COMPUTER ASSISTED QUALITATIVE DATA ANALYSIS SOFTWARE: A PRACTICAL PERSPECTIVE FOR APPLIED RESEARCH

JOSEPH B. BAUGH (Primary Contact)

UNIVERSITY of ARIZONA SOUTH

836 Broken Arrow Lane

Benson AZ 85602

520-331-6351

jbaugh@email.arizona.edu

ANNE SABER HALLCOM

CAPELLA UNIVERSITY

10515 SW 139 Court

Miami, FL 33186

305-408-7371

anne.saber@beckman.com

MARILYN E. HARRIS

CAPELLA UNIVERSITY

10173 Township Park Road

Petoskey, MI 49770

231-348-5826

harri1me@cmich.edu

1. INTRODUCTION

Computer assisted qualitative data analysis software (CAQDAS) holds a chequered reputation to date in academia, but can be useful to develop performance metrics in the field of corporate social and environmental responsibility and other areas of contemporary business. Proponents of using CAQDAS cite its ability to save time and effort in data management by extending the ability of the researcher to organize, track and manage data. Opponents decry the lack of rigor and robustness in the resultant analyses. Research reveals that these opinions tend to be divided by "the personal biography and the philosophical stance of the analyst" (Catterall & Maclaran, 1998, p. 207), as well as "age, computer literacy, and experience as a qualitative researcher" (Mangabeira, Lee & Fielding, 2004, p. 170). A more recent article (Atherton & Elsmore 2007) discussed the continuing debate on CAQDAS in qualitative research:

> The two perspectives both indicate that CAQDAS should be used with care and consideration; in ways that explicitly demonstrate a "fit" between the ethos and philosophical perspective(s) underpinning a research study, on the one hand, and the means of ordering and manipulating the data within CAQDAS on the other. (p. 75)

Despite the ongoing literary debate on the merits of CAQDAS, the use of computer-aided qualitative data analysis has become acceptable to most qualitative researchers (Lee & Esterhuizen; Morison & Moir, 1998; Robson, 2002). However, writers advise that researchers avoid the trap of letting the software control the data analysis (Catterall & Maclaran, 1998). Morison and Moir counseled that CAQDAS is merely one tool in the qualitative data analysis toolbox. No tool should replace the researcher's capacity to think through the data and develop his or her emergent conclusions (Atherton & Elsmore, 2007). On the other hand, Morison and Moir among others (e.g., Blank, 2004; Catterall & Maclaran, 1998; Mangabeira et al., 2004) found the use of qualitative data analysis software can also free up significant amounts of time formerly used in data management and encoding allowing the researcher to spend more time in deeper and richer data evaluation.

Qualitative research studies to develop performance metrics can create huge amounts of raw data (Miles & Huberman, 1994; Robson, 2002). Organizing, tracking, encoding, and managing the data are not trivial tasks and the effort should not be underestimated by the applied researcher. Two methodologies exist to handle these activities and manage the data during the data analysis phase. The first methodology is a manual process, which must be done at times to avoid missing critical evidence and provide trustworthiness in the process (Malterud, 2001), while the second methodology indicates the use of technology for managing the data and avoid being overwhelmed by the sheer amount of raw data (Lee & Esterhuizen, 2000). It is the experience of the authors that some manual processing must be interspersed with CAQDAS. This provides an intimacy with the data which leads to the drawing of credible and defensible conclusions. Thus, a mixed approach that melds manual and automated data analyses seems most appropriate. A basic approach for applying traditional qualitative research methodologies lies in the ability of CAQDAS to support data reduction through the use of a "provisional start list" (Miles & Huberman, 1994, p. 58) of data codes that are often developed manually from the research question.

A rise in the use of CAQDAS for applied research and other nonacademic research fields has been identified (Fielding & Lee, 2002). Since CAQDAS is becoming more prevalent in nonacademic researcher populations and can be useful for developing performance metrics for corporate social and environmental responsibility and solving other complex business issues, it seems prudent at this juncture to discuss how to use the software appropriately rather than rehash the argument for or against using CAQDAS. Selection of and training with an appropriate CAQDAS package can help the

researcher manage the mountains of data derived from qualitative research data collection methods (Lee & Esterhuizen, 2000).

2. FACILITATING QUALITATIVE RESEARCH WITH CAQDAS

As data for performance metrics are collected from study sources, they should be stored in an appropriate file format for importation into the CAQDAS package to facilitate the encoding process. A "provisional start list" (Miles & Huberman, 1994, p. 58) of data codes keyed to the interview guide or survey questions should be established as part of the study protocol. The hierarchical format of this initial list of data codes often develops from adherence to the interview guide or survey and should facilitate insertion of additional codes, as they become necessary.

Lincoln and Guba (1985) recognized that new classifications and categories typically will emerge during the data analysis process, so additional codes may be created as needed during the data analysis phase and added to the data code list. Since data encoding is a "dynamic and fluid process" (Strauss & Corbin, 1998, p. 101), the ease with which data codes can be added, modified, merged and/or eliminated in the selected CAQDAS package with minimal disruption to previously coded and linked data is a beneficial feature of the software and facilitates effective data management. This feature, in turn, supports a more systematic data analysis process that may result in more conclusive findings (Lee & Esterhuizen, 2000). The chosen CAQDAS package should provide a suitable data indexing interface that supports easier data tracking and retrieval. This feature is important for understanding the linkages between the data (Miles & Huberman, 1994).

Effective qualitative data analysis has two primary building blocks that support emergent themes and patterns that are necessary for developing performance metrics: (a) encoding the data and (b) developing memos that are related to the encoded data (Lofland, 1995). The use of memos permits the researcher to capture thoughts and perceptions in field interviews notes about a participant's behavioral characteristics and non-verbal communications, which should address some concerns of basing data analysis strictly on textual data (Catterall & Maclaran, 1998). The researcher can then link these thoughts and perceptions to the appropriate transcribed statements (Lofland & Lofland, 1995), which, in turn, enriches the data, places data back into context, and supports thick descriptions (Creswell, 1998; Denzin, 2001; Geertz, 1973). As concepts, patterns, and themes begin to emerge from the data analysis and interpretation, memos allow the researcher to capture these thoughts, as well, and link them to appropriate encoding levels.

Training on the software is recommended to harness its potential for tracking, organizing and managing data developed from a wide variety of sources. Appropriate usage of CAQDAS can then free the researcher from the burden of managing the raw data and allow him or her time to delve into

the data and observe emergent themes and patterns as they develop. The prototype for a training workshop being developed in conjunction with this paper and other projects has been tested with doctoral students performing qualitative case studies. However, the beneficial attributes of CAQDAS can be leveraged across a variety of researcher populations.

To achieve the desired results of any research study, it is necessary to ground the study within classical qualitative research methodologies and then apply an appropriate CAQDAS package to the data obtained from the field study, stakeholder interviews, Internet surveys, and/or other data sources. Selection of a suitable CAQDAS package requires the researcher to understand the purpose of the study, expected results, and the available timeline in which to perform the study. Absent adherence to accepted qualitative research methodologies, findings by nonacademic researchers may be called into question as to the credibility of the study or research project (Fielding & Lee, 2002). As dissimilar user populations have different needs for their research, it is entirely conceivable that one size literally does not fit all needs in CAQDAS environs.

3. QUALITATIVE RESEARCHER POPULATIONS

Fielding and Lee (2002) described a "juxtaposition of academic and nonacademic environments of CAQDAS use and in the audiences of the resulting analyses" (p. 197), which has raised awareness of diverging paths for the social organization of knowledge production. These divergent paths arise from the different populations that are drawn to use CAQDAS to solve specific problems within their own organizations. Mangabeira et al. (2004) described three trends that have been observed in the evolution of CAQDAS:

1. Greater sophistication in new CAQDAS packages

2. More extensive commercialization of CAQDAS markets.

3. Increasing diversity in the types of users (p. 167).

Commercialization of CAQDAS markets is a concern to academia as CAQDAS was originally derived by qualitative researchers who were seeking to solve specific problems. It is significant in a discussion of researcher populations due to the nature of initial CAQDAS development and the future of qualitative research studies. Mangabeira et al. (2004) expressed concerns about commercialization and the threat to the open flow of information as it pertains to the academic community:

> The commercial success of particular packages might translate into the closure and stabilization of particular approaches to or styles of analysis. Some have worried, too, that a field that was largely based on friendly collaboration and the relatively free flow of

information was becoming competitive with information now being treated increasingly as a commercially sensitive commodity. (p. 167)

Nonacademic groups, such as applied researchers or market researchers, may not share those same concerns as their studies are typically driven and funded by for-profit organizations. Additionally, the research requirements and the rigor with which those requirements are pursued can be different between groups. Fielding and Lee (2002) described three distinct qualitative researcher groups as (a) academic researchers, who conduct rigorous research and publish the results within accepted academic journals; (b) applied researchers with social science backgrounds who are working on specific projects for organizations outside of academic disciplines; and (c) researchers who may have little or no connection to academia, but are researching on behalf of professional practices or for specific organizations. Fielding and Lee described this final group as:

> "Nonacademic adopters" while pointing out that our use of the term nonacademic is not intended pejoratively but simply to reflect lack of social science background. Indeed, we argue that such users are an important potential source of innovation as well as posing challenges to established (academic) community-based understandings of the craft of qualitative research. (p. 204)

Mangabeira et al. (2004, p. 170) described qualitative researcher groups in the CAQDAS pantheon from a different perspective. This typology crosses the boundaries of Fielding and Lee's qualitative research groups and includes (a) "program loyalist" researchers, who are defined as being younger and more computer literate, which allows an easier acceptance and smoother learning curve with CAQDAS packages than other groups, but who may not possess a critical perspective on the qualitative processes and methodologies; (b) "critical appropriators", who work within specific research frameworks and understand the intricacies of robust qualitative research; and (c) "experienced hands", who are well versed in the manual processes of qualitative research, but who attained their computer skills at an older age and may be more reluctant to try CAQDAS.

Figure 1: Mapping researcher groups. © 2009 Baugh, Hallcom & Harris
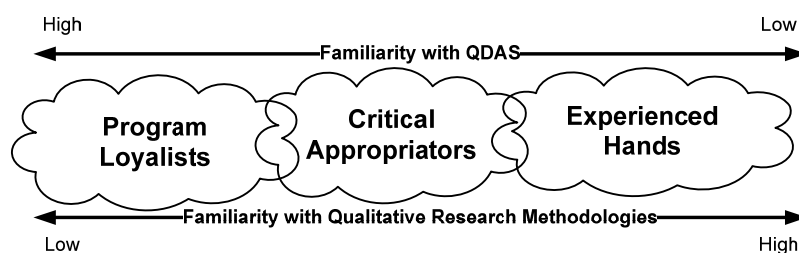
Along a spectrum (Figure 1) bounded at one end by full confidence in the results engendered by CAQDAS and the other end by little or no confidence in the results, the loyalist researcher has a greater familiarity with CAQDAS and exhibit more certainty in the results. The critical appropriators, while they embrace CAQDAS to a certain extent, often exhibit less certainty in the results unless they are obtained with strict adherence to accepted qualitative research standards. The final group, the experienced hands, exhibit skepticism toward results derived through CAQDAS packages and have commensurately less familiarity with the software.

Each of these qualitative research groups, regardless of the chosen typology, has diverse training needs. In addition, the selection criteria of an appropriate CAQDAS package may also be different between these groups. Thus, any training program that addresses the practical perspective of CAQDAS and seeks to ensure adeptness with the software must take these divergent perspectives into consideration.

4. CAQDAS SELECTION

Fielding and Lee (2002) identified three generations of CAQDAS, including first generation software which consisted of word processing and database software that allowed a researcher to search for specific text strings, but did not allow for encoding the data. The second generation software was more robust and provided additional capability to encode the data, retrieve data by those codes and add memos to the encoded text. Second generation CAQDAS is commonly referred to as code and retrieve software. Third generation CAQDAS added the ability to create complex families or networks of codes to support theory building. In most cases, third generation CAQDAS packages are represented by the latest versions of the industry leading products such as Atlas.ti® and NUD.IST® and are classified as code-based theory builders. Other prominent CAQDAS products, classified as conceptual network builders, provide linking and graphic capabilities, but may lack coding and database management features (Catterall & Maclaran, 1998, p. 212).

Once CAQDAS becomes applicable to a particular study or research problem, various packages must be evaluated and an appropriate CAQDAS package selected. Selection criteria may be different depending on the group in which a specific qualitative researcher resides. Fielding and Lee (2002) expressed concerns over the method used by nonacademic researchers to select an appropriate CAQDAS package:

> Such users seem even less likely than those introduced to CAQDAS via academic courses to make a discriminating choice of software. Choice may be a chance matter of a contact in their social network, a call to an acquaintance in a university, or a Web search that yields package or course advertising. (p. 207)

Graduate students and other qualitative researchers may be directed to use a particular CAQDAS package as part of a university's standard software application suite. Some researchers may require more than one CAQDAS package as one described, "Ideally, we should be using multiple programs" (Mangabeira et al., 2004, p. 170) to accomplish different segments of the qualitative data analysis process. Other researchers may be strongly influenced by a specific CAQDAS package used in the training milieu. This should sound an alert to training groups to ensure that the software used for training qualitative researchers is, in fact, effective for the purposes of the training audience.

5. CAQDAS TRAINING

Initial training in CAQDAS should begin with ensuring that trainees fully understand the importance of following a sound qualitative research methodology. Trainees should be apprised early on that CAQDAS complements that methodology. CAQDAS does not replace it.

Blank (2004) developed a course at American University that teaches graduate students the basics of CAQDAS applications. This course covers the basics of qualitative data analysis and concentrates primarily on teaching through practical assignments that involve creating data categories, encoding assigned articles, and developing reports from the data analyses. Blank related that it is possible to teach qualitative data analysis skills in conjunction with training on software and that "students can learn much easier and faster in a classroom" (p. 194).

Fielding and Lee (2002), on the other hand, focus their training efforts "on the user rather than on the software" (p. 197). While these researchers have developed an organized training program to meet the needs of the diverse researcher populations discussed previously in the United Kingdom, they have also found that training across the spectrum of qualitative researchers is still an ad-hoc process for the most part and that the training that is available is distributed sporadically. Fielding and Lee also found that it is easier to ensure a strong qualitative research background in academic audiences, while training offered to nonacademic audiences generally focuses solely on a specific software package and not on teaching qualitative research methodologies.

Blank (2004) also recommended making the training as practical as possible. In an actual training scenario, it is important to understand the audience and tailor the training to meet their specific needs. While this may seem to be a daunting task in groups with mixed experience levels in qualitative research and various levels of computer literacy, it is worth the effort to determine the needs of a particular audience and design training appropriately.

6. USING CAQDAS IN APPLIED RESEARCH

CAQDAS can be very effective in applied research, such as market research, developing performance metrics for corporate social and environmental responsibility, and many other business applications. When designing a qualitative study, it is beneficial to establish an initial list of codes (Miles & Huberman, 1994) from the research design. A naming convention can be developed that indicates code level by capitalization. Level One codes may be typed in all lower-case, while Level Two codes can begin with a capital letter. In more complex encoding schemes, ensuing levels which indicate emerging themes and patterns can be defined in all capital letters or some other alphanumeric schema.

As Blank (2004) discovered, "Students found creating categories and coding to be hard, complicated work. During their coding experience, most discovered that certain categories, however useful they may have seemed in the abstract, did not work well in practice" (p. 192). This is not an uncommon phenomenon. Most CAQDAS packages have an effective code management facility that makes merging and splitting codes a relatively painless process. These codes track and manage quotations from interviews and other sources of qualitative data deemed pertinent to the study design. A similar process can be used in any applied research or other study application, such as that required for developing performance metrics for corporate social and environmental responsibility.

The encoding process begins with assigning initial codes continued with multiple passes through the data assigning progressively higher level codes that emerge during the preceding encoding pass. Memos are attached to encoded data segments during successive encoding passes to accurately track and identify the source of the quotes. After the data are encoded, reports can be run to isolate pertinent data segments by code category for later reference. Most later-generation CAQDAS packages contain graphic and network building capability that support the development of higher level themes and patterns as the performance metrics are finalized. In a graduate student setting where theory building skills may be critical, CAQDAS training should be extended to cover these features.

7. BENEFITS OF CAQDAS

The benefits of using CAQDAS packages for qualitative data analysis are widely touted. The basic advantages of data organization, tracking, and data management have been discussed previously. It is worth repeating that freeing the qualitative researcher from the drudgery of manual data management can free significant time that may be spent in contemplating the data, thinking reflexively on emergent patterns and themes and developing credible conclusions that "make sufficient reference to the context in which the data originated" (Atherton & Elsmore, 2007, p. 69). Codes and groupings assigned through the software act as building blocks for thought and building theory around emerging patterns and themes (Miles & Huberman, 1994; Robson, 2002; Rubin & Rubin, 2005).

Carefully considered use of qualitative data analysis software with its capacity to easily track and retrieve data may allow the qualitative researcher to develop the proper "mindset to bring to the task of making analytical sense of the data" (Lofland & Lofland, 1995, p. 122) and support credible conclusions drawn from the data. Fielding and Lee (2002, pp. 212-214) also described several benefits for both researchers and audiences including new methods of analysis, superior data management capacities, facilitated data analyses, and enhanced legitimacy of qualitative research. Each of these features may be beneficial to the development of sound performance metrics.

Hwang (2008) cited the transparency and reliability of the data analysis processes, which may enhance the credibility of the research processes. Hwang warned that data analysis is not accomplished automatically, but time savings may be garnered through the use of CAQDAS packages as the computer can act as a "research assistant, which can save time and make your work better, especially for large data sets" (p. 524). Hwang also described an additional benefit not mentioned in most of the literature, which may be particularly beneficial in dealing with complex social and environmental issues. This benefit is the effectiveness for using CAQDAS as a project management tool. One of the authors regularly uses CAQDAS products to capture, analyze, compile, and manage project stakeholder requirements and expectations for complex software development and other projects. This usage, which lies well outside the academic realm, highlights the admonishment by Mangabeira et al (2004) that academics must understand new ways that CAQDAS is utilized "to help new users and audiences avoid the pitfalls that are all too familiar to those with a background in qualitative research" (p. 175).

8. LIMITATIONS OF CAQDAS

Time constraints are significant concerns among applied researchers as well as other nonacademic researchers. These researchers may not have the luxury of time afforded to academic researchers due to the pressures of project deadlines and other time related pressures (Fielding & Lee, 2002; Mangabeira et al., 2004). There can be no doubt that third generation CAQDAS is very complex software and a significant learning curve (Blank, 2004; Hwang, 2008) is applicable for mining its full potential for qualitative research. This learning curve adds a considerable time constraint (Catterall & Maclaran, 1998) in mastering the software in order to free the researcher to perform in-depth analyses and interpretation of the data.

Once the researcher has been exposed to the features of the chosen software, the time-consuming task of plowing through the mountain of qualitative data looms before the researcher. Blank (2004) described the basic issue with qualitative data analysis training as being one of "too much data … where many students get stuck. They wonder what they can do with all those field notes, documents, and transcripts. The task may look overwhelming" (p. 187). Graduate students and other academic

researchers are not immune to time constraints as theses, dissertations, and other research study reports may often be postponed while the researcher struggles to learn a specific CAQDAS package (Mangabeira et al., 2004). This temporal phenomenon is not limited to academia, but may play a larger role in the nonacademic research arena.

One of the primary concerns of established qualitative researchers in academia is that nonacademic researchers may not be able to cope sufficiently with the massive amounts of data generated by Internet surveys and other means of mass data collection. Fielding and Lee (2002) described "pressures not to code, or even load all the data (particularly in market research), as well as pressures against systematic or sufficiently precise coding" (p. 211). The fluid nature of business studies in applied and market research arenas combined with deadline driven pressures will continue to drive temporal and other pressures for nonacademic researchers. Gaining a better understanding in advance of how to perform sound qualitative research processes in conjunction with learning the selected CAQDAS package can help diminish the effect of this temporal limitation for both academic and nonacademic researchers.

9. RECOMMENDATIONS

Qualitative researcher groups need to work together (Figure 2) to cross-pollinate ideas and methodologies across the three primary groups: (a) Academic Researchers, (b) Applied Researchers, and (c) Nonacademic Adopters. Academic researchers can provide necessary training and support to the other two groups of researchers, while simultaneously gaining a better understanding of practical issues that emerge from applied research and nonacademic studies. Applied researchers can also provide support to academic researchers while collaborating with nonacademic adopters to develop pertinent studies that solve specific business problems. Nonacademic adopters will gain a better understanding of sound research practices while working on the cutting edge of qualitative research in consumer markets. Taken as a holistic approach, the diverse groups of qualitative researchers can use CAQDAS in new and innovative ways to stretch the boundaries of qualitative research while seeking to recognize and better comprehend complex issues in contemporary turbulent market conditions.

Figure 2: Research population interactions. © 2009 Baugh, Hallcom, & Harris.
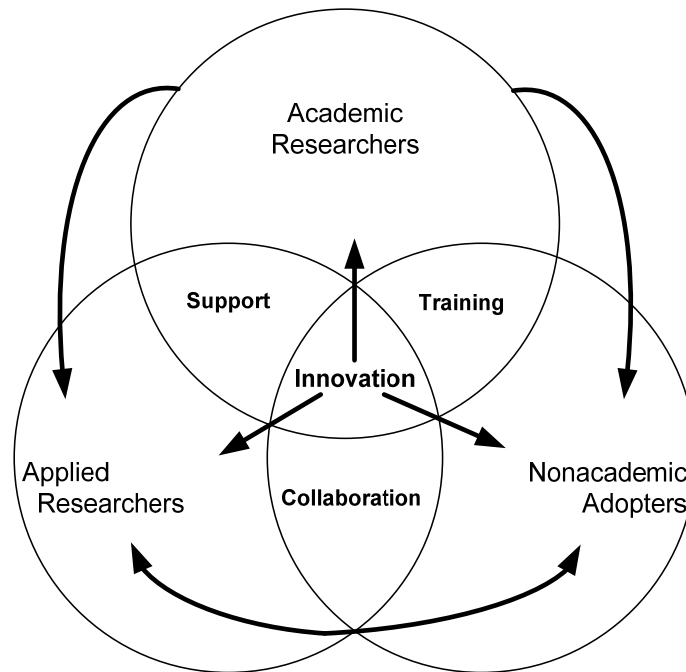
10. CONCLUSION

Qualitative research studies and CAQDAS have become valuable tools in the research arsenal. Qualitative studies are more prevalent in today's academic environment where adherence to robust methodological practices is observed. However, with the burgeoning of other categories of qualitative researchers – some far removed from academia – conclusions drawn from applied research studies may be sketchier. Thus, qualitative researchers require training not only in CAQDAS packages, but in sound qualitative research methodologies and practices as well.

CAQDAS can support a better understanding of complex and subjective variables drawn from open-ended questionnaires, interviews, Internet surveys, and other means of qualitative data collection. In this manner, CAQDAS can play a viable role in developing performance metrics of corporate social and environmental responsibility and solving other business problems. Providing appropriate training to diverse qualitative researcher populations can support better research into social and business issues. The academic qualitative research community must step up to its responsibility to engender awareness of best practices in qualitative research, while remaining open to innovations in qualitative research discovered by nonacademic researchers. Academia must support the training of nonacademic researchers in both qualitative methodologies and appropriate CAQDAS packages. Taking

responsibility in this manner will provide a higher visibility for academia on complex issues with which practitioners struggle daily to develop and implement relevant solutions.

11. REFERENCES

Atherton, A., & Elsmore, P. (2007). Structuring qualitative enquiry in management and organization research: A dialogue on the merits of using software for qualitative data analysis. *Qualitative Research in Organizations and Management, 2*(1), 62-77/

Blank, G. (2004). Teaching qualitative data analysis to graduate students. *Social Science Computer Review, 22*(2), 187-196.

Catterall, M., & Maclaran, P. (1998). Using computer software for the analysis of qualitative market research data. *Journal of the Market Research Society, 40*(3), 207-222.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among the five traditions*. Thousand Oaks, CA: Sage.

Denzin, N. K. (2001). *Interpretive interactionism* (2nd ed.). Thousand Oaks, CA: Sage.

Fielding, N. G., & Lee, R. M. (2002). New patterns in the adoption and use of qualitative software. *Field Methods, 14*(2), 197-216.

Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.

Hwang, S. (2008). Utilizing qualitative data analysis software: A review of Atlas.ti. *Social Science Computer Review, 26*(4). 519-527.

Lee, R. M., & Esterhuizen, L. (2000). Computer software and qualitative analysis: Trends, issues, and responses. *International Journal of Social Research Methodology, 3*, 231-243.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lofland, J. (1995). Analytic ethnography: Features, failings, and futures. In A.M. Huberman, & M.B. Miles (Eds.), *The qualitative researcher's companion* (pp. 137-170). Thousand Oaks, CA: Sage.

Lofland, J., & Lofland, L. H. (1995). *Analyzing social settings: A guide to qualitative observation and analysis* (3rd ed.). Belmont, CA: Wadsworth.

Malterud, K. (2001, August 11). Qualitative research: Standards, challenges, and guidelines. *The Lancet, 358*, 483-488.

Mangabeira, W. C., Lee, R. M., & Fielding, N. G. (2004). Computers and qualitative research: Adoption, use, and representation. *Social Science Computer Review, 22*(2), 167-178.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Morison, M., & Moir, J. (1998). The role of computer software in the analysis of qualitative data: Efficient clerk, research assistant, or Trojan horse? *Journal of Advanced Nursing, 28*, 106-116.

Robson, C. (2002). *Real world research: A resource for social scientists and practitioner-researchers* (2nd ed.). Malden MA: Blackwell Publishing.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data* (2nd ed.). Thousand Oaks, CA: Sage.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.