Qualitative Data Analysis:

Comparing Results From Constant Comparative and Computer Software Methods

Jim Vander Putten

Dept. of Educational Leadership

Amanda L. Nolen

Dept. of Teacher Education

University of Arkansas-Little Rock

Paper presented at the Conference of the Association for the Study of Higher Education, November 5-8, 2008 in Jacksonville, FL. Please do not distribute without permission of the authors. Comments and suggestions are welcome.

Address information:
Department of Educational Leadership,
Higher Education Program
2801 S. University Ave.
University of Arkansas-Little Rock
Little Rock, AR 72204-1099

(501) 569-3549 (office)
(501) 569-3547 (fax)
jvputten@ualr.edu

Qualitative Data Analysis:

Comparing Results From Constant Comparative and Computer Software Methods

Introduction

Tracing the origins of qualitative research, several seventeenth century authors laid the foundation for current philosophies and practices (Ritchie & Lewis, 2003). Descartes' 1637 book *Discourse on Methodology* included the important idea that researchers should distance themselves from sources of bias that may impair their analytical abilities. During the same time period, Isaac Newton and Frances Bacon discussed the usefulness of direct observation on knowledge generation. In the eighteenth century, David Hume asserted that knowledge emerged from experiences and is acquired through the senses, and Immanuel Kant extended this to include various interpretations of what is sensed, and that knowledge can be derived from thinking about experiences.

More recent ideas about qualitative research emerged in the late 1800s and early 1900s in anthropological studies of indigenous cultures, and sociological studies of the poor in Europe and in large U.S. cities (Creswell, 2008). According to Creswell (2008), "the actual use of qualitative research in education is most apparent during the last 30 years…" and recent historical developments can be categorized in terms of philosophical ideas (e.g., advocating the naturalistic paradigm as an alternative to traditional research), procedural developments (e.g., advancing types of qualitative research and research designs), and participatory and advocacy practices (e.g., exploring issues about racial identity, feminist perspectives, and GLBT sensitivity).

The past 15 years have witnessed the emerging availability of a variety of computer software programs for qualitative data analysis in social sciences research, documented in part by Weitzman and Miles' (1995) review of 24 software programs and resulting typology for handling qualitative data. Lewins and Silver (2006) updated this typology to reflect the increased functionality of currently available computer software programs in three primary categories: *Code-based Theory Building software* (i. e., allowing the researcher to test relationships between issues, concepts, and themes), *Text Retrievers* (e.g., features including content analysis tools, word frequencies, and word indexing), and *Textbase Managers* (e.g., features including keyword proximity plotting, and chart and graph building). These programs provide researchers with the technology to locate text related to codes and themes, make comparisons among code labels, conceptualize different levels of abstraction in data analysis, and create graphic representations of codes and themes efficiently and effectively (Creswell, 2007).

An increasing amount of scholarship on the philosophical and methodological aspects of qualitative research in the recent past has been paralleled by an increasing amount of qualitative research being conducted in education. Fallon (2006) observed that “a broad review of the past several decades reveals a predominance of quantitative research in the early and mid-20th century followed by a sharp increase in qualitative research in the late 20th century and continuing to predominate to the present day.” (p. 145) Furthermore, Shavelson and Towne’s (2002) National Research Council report noted the popularity of qualitative methods in observing “the current trend of schools of education to favor qualitative methods, often at the expense of quantitative methods, has invited criticism” (p. 19).

Eckardt’s (2007) descriptive research on the methodologies of research paper proposals submitted to, and accepted for AERA conference presentation provided some evidence of this shifting balance. Data from the 2006 AERA conference indicated that of 8,612 papers submitted, 37.8% used exclusively qualitative methods, 28.7% used exclusively quantitative methods, and 32.6% used mixed methods or a conceptual/theoretical approach that did not involve analysis or presentation of data. These percentages were remarkably similar (37.9%, 30.4%, and 31.8%, respectively) for the 4,550 papers accepted and presented at the conference.

The increased amount of published research in education that utilizes qualitative methodology coupled with the increase in availability and ease of use of computerized data analysis software raises questions about issues of trustworthiness and accuracy of results compared to those obtained by more traditional qualitative methods. This paper addresses one contextual question for researchers to consider as a qualitative research study is being conceptualized: Should the data analysis plan employ a manual constant comparative approach or a qualitative data analysis computer software approach? To provide insights into these decisions, one overarching research question guided this study: How do qualitative research results compare when obtained by analyzing the same data using computer software and a constant comparative manual approach?

Literature Review

The conceptual framework for this study is grounded in three areas of literature: constant comparative data analysis, validity in qualitative research, and uses of computer software for qualitative data analysis.

Constant comparative data analysis

Constant comparative data analysis is the process by which the researcher moves back and forth between the data and the field to gather information about a particular concept that will then be coded into categories, properties, and hypotheses. Glaser and Strauss (1967) identified four stages of the constant comparative method: (1) comparing incidents applicable to each category, (2) integrating categories and their properties, (3) delimiting the theory, and (4) writing the theory. These stages are not linear; rather, they overlap throughout the data collection and analysis activities.

While the researcher is involved in first two stages, he or she is engaging in an open coding process. When confronted with new data the researcher asks “What is the meaning of this?” or “How does this fit?” In doing so, the researcher creates exhaustive categories with scrutiny to the dimension of each category. Each new occurrence is examined for similarities and differences. These categories are then reduced by collapsing and combining based on meaning and relationships. This induction leads the researcher to begin to develop concepts and theory. In the third and fourth stages, the researcher further “delimits” or refines the theory until saturation is reached (Glaser & Strauss, 1967).

This deductive and inductive process allows the research to achieve greater precision and consistency in the resulting concepts and theory development (Corbin & Strauss, 1990). The precision and consistency results from the researcher continually challenging categories with fresh data, examining any variations in patterns and grouping only like phenomena. For example, a researcher might notice that participants engage in Behavior A under Condition A. If, however, the researcher also notices that when under stress the participants engage in Behavior B under Condition A, a variation of the original pattern of behavior emerges. Finding these patterns and variations of these patterns help the researcher to give order to the data and assist with category and theory development.

Validity

While validity is historically a term valued by the quantitative research literature, with the increase in popularity of qualitative research in the 1970s and 1980s, researchers were compelled to demonstrate the legitimacy and accuracy of their findings in a quantitative world. In this new application, validity takes on implications of accuracy and trustworthiness. Miles and Huberman (1994) remarked that validity in qualitative methods should not be linked to truth as for positivists, but rather to trustworthiness – a matter of persuasion where the researcher went to great lengths to make the research process visible and auditable. While qualitative data collection, analysis, and interpretations are viewed as 'looser' than those of quantitative methods, there is no less of a need to ensure the reader of the accuracy and credibility of the study. Establishing validity in this case means the researcher must bracket the observer bias from what is being observed (Shank, 2006).

Westphal (2000) investigated Lincoln and Guba's (1985) concept of trustworthiness of qualitative data analysis results, and used the QSR NUD.IST software package to analyze natural resources data. She identified a series of strategies to increase trustworthiness of qualitative research results, including searching for rival information, linking findings to data and theory, and conducting coding checks. The findings led Westphal to conclude that software can increase the trustworthiness of results and conclusions obtained through qualitative research.

Another way validity has been construed quantitatively is in terms of accuracy, and in this case, refers to the ability to measure or capture a construct in a precise manner that is free of error. Accuracy is a concern for qualitative researchers as well, but tempered by what Hammersley referred to as the conundrum of precision (Hammersley, 1987). Attempting to become too precise and exact will interfere with the truthfulness of the findings. For instance, measuring the length of a pause in conversation in milliseconds is precise; but describing that pause as "the teacher paused a beat before continuing her instructions:" is more accurate. The first measure can be objectively verified certainly, but in its precision, it loses the meaningfulness or truthfulness of the second account. Consequently, the second description, while less precise, is a more accurate measure of the phenomenon. When using qualitative analysis software that can group terms with great precision based on syntactic and linguistic algorithms, one must be concerned with those groupings being so precise they lose connectedness with the overall meaning of the context.

Qualitative data analysis software

Creswell (2007) identified several advantages of using computer software for qualitative data analysis: providing an organized storage system to easily locate data; facilitating close reading of the data, and generating concept maps to visualize relationships between codes and themes. Conversely, several disadvantages of using computer software have been cited: steep learning curves to use software effectively; insertion of a machine between researcher and data can create uncomfortable distance (Fielding & Lee, 1998), and available software may not have desired data analysis features and capabilities (Creswell, 2007).

Gilbert (2002) investigated researchers' experiences when transitioning from manual data analysis to using computer software, and identified three stages in relation to the data: the tactile-digital divide (shifting from paper to computer screen), the coding trap (becoming mired in extensive coding facilitated by the software), and the metacognitive shift (learning to think about software processes to complete qualitative research). Participants in Gilbert's study viewed qualitative data analysis software as a 'cognitive tool' that is influenced by the researchers' skill in accurately using sophisticated software features, and this underscores the possibility of validity issues in qualitative research conducted by novice software users. Catterall and Maclaran (1997) identified limitations of computerized analysis of focus group data and advocated the combination of computerized and manual qualitative data analysis. However, their research design involved adding manual data coding to computerized coding that generated two separate sets of analyses and interpretations that were reported individually, and no comparison of the two data analysis methods was conducted to check for validity of results. Thompson (2002) identified a lack of clarity in describing the mechanical and conceptual analysis processes used as a major weakness in reporting qualitative research results obtained by using computer software. The author used the HyperQual 2 computer software to analyze five excerpts of one long interview, and included graphic depictions as a model for researchers to make each step of the data analysis process transparent.

Morison and Moir (1998) conducted a two-stage qualitative study using manual data analysis in a grounded theory approach to develop hierarchical coding trees that were subsequently entered into the QSR NUD.IST computer software package for additional data analysis. Combining manual and computerized data analysis methods in sequential order is a novel approach, but the subjective decision-making process of manual coding used to develop parameters for subsequent computerized data analysis significantly limits the software capabilities and data analysis possibilities, and raises questions about the validity of results.

Focusing specifically on research that compared qualitative methods, Leech and Onwuegbuzie (2007) applied seven different data analysis tools to one piece of qualitative data to illustrate the types of analyses. Some analyses were completed by hand (e.g., taxonomic analysis, componential analysis) and some used the search command of word processing software (e.g., Keywords-in-Context, Word Count), however, none involved specialized qualitative data analysis software. Carley (1993) used computer software to complete content analysis (extracting concepts from texts) and map analysis (extracting concepts and analyzing the relationships between them) to obtain research results of the same data. Summarizing her thorough technical discussion of text coding decisions, she observed that "each technique may lead to different interpretations of the texts." (p. 103)

Surprisingly, no research has been conducted to compare research results obtained by manual methods and computer software analysis of one entire qualitative dataset. The increasing prevalence of qualitative research in education combined with increasing sophistication of computer software programs for qualitative data analysis raises important questions about the nature of research results obtained from various methods of data analysis. How do results compare when using different data analysis methods on the same data? What are the implications of obtaining different sets of results? What does this mean for the trustworthiness of results for higher education researchers and practitioners? Wolcott's (1994) question is useful to consider: How do we know we are getting it right?

Methods

Data Source

The data for this study is comprised of 115 personal introductions that new subscribers to the Working-class Academics (WCA) electronic mail listserv posted during an eleven-year period from 1994 through 2005. This listserv was created to establish an online community to discuss issues involving significant social class mobility for graduate students and faculty from poverty- and working-class backgrounds. The message that greets new subscribers states:

> "Welcome to the Working-class-list. This is a discussion list dedicated to the discussion of issues pertaining to being an academic who is from a working-class background. These issues may be personal, research oriented, classroom oriented, or societal oriented. Before this list went 'public', we discussed such things as 'coming-out' re: our class backgrounds when teaching, reactions of non-working-class colleagues to class differences, etc. Please introduce yourself to the rest of the list and let us all know a little about yourself and your work."

This dataset was used because of its size, and the analysis of 115 different pieces of qualitative data provided a larger number of data points from which to investigate issues of comparative results and validity.

Manual constant comparative data analysis methods

The constant comparative (Glaser & Strauss, 1967) data analysis was completed first, and Tesch's (1990) eight step qualitative process was used to manually code the data for aspects of the context of working-class identity. After an initial coding of the data generated topics and themes, a second review was completed for data reduction and to shift to a conceptually oriented approach to coding the data. Finally, a third review was completed to identify broad category codes and specific sub-category codes, and to make coding assignments in the data.

Computer software data analysis methods

After the data were entered into an Excel spreadsheet, the computer software program Tropes version 7.0 (Semantic Knowledge, 2007) was used for analysis. The software's statistical and linguistic algorithms clustered and classified the data and identified trends through concept maps or constellations. Key concepts were extracted from the data using the language processing text analysis technology. Tropes 7.0 analyzed the text as phrases and sentences whose grammatical structure provided a context for the meaning of the responses. Once the concepts were extracted, they formed the foundation for the categories used to reduce the extractions. A combination of statistics-based and linguistics-based methods were employed to determine how frequently terms occur in the data and establish a semantic network which created a probabilistic analysis of the co-occurrence of terms. Finally, a constellation of categories was produced to show the relationship and relative importance of the categories across the 115 responses.

Results

Constant comparative data analysis results

The inductive constant comparative manual data analysis process produced two levels of data codes. First, 46 initial narrow sub-categories were generated that were reduced to 5 broad coding categories and applied to the data for understanding self-described working-class identities: 'Parental Influences,' 'Childhood Hometowns,' 'Family Relations,' 'Individual Lived Experiences,' and 'Spousal Support.'

In the broad category '*Parental Influences,'* one graduate student subscriber to the listserv explained:

> "My father is a mechanic and an auto parts wholesaler. He is in ill health and is frequently 'between jobs.' My mom works in a furniture store, a doctor's office and a beauty salon. I am the first in my family to go to graduate school, and only the second in my extended family to graduate from college. I love my family, but they bring chaos to my life. I feel so guilty for saying this, but whenever I talk to them, all of the feelings of insecurity and instability that marked my youth come back to me. They are always on the edge of flying apart and talking to them leaves me feeling hopeless, like there is nothing I can do to make things different for myself or for them. I know it doesn't make sense, but I started to resent even being in grad school because it was keeping me poor. Not having money is not the issue. It's the feeling of helplessness, of recognizing my parents' lives in the excuses I have to make to creditors and utility companies. I really, really hate it."

This category reflects the myriad influences, positive and negative, that parents and grandparents in working-class occupations have on their children and their mental health and educational aspirations (See Table 1 for sub-category coding assignments that emerged from the manual analysis of this text excerpt using a constant comparative approach).

In the category '*Childhood Hometowns*,' a tenure-track Assistant Professor explained,

> "My dad used GI benefits for a low-interest loan to buy a house in Hamilton Township, just outside Trenton, NJ, because he couldn't afford a house in Trenton. It was one of those little split-level developments that sprang up in the 1950s. We lost the house when a highway came through, and bought a ranch house in another neighborhood. From the outside, it certainly looked like the middle-class American dream. But inside the front room was empty -- until an uncle gave us his couch -- and the rest of the furniture

> was old and shabby. Though he has been there since 1969, the mortgage is still not paid off because he has had to refinance the house several times to get money."

Many of the comments in this category contextualized childhood hometowns in terms of labor history, including the strength of labor unions and the industries that served as primary employers of the working-class. In addition, this exemplary comment and the comment used to illustrate the next broad coding category provide two sides of the narrow code 'Stability' in the '*Childhood Hometowns*' category. (See Table 2 for sub-category coding assignments that emerged from the manual analysis of this text excerpt using a constant comparative approach).

The broad category of '*Family Relations*' expands on the labor history context for many participants, including this graduate student:

> "it took me ages to come to see my family as not the pretentious "proper" types they acted like, but were just ordinary working class people who aspired to a better station in life. On my dad's side of the family, all of them were born into sharecropping parentage, though some of my younger uncles came along after grandpa had started working for International Harvester, a good union job. Still, his southern politics wouldn't allow any real sense of union solidarity because collective action was viewed as a type of communism. When my mom finally remarried (a truck driver for a beer distributor), we moved out to the white suburbs and I met lots of folks whose parents were "working class" - electricians and construction people in union jobs - who were in way better shape than my family was. Owned houses and had swimming pools in the back yard and re-did their living rooms every three years, etc. But the union types were working class."

Several comments in this category focused on issues related to sibling relationships, experiences with public assistance programs, and the influences of union membership on family values. As illustrated in this comment, several respondents discussed comparisons of material wealth with other families as a strategy to 'locate' their family in stratifications of socioeconomic status, and indicated levels of social class consciousness among family members and their relationships. (See Table 3 for sub-category coding assignments that emerged from the manual analysis of this text excerpt using a constant comparative approach).

The '*Individual Lived Experiences*' category provided the largest number of sub-categories (n=17) generated in the constant comparative data analysis process, and this experience recounted by a graduate student focused on the intersection of graduate education and social class origins:

> "I have been fortunate to be among other graduate students from the working class. Thus the elitist attitude of many of our profs is somewhat easier to bear, knowing that we're not alone. Our informal support group helps to save our self-esteem, not to mention our sanity! (It also gives us the opportunity to decompress from the frustrations of dealing with the privileged darlings who are our students.) The elitist attitude annoys me, yes, but I've never really felt isolated."

The teaching role is often a point of class conflict for working-class academics, and particularly for those in elite, selective institutions with students from privileged backgrounds. Because the listserv's focus was on working-class academics, many of the sub-categories in this broad category related to educational experiences. However, a number of sub-categories addressed other issues, such as work experiences, finances, religion, personal health, and social class differences (See Table 4 for sub-category coding assignments that emerged from the manual analysis of this text excerpt using a constant comparative approach).

The final broad category focused on '*Spousal Support*,' and excerpts of a lengthy comment from one Professor of Computer Science illustrated this:

> "… I met my future wife there [at college], … and whose father was an auto mechanic. Ever since then, we have clung to each other like the country song says, "Two Sparrows In A Hurricane"… I got my first teaching job at a small college on a hilltop in Tennessee. We lived within our means, a good working-class value. People at work were friendly enough, but we didn't socialize. We socialized only at department Christmas parties, graduation parties, and such. And what was worse, my wife worked at the college as a secretary of all things! We lived like blue-collar people, so I wasn't really accepted at the college; and I was a professor with a Ph.D. so we weren't really accepted in our neighborhood…"

The sub-categories that comprise this broad category focus on various aspects of interpersonal relations. As addressed in this exemplary comment, common issues in this category involve relationship formation, spouse's social class background, occupational and educational levels, and personal values regarding lifestyles and residential location. (See Table 5 for sub-category coding assignments that emerged from the manual analysis of this text excerpt using a constant comparative approach).

In summary, the 46 initial narrow sub-categories and 5 broad coding categories that were generated from the manual constant comparative data analysis confirm Barker's (1995) findings from a qualitative study that identified a series of themes describing working-class academics; they (a) possess fewer financial resources, (b) experience feelings of invisibility due to the myth of classlessness, (c) have few working-class based professional and social support systems, (d) experience insecurity about their intellectual ability, and (e) often have fears of inadequacy in social and professional situations. In addition, the results of this analysis mirror the issues identified in a series of edited volumes of essays written by working-class academics (Ryan & Sackrey, 1984; Tokarczyk & Fay, 1993; Dews & Law, 1995; Shepard, McMillan, & Tate, 1998; Welsch, 2005; Muzzati & Samarco, 2006), that identify important characteristics and descriptors from a personal viewpoint.

Computer software data analysis results

The Tropes computerized data analysis process yielded six code levels of increasing specificity: 19 broad first-level codes, 118 narrower second-level codes, 254 third-level codes, 287 fourth-level codes, 47 fifth-level codes, and 3 codes at the narrowest level. For example, the broad first-level code 'Education & Work' contained three narrower second-level codes (Education, Employment & Work, Study) and four more levels of codes of increasing specificity. This was not an unduplicated count, however, and some results were generated in multiple broad categories. Despite this initial duplication, the large number of codes and coding categories generated several accurate and detailed data analyses.

Coding. The manual constant comparative data analysis was conducted from the individual member's point of view to explain the individual's perspective on the intersection of academic career and social class identity. However, the corresponding computerized data analysis did not begin at the same broad individual level. Instead, the Tropes software analysis provided initial results that were divided by 19 components of individual perspective into more specific coding categories. This treatment allowed the analysis to parse the data more finely, although unevenly. For instance, the two most populated categories were ‘Education & Work’ (164 occurrences) and ‘People & Persons’ (123 occurrences); the two least populated categories were ‘Nature & Wildlife’ (3 occurrences) and Sports (2 occurrences). This is not surprising, given the nature of the data on working-class academics discussing their work lives, however, these broad first level categories provide limited insights into the data due to their generic de-contextualized nature. ‘Education & Work’ for instance, could apply to any qualitative data and do not supply the researcher with any insight into the phenomenon in question. Therefore, unlike constant comparative analysis, the broad first level categories cannot be seen as the end result, and must be deconstructed by the researcher for accuracy and meaning.

For example, the initial computerized data analysis process produced several inaccurate coding assignments. In one example, the software misidentified the word “plant” as vegetation rather than a place of employment. Two text excerpts were “the plant manager took note of me and encouraged me to go to college” and “my father worked in a coal-fired electricity generating plant outside Pittsburgh.” The Tropes software coded these excerpts on four levels of increasing specificity as ‘Nature & Wildlife’ > ‘Plants’ > ‘Plant Parts’ > ‘Vegetation.’ This required researcher expertise to create a new code and re-code the text as ‘Education & Work’ > ‘Employment & Work’ > ’Plant.’ In another instance, the software interpreted the word “vermin” in the context of ‘Nature & Wildlife’ in the comment “we are treated as utterly contemptible, incompetent, stupid, greedy, subhuman vermin.” As a result, researchers’ expertise was again required to accurately identify this as a description of administrator treatment of faculty and re-code the text as ‘Education & Work’ > ‘Employment & Work.’ These examples demonstrate the importance of the researcher applying background knowledge to the analysis and decision making process and demonstrate that the role of the researcher is not supplanted by the computer analysis software.

However, after re-coding inaccurate coding assignments, the 19 computer-generated broad coding categories (compared to 5 generated manually) enabled the researchers to analyze the data in greater detail, and a more comprehensive approach. As a result, many of the 14 additional broad coding categories provided important supplemental dimensions to the concept of social class identity; codes comprising ‘Crisis & Conflicts’ facilitated a more sophisticated portrayal of family dynamics for working-class academics, and codes within ‘Arts & Culture’ gave richer meaning and depth to descriptions of working-class culture. (See Table 6 for comparison of coding categories generated by each data analysis method).

Connecting concepts. The computer software analysis was more effective in establishing connections between concepts in the data that were not recognizable in the manual constant comparative analysis. For instance, in the broad first-level code ‘Properties & Characteristics,’ the software identified 10 occurrences of the reference term ‘Difference,’ whereas the concept of differences was not addressed by the researcher in manual analysis. When the extracted text excerpts were viewed in aggregate in the computer software, it became clear that respondents’ self-perceptions of differences occurred in family settings and throughout the K-20 educational pipeline in various contexts (e.g., "The differences in all our families' lives has put a strain on our relationships…," "I found it difficult to relate to many of the undergraduates there because of class differences", and "I entered a private graduate school where I first became conscious of class differences."). The researchers' bias based on the knowledge of the previous research literature in combination with personal experience prevented recognition of this phenomenon because differences were only thought to have occurred at the undergraduate level. The software allowed for recognition and analysis of relationships among concepts without the filter of bias.

In addition, the 'Bundles' data presentation tool in the Tropes software enabled additional analysis and interpretation of chronological relationships within the data. This tool generated a constellation graphic representing the density of word occurrences that appeared in proximity to the reference term 'Difference.' Results indicated that the words 'Work,' 'School,' 'Life,' 'Religion,' and 'Luck' appeared in descending order of frequency preceding the reference term 'Difference' in the data, and identified a variety of ways in which the respondents constructed their perceptions of being different. To a certain extent, these findings are intuitive, given the context of social class differences in the academic work environment. The words 'Job,' 'Better,' 'Small,' 'Subject,' 'Live,' and 'Strain' appeared in descending order of frequency following the reference term. (See Figure 1 for constellation graphic from computerized data analysis of this code). The sophisticated data representation tools in the software allows for deeper levels of analysis that are too labor intensive for the individual researcher.

Linguistic Analysis. The software provided a linguistic analysis of the style of the text in the dataset. According to the statistical indicators retrieved during the analysis, the text was identified as 'Enunciative;' defined as 'setting some influence or revealing a point of view.' With this dataset, this finding was merely confirming that the data were indeed introductions. This feature could be useful to evaluate the alignment of style diagnosis to the nature of the data, in instances when the data is unfamiliar to the researcher.

In summary, the software produced a greater range in the levels of descriptive specificity of the codes, with some being more general (e.g., computer-generated code ‘People & Persons’ vs. constant comparative method-generated code ‘Parental Influences’), and some being more specific (e.g., three computer-generated coding levels for ‘Medicine & Health’ comprised of eleven codes vs. one constant comparative method-generated code ‘Personal Health’). While the computer software offers more precision in analysis and generation of coding categories and in the identification of hidden relationships among concepts; it does so at the expense of the researchers’ expertise and understanding of the context. In order to establish trustworthiness of the findings, the computer analysis findings must always be refined by the researcher.

Conclusions

This study began as an examination of a comparison of constant comparative analysis with a qualitative analysis computer software. One similarity in the two data analysis processes focused on the need to conduct content analysis of the data regardless of the mediating influence of the computer screen. Close reading of the data was necessary in both instances, whether in the software results window or on paper. The use of the software for analysis did not relieve the researchers of the responsibility to be intimately familiar with the data or the research literature with which to interpret and contextualize the findings.

In the constant comparative manual analysis, each piece of data was examined, analyzed and carefully placed within the context of the gestalt of the phenomenon. The researchers made conscious decisions about the importance and weight of each word, phrase and statement. With each pass through the data, the researchers themselves were changed and became more informed. Therefore, the focus of the researchers' attention became narrower and more focused. However, with the computer software, the same iterative process was not apparent. Each piece of data was examined and analyzed simultaneously, without the benefit of an overall context to give meaning. Each word, phrase and statement were treated democratically and given equal weight. Without the same narrowing of focus, the software enabled the researchers to see connections and relationships that were not apparent in the manual constant comparative approach.

The constant comparative manual method yielded results from the perspective of the individual participant, and the 5 themes that emerged from the data were all accurate broad descriptors of social class identity. In contrast, the computerized data analysis yielded 19 broad descriptors that were not specific to this data, but the software enabled data analysis and interpretation in much more depth and involved much narrower coding categories in order to access a comparable level of specificity. The text identification tools unique to Tropes and other computer software packages provide additional sophisticated data analysis capabilities. The software was able to produce accurate word counts of reference terms that would be difficult to achieve using the constant comparative method. This enumeration is one method of identifying salient reference terms that might otherwise go unnoticed. When used in combination with the graphics tools, the computer analysis processes allowed the researchers to complete horizontal data analysis (i.e., connections among narrow coding categories) and identify important connections among narrower coding categories. Analyzing these connections both vertically (i.e., connections between broad and narrow coding categories) and horizontally provides additional dimensionality that would be difficult to achieve manually.

Leech and Onwuegbuzie (2007) observed that “most researchers use one type of [qualitative] analysis and hope the results are trustworthy,” and issued a call to use two or more data analysis tools to triangulate results. The starkly different results obtained by the manual and computerized data analysis methods in this study illustrate the rationale for that call, and raise interesting questions about the trustworthiness of previous qualitative research results obtained solely by manual data analysis methods. These differences in results also raise questions about which approach is ‘better.’ The manual analysis employed a constant comparison of each new piece of data with the context and the body of data in total. Consequently, this process has an inherent trustworthiness similar to member checking or triangulation among sources. The resulting scope of the findings is narrow and vertical (i.e., connections between broad and narrow coding categories) in organization and influence. However, the computer software analysis results are much broader in scope and enables horizontal (i.e., connections among narrow coding categories) as well as vertical analysis. Whereas the software is more efficient in terms of time and energy, there is a heavier burden upon the researcher to have expert knowledge about the phenomenon under examination. Leech and Onwuegbuzie (2007) provided a useful reminder for researchers considering the use of computerized qualitative data analysis software programs: “[these] programs can help researchers to analyze their data, but they cannot analyze the data for researchers.” (p. 578)

These findings suggest that researchers need to be deliberate in the choice of qualitative data analysis tools, considering their knowledge of the research literature, understanding of the context, and subsequent ability to recognize relationships in horizontal and vertical analyses. In this way, researchers can generate trustworthy qualitative results that can be relied on with confidence.

References

Barker, J. (1995). White working-class men and women in academia. *Race, Gender & Class*, 3(1), 65-77.
Carley, K. (1993). Coding choices for textual analysis: A comparison of content analysis and map analysis. *Sociological Methodology*, 23, 75-126.
Catterall, M., & Maclaran, P. (1997). Focus group data and qualitative analysis programs: Coding the moving picture as well as the snapshots. *Sociological Research Online*, 2(1). Retrieved September 26, 2008, from http://www.socresonline.org.uk/socresonline/2/1/6.html.
Creswell, J. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.
Creswell, J. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson.
Dews, B., & Law, C. (Eds.), (1993). *This fine place so far from home: Voices of academics from the working class*. Philadelphia: Temple University Press.
Eckardt, N. (2007). The prevalence of qualitative methodology at AERA's annual meeting and the potential consequences. *Teachers College Record*, Date Published: November 08, 2007. http://www.tcrecord.org ID Number: 14741, Date Accessed: March 11, 2008.
Fallon, D. (2006). The buffalo upon the chimneypiece: The value of evidence. *Journal of Teacher Education*, 57(2), 139-154.
Fielding, N., & Lee, R. (1998). *Computer analysis and qualitative research*. London: Sage.
Freeman, M., deMarrais, K., Preissle, J., Roulston K., & St. Pierre, E. (2007). Standards of evidence in qualitative research: An incitement to discourse. *Educational Researcher*, 36(1), 25-32.
Gilbert, L. (2002). Going the distance: 'Closeness' in qualitative data analysis software. *International Journal of Social Research Methodology*, 5(3), 215-228.
Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Thousand Oaks, CA: Sage.
Hammersley, M. (1987). Some notes on the terms "validity" and "reliability." *British Educational Research Journal, 13*(1), 73-81.
Harper, S., & Museus, S. (Eds.). (2008). *Using qualitative methods in institutional assessment*. New Directions for Institutional Research, No. 136. San Francisco: Jossey-Bass.
Leech, N., & Onwuegbuzie, A. (2007). An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly*, 22(4), 557-584.
Lewins, A., & Silver, C. (2007). *Using software in qualitative research: A step-by-step guide*. Thousand Oaks, CA: Sage.
Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.
Miles, M., & Huberman, A. (1994). Qualitative Data Analysis (2nd edition). 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(1), 106-116.
Muzzatti, S., & Samarco, C. (2006). *Reflections from the wrong side of the tracks: Class, identity, and the working class experience in academe*. Lanham, MD: Rowman & Littlefield.

Ritchie, J., & Lewis, J. (2003). *Qualitative research practice: A guide for social science students and researchers*. Thousand Oaks, CA: Sage.

Ryan, J., & Sackrey, C. (1984). *Strangers in paradise: Academics from the working-class*. Boston: South End Press.

Semantic Knowledge (2006). Tropes (Version 7.0) [Computer software]. Paris: ACETIC.

Shank, G.D. (2006). *Qualitative research: A personal skills approach*, (2nd edition). Upper Saddle River, NJ: Pearson.

Shavelson, R., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Shepard, A., McMillan, J., & Tate, G. (1998) *Coming to class: Pedagogy and the social class of teachers*. Portsmouth, NH: Heinemann.

Tesch, R. (1990). *Qualitative research: Analysis types and software tools*. New York, NY: Falmer.

Thompson, R. (2002, May). Reporting the results of computer-assisted analysis of qualitative research data. *Forum: Qualitative Social Research* [On-line Journal], 3(2). Available at: http://www.qualitative-research.net/fqs-texte/2-02/2-02thompson-e.htm [Date of Access: February 18, 2008].

Tokarczyk, M., & Fay, E. (Eds.), (1993). *Working-class women in the academy: Laborers in the knowledge factory*. Amherst: University of Massachusetts Press.

Weitzman, E., & Miles, M. (1995). *Computer programs for qualitative data analysis*. Thousand Oaks, CA: Sage.

Welsch, K. (2005). *Those winter Sundays: Female academics and their working-class parents*. Lanham, MD: University Press of America.

Westphal, L. (2000). Increasing the trustworthiness of research results: The role of computers in qualitative text analysis (pp. 106). In Bengston, D. (Ed.) Applications of computer-aided text analysis in natural resources. General Technical Report. NC-211. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

Wolcott, H. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Thousand Oaks, CA: Sage.

Table 1: Emergent Codes From Constant Comparative Analysis Category ‘Parental Influences’

Data excerpt	Code for excerpt
Father is a mechanic	Employment
Is frequently ‘between jobs’	Unemployment
Second in my extended family to graduate from college	Education
They bring chaos to my life	Experiences
Talking to them leaves me feeling hopeless	Influences

Table 2: Emergent Codes From Constant Comparative Analysis Category ‘Childhood Hometowns’

Data excerpt	Code for excerpt
couldn't afford a house in Trenton	Affordability
lost the house when a highway came through	W-C Neighborhood
looked like the middle-class American dream	M-C Neighborhood
he has been there since 1969	W-C Stability

Table 3: Emergent Codes From Constant Comparative Analysis Category ‘Family Relations’

Data excerpt	Code for excerpt
just ordinary working class people	Family Class Consciousness
southern politics wouldn't allow any real sense of union solidarity	Union Family
who were in way better shape than my family was	Family Comparisons

Table 4: Emergent Codes From Constant Comparative Analysis Category ‘Individual Lived Experiences’

Data excerpt	Code for excerpt
fortunate to be among other graduate students from the working class	Graduate School Experiences
elitist attitude of many of our profs	Faculty interaction
informal support group helps to save our self-esteem	Emotional Support
frustrations of dealing with the privileged darlings	Social Class Differences

Table 5: Emergent Codes From Constant Comparative Analysis Category ‘Spousal Support’

Data excerpt	Code for excerpt
I met my future wife there [at college]	Educational Experiences
whose father was an auto mechanic	Class Background
we have clung to each other	Mutual Support
We lived within our means, a good working-class value	Values
my wife worked at the college as a secretary	Spousal Occupation

Table 6: Comparison of Broad Coding Categories From Computerized Data Analysis and Manual Constant Comparative Data Analysis

Computerized Data Analysis		Manual Constant Comparative Analysis
Education & Work	(164 occurrences)	Parental Influences
People & Persons	(123 occurrences)	Childhood Hometowns
Properties & Characteristics	(84 occurrences)	Family Relations
Other Concepts	(68 occurrences)	Individual Lived Experiences
Behaviors & Feelings	(68 occurrences)	Spousal Support
Politics & Society	(66 occurrences)	
Times & Dates	(65 occurrences)	
Health, Life, & Casualties	(64 occurrences)	
Countries & Locations	(53 occurrences)	
Business & Industry	(51 occurrences)	
Arts & Culture	(21 occurrences)	
Sciences & Techniques	(18 occurrences)	
Things & Substances	(17 occurrences)	
Thinkings & Cognition	(15 occurrences)	
Communications & Medias	(12 occurrences)	
Crisis & Conflicts	(9 occurrences)	
Agriculture & Environment	(8 occurrences)	
Nature & Wildlife	(3 occurrences)	
Sports	(2 occurrences)	

Figure 1. Constellation of Self-Identity Concepts From Computerized Data Analysis of the Code "Difference."