

SYMPOSIUM EDITORIAL

Computer-assisted analysis in the social sciences: A unique strategy for mixed research?

EUGENIO DE GREGORIO

and

FRANCESCO ARCIDIACONO

*It is the brain, the little gray cells on which one must rely.
One must seek the truth within – not without*

(Hercule Poirot, by Agatha Christie)

The use of computer science in social research, and in particular psychology, has grown enormously in the last 20 years. Analysis techniques and very complex software, which make it possible to put into practice different approaches based on both experimental quantitative research and as a support for quality research, have been available for quite some time. However, different authoritative researchers have recently advocated the need to think about (and consequently put into practice current operational research methods) integrating the different approaches.

The use of computers in the social sciences

On looking at qualitative research methodology various teams of researchers have defined the use of computer programs as a substantially new trend. (Fielding & Lee 1991, 1998; Coffey, Holbrook & Atkinson 1996; Bringer, Johnston & Brackenridge 2006; Lewins & Silver 2007; Baze-

ley 2007). The acronym CAQDAS (Computer Assisted Qualitative Data Analysis Software) represents all the programs which allow us to analyse qualitative data. CAQDAS helps the work of the researcher in the following ways:

- by making some of the phases entirely automatic;
- by making it possible to compare the interpretations of different analysts working on the same material;
- by memorising all the different phases in the analysis by making notes (files in which the researcher can write down definitions, ideas, explanations, ethnographic notes, routine procedure to share with other members of the team);
- by speeding up the creation of graphs, tables, relationship networks for the summing up of the results.

Most of the CAQDAS available today allows us to deal with, integrate and analyse informa-

tion in an extremely versatile way. The differences between the main kinds of software concern on one hand the theoretic models of the respective founders and on the other the potentiality of each of them as regards the analytical needs of the researchers (Barry 1998). Since the beginning of the nineties the programs have been conceived with reference to particular areas of application and interest (Weitzman & Miles 1995):

- Text retrieval and textbase managers
- Code-based theory building software
- Conceptual network builders

In the midst of this complexity, we can say that the distinctions between programs are less clear. Most available programs offer multiple, integrated perspectives and can be used in a non-exclusive way between the various levels.

As Lewins and Silver (2007) have suggested:

... distinctions have become blurred. This has happened, for example, between software formerly labelled code and retrieve and the more extensive functionality in code-based theory builders. Now very few software programs remain solely in the code and retrieve category. Similarly, some of the code-based theory building software has taken on tools more traditionally featured in text retrievers ... (2007: 7).

Transversally, results processing is based initially on a quantitative approach (frequency calculations, contingency tables, functions with regard to the ranking and ordering of frequency). In the course of time, by continually trying to perfect software programs that allow better and better data analysis and produce a better 'fit' between results, data and theoretic models, the world of software programmers and that of social scientists have drawn more and more apart.

Similar attention to the cognitive and informative complexity of theory has not been brought to experimental quantitative research. Today we can say that the most sophisticated techniques of quantitative data analysis succeed

with difficulty in grasping the complexity of human action in everyday life; on other hand, the push towards a increased emphasis on the interpretative aspects of the research process means we risk not being able to grasp the 'general laws' that govern human social processes. Although integrating methodology and tools has been in the minds of various scholars for some years, up to now this has only resulted in works of a taxonomic nature, as well as good intentions and hypotheses as to how carry it out; all this, however, has rarely brought about a good piece of integrated research (Johnson & Onwuegbuzie 2004).

Until now the growth of software products which act as a support to research in the social sciences has proceeded in a parallel direction with the increasingly complex 'ways of thinking' of those who carry out social and psychological research (Bazeley 2006). The most recent programs for the analysis of quantitative data include functions to export the results (codes, topics, conceptual networks) into programs which produce results in a quantitative and statistical format (this is the case of ATLAS.ti e di MAXQ-DA); work functions and inter-subject comparisons between the different researchers who are dealing with the same data are, however, still not used a great deal, thus there is a risk of obtaining an *emerging theory* that may represent *the emerging theory for researcher X* but contains little material that can be generalised, as it is too tied to *his/her* own specific content and circumscribed aims, *his/her* own interpretation of reality, *his/her* theoretic preferences and epistemological options. In a similar way, some systems used for statistical analysis contain a logic that is 'qualitative' to the extent that they make it possible to work with non-structured data (texts of different lengths), data on which, however, even very complex and advanced statistical analyses can be carried out (Spad and TalTac, for example). This kind of analysis often needs a great deal of onerous work in order to separate the texts from their natural contexts: here every single word has a meaning

that depends on the other words produced in the same context, on the time and place of production and reproduction, on the communicative intent of the person who produced them, on the meaning attributed by the source and by the semantic knowledge of the speaker/reader¹.

In our opinion, those who create such software programs are inclined to reproduce the complexity of the thought processes which are characteristic both of the researchers who deal with the problem or specific set of problems, and of those who participate in the research. In this way the complexity of software programs (complexity understood as complexity of use and complexity in terms of coefficients, indices or networks of theoretical connections) reflects the ‘cognitive complexity’ and ‘psychosocial complexity’ of human activity, including the process of acquiring knowledge in human and social terms.

Which tool?

Here, as in the whole of this work, ‘tools’ are often mentioned. It is necessary to spend a few words on this term and on its implications for a *mixed/blended approach*. We will take the obvious as a starting point: the human mind (every human mind, including the researcher’s) is rooted in an environment. According to cultural psychology this environment is populated by others, but also by *tools*, the artifacts (Bruner 1991). These are both products of the human mind and accessories which are useful for a knowledge of the environment. As far as they are products of the human mind (cognitive artifacts), tools incarnate the ways in which human beings during the centuries have understood reality and have established a relationship with their environment (for example developments in the sector of communication technology, from the land lines of the sev-

enties and eighties to smaller and smaller mobile phones and the most recent form of wireless video communication. The software that has come to aid in this research does not escape this definition and Bateson’s metaphor about a blind man’s stick is still valid:

Suppose I am a blind man, and I use a stick. I go tap, tap, tap. Where do I start? Is my mental system bounded at the hand of the stick? Is it bounded by my skin? Does it start halfway up the stick? Does it start at the top of the stick?

(Bateson 1972: 459 in Mantovani 1998: 121).

Since Vygotsky’s contribution (1978), theorists of cultural psychology have seen the tools of the physical world as extensions of the mind and as elements of a context that are part of the mind (Cole 1995; Cole 1996). The difficulty in tracing a boundary between the stick and the mind of the blind man represents the connection between social action, the person’s aims and context and the knowledge that results from this interaction. “External” realities become a part of us when they act as tools to carry out our plans [...] Artifacts are plans which have been given substance’ Mantovani (1998:122).

So why cannot software tools and the programs which allow them to operate be considered in this light? The choice of an approach reflects a cognitive style, a preference, the direction taken by the researcher as regards social reality; the consequent choices of the methodological system, including research strategies and technical analysis of the data are – as a result – the extensions and implications of those preferences and meanings. Thinking in terms of *blended can*, however, result in research of a similar cognitive complexi-

1 It is necessary to mention the choices offered by some important firms that produce software for quantitative analysis: they sell in fact modules that can be added onto the usual programs that allow, on different levels, qualitative and integrated data analysis and also text mining and data mining software. This is an excellent alternative to to a makeshift bridge consisting of other compatible software used to cross over from one from one program to another (Microsoft Office ExcelTM and WordTM are the most popular). However this is a relatively new market with a cost that not everybody can afford.

ty into the software and technical tools that are part of scientific culture.

Concretely, technologies have an impact on ontological, epistemological and methodological standpoints; our attention should be directed towards the way technologies and users *jointly* build our knowledge of phenomena, not only reactively. Qualitative and quantitative researchers have to take into accounts the development of technologies; they have to choose what they need, which issues and troubles, which potentialities and advantages, which help or limitations.

According to Brown (2002: 2):

The objective here is to consider how do we arrive at a solution that all data are captured and processed in a range of high quality, digitized formats, without altering the fundamental characteristics of good qualitative research (... open-ended process, flexibility, and strong orientation towards inductive reasoning that are ideally suited to exploring the construction of human meanings in the context of their own making).

Technology has a great impact on the whole process of research. Software offers great assistance, but the majority and hardest part of the work must still be undertaken by the researcher.

Researchers have a great responsibility: they can drive the process of research, but can also suggest answers. Likewise, researchers can build methodology but can also (en) literally (en) build data. A strategy to accommodate these risks is reflexivity².

Coding is a procedure through which texts, videos, pictures (and other materials) are synthesised and brought back to fundamental conceptu-

al core categories. There are two main strategies to assign codes to documents to analyse.

- a) Theory-driven: The choice of the portion of text depends on the theoretical dimensions that the researcher wants to increase the value of documents.
- b) Grounded theory: Themes emerge from the data as read by the researcher.

In both cases, the first most important step is to assign codes to contents. According to grounded theory methodology, a theory 'emerges' from data; according to your point of view researchers identify themes. A theory-driven approach allows the researcher to complete or better define his/her theoretical issues. But, in both cases fragmentary procedure will make you blind in front of the whole document. 'An account of themes "emerging" or being "discovered" is a passive account of the process of analysis, and it denies the *active* role the researcher always plays in identifying pattern/themes, selecting which are of interest, and reporting them to the readers' (Braun & Clarke 2006: 80). Most research projects need to be close to the data (Lewins & Silver 2007), continually interact with the data and contain options for evaluation. Software programs and the computer permit this kind of contact, a sort of 'dough rolling', in the sense of rolling out the acquired information which is not without its risks as regards the quality of the overall research³.

This is the most plausible of the criteria used to obtain knowledge of the complexity of reality as seen by the social actors and experienced in their activities (Brewer and Hunter 2006) This kind of interaction has also the advantage of allowing an intuitive and flexible approach to the data, obviously within limits. This is emphasised by Bazeley (2007: 3):

2 'The importance of reflexivity in psychological studies has gained increased recognition as researchers have begun to acknowledge themselves as co-creator of the knowledge [...] As researcher attempts to balance the potential benefits of researcher involvement with a commitment to accurately represent their respondents' voice, you can enlist the strategy of reflexivity as mediator of the research process'.

3 We are thinking of the criteria that defines the quality of good research and of the fact that it exists both for the quantitative and for the qualitative approach, but not yet for research using the *mixed/blended method*.

There are procedures that can be used, too, to check for completeness, and use of a computer makes it possible to test for negative cases (where concepts are not related). Perhaps using a computer simply ensures that the user is working more methodically, more thoroughly, more attentively. [...] Even so, human factors are very much involved, and computer software cannot make good work that is sloppy, nor compensate for limited interpretative capacity. As much as 'a poor workman cannot blame his tools', good tools cannot make up for poor workmanship.

*"Eh bien, since you are too stupid to guess, I will tell you.
I am human, am I not? I can be the machine if it is necessary.
I can lie back and think. I can solve the problems so.
But I am human, I tell you. And the problems concern human beings"*

(Hercule Poirot, by Agatha Christie)

Eugenio De Gregorio

(University of Rome 'La Sapienza', Italy)

Francesco Arcidiacono

(University of Neuchâtel, Switzerland)

References

- Barry CA (1998) Choosing qualitative data analysis software: Atlas/ti and Nudist compared, *Sociological Research On-line* 3(3); available on www.socresonline.org.uk/socresonline/3/3/4.html.
- Bateson G (1972) *Steps to an Ecology of Mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*, University of Chicago Press: London.
- Bazeley P (2006) The contribution of computer software to integrating qualitative and quantitative data and analyses, *Research in the Schools*, 13: 64–74.
- Bazeley P (2007) *Qualitative Analysis with NVivo*, Sage: London.
- Bringer JD, Johnson LH and Brackenridge CH (2006) Using computer-assisted qualitative data analysis software to develop a grounded theory project, *Field Methods*, 18: 245–266.
- Braun V and Clarke V (2006) Using thematic analysis in psychology, *Qualitative Research in Psychology*, 3: 77–101.
- Brewer J and Hunter A (2006) *Foundations of Multimethod Research. Synthesizing Styles*, Sage: London.
- Brown D (2002) Going digital and staying qualitative: Some alternative strategies for digitizing the qualitative research process [Electronic Version]. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 3. Retrieved April 7th, 2007, from <http://www.qualitative-research.net/fqs-texte/2-02/2-02brown-e.pdf>.
- Bruner J (1990) *Acts of Meaning*, Harvard University Press: Cambridge MA.
- Coffey A, Hollbrook B and Atkinson P (1996) Qualitative data analysis: Technologies and representations. *Sociological Research On-line* 1(1); available on www.socresonline.org.uk/socresonline/1/1/4.html.
- Cole M (1995) Culture and cognitive development: From cross-cultural research to creating systems of cultural mediation. *Culture & Psychology*, 1: 25–54.
- Cole M (1996) *Cultural Psychology*, Harvard University Press: Cambridge MA.
- Fielding NG and Lee RM (1991) (eds) *Using Computers in Qualitative Research*, Sage: London.
- Fielding NG and Lee RM (1998) *Computer Analysis and Qualitative Research*, Sage: London.
- Johnson R and Onwuegbuzie AJ (2004) Mixed methods research: A research paradigm whose time has come, *Educational Researcher*, 33: 14–26.
- Lewins A and Silver C (2007) *Using Software in Qualitative Research. A step-by-step guide*, Sage: London.
- Mantovani G (1998) *L'elefante invisibile. Tra negazione e affermazione delle diversità: scontri e incontri multiculturali*, Firenze: Giunti.
- Russell GM and Kelly NH (2002) *Research as Interacting Dialogic Processes: Implication for reflexivity* [Electronic Version]. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 3. Retrieved April, 6th, 2007, from <http://www.qualitative-research.net/fqs-texte/3-02/3-02russellkelly-e.htm>.
- Vygotsky LS (1978) *Mind in Society*, Harvard University Press: Cambridge MA.
- Weitzman E and Miles M (1995) *A Software Source Book: Computer programs for qualitative data analysis*, Sage: Thousand Oaks CA.