differences between qualitative and quantitative approaches are more likely to be preserved, rather than integrated, during data analysis.

I believe that the advantages of software use, discussed above, are bringing a new kind of methodological thinking to the field of qualitative research. It may not be a question of radical change, but computer-assisted qualitative data analysis seems to support more explicit and well-documented analysis in general. In spite of this, it is still possible to contradict the argued methodological progress completely. One can plausibly claim that computer assistance is not taking qualitative research in a new direction. Rather, it is causing qualitative analysis to revisit many old debates (Fielding & Lee 1998, 181). The controversial question related to structured and unstructured analysis techniques can clearly be seen as an example of an older debate. Also, the opposing opinions concerning computer use can be approached from a similar viewpoint. Those who are cautious about systematic methodological approaches are probably more likely to criticise computer assistance as well (Kelle & Laurie 1998; Luomanen & Räsänen 1999). In this way, it is surely highly problematic to make claims about current methodological development in terms of evolution, or as progress toward a higher state of data analysis.

While noting the problems, it can be argued that the utilisation of computer software in qualitative data analysis offers something new also in a methodological sense. As has been discussed, qualitative research might benefit in general from a greater explicitness about analytic procedures. In computer-assisted data analysis, the transparency of operations becomes a rule even if the researcher has not methodologically committed to any established structured approaches. It follows that when the use of software becomes more commonplace, methodological discussion and conversion of ideas among qualitative researchers also should become easier.

There are probably numerous influences, but it is reasonable to argue that the current technological developments and their utilisation for research purposes can contribute considerably the principles by which qualitative data analysis are appraised. It may well be that several new sets of criteria for judging analytic operations shall be constructed in the near future. At least it is very likely that research ideologies approving any kind of data parsing, whether systematic and detailed or not, will soon lose their popularity, and perhaps even their acceptability, as scientific approaches.

Summary and discussion

It was argued in this paper that computer-assisted data analysis techniques can provide possibilities for developing the often impressionistic qualitative analysis procedure into something more integrated, explicit and systematic. It was also claimed that the use of software packages might help researchers to take the validity and reliability issues of research into more careful consideration.

Computer use was examined from a methodological framework and from the viewpoint of concrete operations of data analysis. Methodological aspects of computer assistance were found quite dissimilar to a basic perspective on qualitative data analysis, but many connections between structured approaches and computer-based techniques were found. It was suggested that computer-based data analysis is best conducted by following certain structured methodological principles. However, it is not impossible to use software for other purposes. They do not force users to work according to particular guidelines, but they do urge them to conduct a systematic and properly documented data analysis. In this way computer use could increase the openness and transparency of analytic procedures in general.

The systematicity of computers enables easier data management throughout all phases of data analysis. However, there are many issues that were not actually discussed in the paper but that certainly affect shifting from manual to computer-based analysis or choosing a suitable software package. For example, the kind of project the researcher has and the databases he or she is using should be taken into consideration. Different programs tend to be good for different kind of analyses (e.g. Weitzman & Miles 1995; Fielding & Lee 1998, 129-130). Naturally, the more general issues related to research problems, analysis styles and methodological approaches should also be considered carefully. In addition to such questions, it is important to understand that some qualitative researchers might find a computer environment more or less awkward even if software packages have become quite user-friendly during the last years. Therefore, it is important to consider what kind of a computer user one is. It has been said that computer-based analysis is probably best suited for those whose present level of computer use is already high, and who are generally motivated to learn new working techniques (Luomanen & Räsänen 2000, 169-170; Fielding & Lee 1998, 69-70).

Some people may also have unrealistic expectations of software. Even though some of the restrictions related to computer use were mentioned above, it should be stressed that there is not, and arguably never will be, a computer program that will analyse your data for you. In other words, the quality of data inspection, coding procedures and final conclusions will always depend upon the researcher's personal input and methodological capabilities.

In sum, then, computer-assisted qualitative data analysis should be understood to represent only an attempt to make the analysis procedure explicit by offering a special toolkit for data processing. Software development can contribute to the progress of qualitative analysis methods in general, but from the point of view of a researcher, computer use surely does not represent a methodological principle in itself. By contrast, however, computer programs provide facilities that can help qualitative researchers to follow and develop the procedures involved in analytic considerations.

References

Alasuutari, Pertti. 1994. *Laadullinen tutkimus* (Qualitative research). 2. painos. Tampere: Vastapaino.

Alkula, Tapani, Seppo Pöntinen and Pekka Ylöstalo. 1994. *Sosiaalitutkimuksen kvantitatiiviset menetelmät* (Quantitative methods in social research). Juva: WSOY.

Bacon, Jean. 1993. Editor's introduction. In: *Readings in population research methodology,* edited by Donald J. Bogue, Eduardo E. Arriaga, Douglas L. Anderton and George W. Rumsey. Vol. 6. Advanced basic tools. Chapter 24, pp. 1-2. New York: United Nations Population Fund.

Bogue, Donald J. 1993. Role of qualitative method in demographic research. In: *Readings in population research methodology,* edited by Donald J. Bogue, Eduardo E. Arriaga, Douglas L. Anderton and George W. Rumsey. Vol. 6. Advanced basic tools. Chapter 24, pp. 2-7. New York: United Nations Population Fund.

Dey, Ian. 1993. *Qualitative data analysis. A user-friendly guide for social scientists*. London: Routledge.

Eskola, Jari and Juha Suoranta. 1998. *Johdatus laadulliseen tutkimukseen* (Introduction to qualitative research). 2. painos. Tampere: Vastapaino.

Fielding, Nigel G. and Raymond M. Lee. 1998. *Computer analysis and qualitative research.* London: Sage.

Glaser, Barney G. and Anselm L. Strauss. 1967. *The discovery of grounded theory: strategies for qualitative research.* New York: Aldine

Kelle, Udo and Heather Laurie. 1998. Computer use in qualitative research and issues of validity. In: *Computer-aided qualitative data analysis. Theory, methods and practice*, edited by Udo Kelle, pp. 19-28. London: Sage.

Lonkila, Markku. 1998. Grounded theory as an emerging paradigm for computer-assisted qualitative data analysis. In: *Computer-aided qualitative data analysis. Theory, methods and practice*, edited by Udo Kelle, pp. 41-51. London: Sage.

Luomanen, Jari and Pekka Räsänen. 1999. Tietokoneavusteinen laadullinen analyysi — kehitys ja nykytilanne (Computer-assisted qualitative analysis — development and current situation). *Kulttuurintutkimus* 16(4):3-12.

Luomanen, Jari and Pekka Räsänen. 2000. *Tietokoneavusteinen laadullinen analyysi ja QSR NVivo –ohjelmisto* (Computer-assisted qualitative analysis and the QSR NVivo software). Sosiologian tutkimuksia A 23, Sociological research. Turku: Turun yliopisto.

Miles, Matthew B. and Michael A. Huberman. 1994. *Qualitative data analysis. An expanded sourcebook.* Second edition. London: Sage.

Mäkelä, Klaus. 1990. Kvalitatiivisen analyysin arviointiperusteet (The evaluation principles of qualitative analysis). In: *Kvalitatiivisen aineiston analyysi ja tulkinta* (Analysis and interpretation of qualitative materials), edited by Klaus Mäkelä, pp. 42-61. Helsinki: Gaudeamus.

Newman, Isadore and Carolyn R. Benz. 1998. *Qualitative-quantitative research methodology. Exploring the interactive continuum.* Carbondale: Southern Illinois university press.

Raunio, Kyösti. 1999. *Positivismi ja ihmistiede. Sosiaalitutkimuksen perustat ja käytännöt* (Positivism and human science. Foundations and practises of social research). Helsinki: Gaudeamus.

Richards, Lyn. 1999. *Using NVivo in qualitative research.* Melbourne: QSR.

Roos, J.P. 1989. Kovaa kvalitatiivista (Hard-edge qualitative). *Sosiologia* 26(2):141-143.

Seidel, John J. 1991. Methods and madness in the application of computer technology to qualitative data analysis. In: *Using computers in qualitative research*, edited by Nigel G. Fielding and Raymond M. Lee, pp. 107-116. London: Sage.

Strauss, Anselm L. 1987. *Qualitative analysis of social scientists.* Cambridge: Cambridge University Press.

Strauss, Anselm L. 1997. Foreword. In: Muhr, Thomas: *ATLAS/ti. Visual qualitative data analysis management model building. Short user's manual.* Scientific software development: Berlin, 1.

Strauss, Anselm L. and Juliet Corbin. 1990. *Basics of qualitative research: grounded theory procedures and techniques.* London: Sage.

Tashakkori, Abbas and Charles Teddlie. 1998. *Mixed methodology. Combining qualitative and quantitative approaches.* Applied social research methods series, volume 46. London: Sage.

Tesh, Renata. 1990. *Qualitative research: analysis types and software tools.* London: Falmer press.

Toivonen, Timo. 1999. *Empiirinen sosiaalitutkimus. Filosofia ja metodologia* (Empiric social research. Philosophy and methodology). Porvoo: WSOY.

Töttö, Pertti. 1999. Kvalitatiivisen ja kvantitatiivisen tuolle puolen? Metodipoliittinen puheenvuoro (Beyond qualitative and quantitative? On methods and approaches). *Sosiologia* 36(4):280-292.

Weizman, Eben A. and Matthew B. Miles. 1995. *A sourcebook for computer programs for qualitative data analysis.* London: Sage.