All Volumes (2001-2008)

The Osprey Journal of Ideas and Inquiry

2007

Qualitative Data Analysis: Comparison Between Traditional and Computerized Text Analysis

Melissa A. Bright
University of North Florida

Dawn O'Connor
University of North Florida

Qualitative Data Analysis: Comparison Between Traditional and Computerized Text Analysis

Melissa A. Bright
and
Dawn P. O'Connor

Faculty Sponsor: Dr. Robert Smallwood, Associate Provost

Most often quantitative data is used in the social sciences. Sometimes, however, researchers must use qualitative data for exploration. In a study on communication effectiveness we used both quantitative and qualitative data. After analyzing the quantitative data we used both a traditional text analysis and a computerized text analysis. Our aim in this paper is to discuss variance, or lack thereof, between outcomes utilizing traditional and computerized text analysis.

Traditional text analysis

For analysis of our qualitative data we used a traditional text analysis method based on the classic Grounded Theory. Traditional text analysis (TTA) is a process where data coders place each piece of data into various categories. Grounded Theory (GT) is a set of incorporated abstract hypotheses generated to produce a general theory using large amounts of data on a topic of interest (Glaser & Holton, 2004). In other words, grounded theory is a method used to generate theories based on qualitative data. Glaser and Holton (2004) reported that researchers often aim to produce a conceptual theory that shows a pattern of responses relevant to a problem or question being studied. Additionally, these researchers state that grounded theory produces "…a set of carefully grounded concepts organized around a core category and integrated into hypotheses."

As with any social science method, there are benefits and limitations to using TTA. Traditional text analysis, as opposed to computerized text analysis, is beneficial for several reasons. First, TTA is available to everyone- including those without access to automated software. Personnel in start-up organizations, for example, have ability to code data without having an expense of automated software. In addition, an investigator with little data coding experience can learn to traditionally code data utilizing GT methods. The GT method of data coding can be perfected with use as a coder gains coding experience.

A second benefit of TTA is human interpretation of data. Traditional text analysis allows for humans, as opposed to computers, to interpret meaning from data. For example, given a response: 'I love my professor and hate my professor's assistant," a computerized method may interpret this sentence to be both positive and negative qualifying statements for professor (i.e., a positive qualifier for professor and a negative qualifier for professor). A TTA coder may interpret this sentence to be two separate ideas (i.e., about a professor and a professor's assistant).

A third benefit of TTA is an ability to find words not in an English dictionary. A word, for example, that qualifies as slang may not be interpreted correctly when using a computerized analysis method. Additionally, if a participant responds with incorrect spelling, then that response may not be coded correctly by a computerized method.

A final benefit to TTA is an ability to better utilize automated systems based on knowledge gained from GT. Through use of GT data coding, for example, a coder will have perfected his or her ability to combine categories and will have a better idea of how to best organize data. Thus, when using an automated data analysis method a data coder can see patterns emerge even though this coder isn't as close to that data as when he or she hand codes data.

A limitation of TTA is a tendency for data coders to "read between the lines" when coding data. That is, individuals have an ability to infer (i.e., fill in the blanks)

information from what they read based on what is actually written. Individuals who read between the lines, compared to individuals who only extract exact categories listed by researchers, may more accurately categorize responses. Reading between the lines, however, while coding data may skew results. That is, individuals who infer information may inappropriately fill in blanks with what they *think* is meant by what is written not what was *intended*. Skewed results, for example, could be due to incorrectly categorizing data because of an assumption when not taking question responses literally.

A second limitation of TTA is the amount of people and time needed to code data. That is, hand coding data takes several data coders and possibly multiple days to review and code qualitative responses. On a project, for example, for which there are four qualitative questions and 200 responses to each question, hand coding data may take 10 data coders (2 coders per question to insure inter rater reliability) up to three hours to review and code responses to one question. Furthermore, if there are discrepancies in previously coded responses a third rater must review and code that data. In short, a small project may take 10 data coders and 30 to 40 hours to code data when utilizing a TTA method of coding data.

A final limitation of TTA is categorization assignment. Often, a category of "other" is selected due to a limited number of predetermined categories. One goal researchers have when looking at qualitative data is to create appropriate categories in which response options will fit. Not uncommon, however, is to find that all responses do not fit into predetermined categories. A qualitative data coder, for example, may be forced to categorize the response "I like to use web sites for gathering information because web sites are usually well organized." as "other." This may happen if researchers failed to consider "well organized" as a category option when creating a survey investigating communication with clients.

Computerized text analysis

Qualitative data can be analyzed by either a traditional text analysis method (i.e., by hand) or a computerized text analysis method (i.e., by computer). For purposes of this paper we used both traditional and computerized methods. Computerized test analysis (CTA) is exactly what an individual would expect – analysis done by use of a computer program. Researchers have also used terms such as computer assisted qualitative data analysis (Seale, 2005) and computational content analysis (Burton, 2002). There are several different computer programs used for CTA such as Computer-Assisted Evaluative Text, Program for Linguistic Content, and Map Extraction Comparison and Analysis (Wade, 2002). For purposes of this paper we will discuss CTA using SPSS Text Analysis for Surveys™ 2.0.

Just as there are benefits and limitations to TTA, there are benefits and limitations to using CTA. According to Seale (2005), using CTA helps with speed, rigor, team research and generalizability. As mentioned, when using traditional text analysis (TTA), researchers must have data coders who are willing to sit for great lengths of time and hand code each participant's response. This method is generally slow. With CTA, however, a computer may find commonalities and themes from a large quantity of data in a short amount of time.

A second benefit of CTA is an alleviation of rigor. When using TTA, researchers rely on humans to count frequencies and search for outliers. This method takes a great deal of effort and usually results in some degree of human error. With CTA, however, a computer may count frequencies and find outliers with little to no effort - figuratively speaking.

A third benefit of CTA is an ease or improvement in team research. According to Seale (2005), team research relies on use of a single data set accessible by all parties. When using TTA, researchers must have many copies of data for many individuals to categorize. When using CTA, there may be

only one copy of data which can be used by many individuals.

A final benefit of CTA is an option of using a large sample size. When using TTA, researchers may use only a small sample of data to ensure timeliness of data analysis. When using CTA, however, researchers may use a large or small sample of data and still ensure timeliness. In other words, by using a computer, researchers can analyze large amounts of data otherwise deemed too time consuming for hand coding. Additionally with a large sample size, researchers may have more generalizeable results than with a small sample size.

A limitation of computerized text analysis is a possibility of incorrect coding based on syntax (i.e., sentence structure). Researchers use computer analysis software to locate and highlight common words based on syntax. That is, researchers use computer programs to extract meaning based on a subject of a sentence found within responses. Locating information based on syntax works well unless an author incorrectly structures a sentence within their response. If, for example, a participant organizes a sentence such that an object comes first, then a verb, then a subject, that participant's response may be coded incorrectly. The following sentence is an example: "The email was sent by my professor." In this sentence, the subject should be professor but because of syntax, the subject is email. In this case, this response may be coded incorrectly using a computerized text analysis.

Another limitation of computerized text analysis is a possibility of missing valuable categories and patterns found within qualitative data. Computerized text analysis, for example, may not find relationships or categories based on word usage. That is, uncommon words found throughout open-ended responses may not be highlighted by a computerized text analysis program. However, this same information may be identified by hand coding data.

Method

Participants

Approximately 1954 Undergraduate and Graduate students from the University of North Florida participated in this study. We asked participants to complete a survey titled "Tell us What You Think…It Matters!" via their email. We told participants that information obtained from this survey would be used in efforts to better communicate with the University of North Florida community. In exchange for their participation, participants had an opportunity to put their name in a drawing for a free airline ticket.

Of the 1954 participants, 215 (11%) identified themselves as freshman, 259 (13%) sophomore, 589 (30%) junior, 544 (28%) senior, and 342 (18%) graduate or post-baccalaureate classification. Of the 1954 participants, 743 (38%) identified themselves as belonging to the College of Arts and Sciences, 453 (23%) belonging to the Coggin College of Business, 324 (17%) to the College of Education and Human Services, 266 (14%) to the Brooks Brown College of Health, 130 (07%) to the College of Engineering, Computer Sciences and Building Construction, and 35 (02%) as undecided. Researchers obtained permission from the University of North Florida's Institutional Review Board prior to beginning this experiment.

Procedure

Participants received, via email, an invitation to complete a web-based survey which included 22 quantitative (i.e., multiple choice) questions and 8 qualitative (i.e., open ended response) questions. After completion of this survey, participants forwarded their responses to a main database. Participants' surveys contained no identifying information. Additionally, we offered participants an opportunity to sign up for a chance to win a free airline ticket. For participants who chose to sign up for this give away, those participants submitted personal information (i.e., email address and phone number) to a separate database. For participants who chose not to sign up for this give away, there was no

personal identifying information forwarded to a separate database.

We imported answers for quantitative questions into SPSS for Windows™ 13.0 and analyzed those responses accordingly. For purposes of this paper, we will only address processes pertaining to responses for qualitative questions. We used both a traditional text analysis method (i.e., by hand) based on Grounded Theory and a computerized text analyses method (i.e., SPSS Text Analysis for Surveys™ 2.0).

Traditional Text Analysis.

As outlined by W. Lawrence Neuman, (2006) there are three common phases to hand coding data. Phase one is referred to as "open coding." Open coding takes place during the first reading of data. During this first reading a coder scans data for themes and common terms. Phase two is referred to as "axial coding." During this second reading a coder looks at actual data and assigns a label to themes found in phase one. Phase three is referred to as "selective coding." During this third reading a coder looks selectively for obvious illustration of themes in which to make final comparisons. As a first step for our traditional text analysis, two researchers independently brainstormed categories based on data collected for this study. That is, researchers looked over an output of data and each determined approximately ten categories per question (80 categories total). Researchers then collaborated on their categories and created one list of categories for each question. In other words, researchers created a list of categories in which each participant response would fit.

Researchers then pilot tested these categories with two students. These students were asked to use these given categories and code each response for two questions. If these students were unable to distinguish categories or were unable to use every code at least once, then researchers would need to re-create categories. These students had no problem assigning codes to each response.

We then created data printouts for each qualitative question. There were eight qualitative questions. We printed two copies of each question so that two data coders could code each question to insure inter rater reliability. There were eight qualitative questions of which we made two copies; therefore, we printed 16 copies in total.

We recruited 16 honors students to participate in qualitative data coding. We chose students who would be objective toward survey responses. Students were invited to a free lunch and a day of data coding. We then gave students detailed instructions which included respecting confidentiality and procedural rules. For each category there was a code. A category, for example, titled "convenience" could be coded "A." Each student was assigned one question for which that student coded responses. We sat with each student individually to ensure that student understood his or her responsibilities.

With two copies of each question, we gave one copy of each question to two data coders. We set up all questions to be coded forward and reverse. That is, one coder started with participant #1, a second coder started with participant #1954. This was in attempts to control for acquiescence while coding data. Additionally, paired data coders (i.e., two data coders working on the same item) were separated to ensure these coders did not converse about category options. After reviewing each set of codes for each question (i.e., 2 per question) and determining where there were discrepancies in answers, a third rater rated each question. After hand coding was complete, we entered codes into an SPSS for Windows™ 13.0 spreadsheet.

Computerized Text Analysis.

Data coding utilizing computerized text analysis we used SPSS Text Analysis for Surveys™ 2.0. We imported the eight qualitative data questions and responses used in our survey from a comprehensive data file in SPSS for Windows™ 13.0. Beginning with question one (i.e., "Please tell us why your 1st choice is the most effective for services."), we extracted preliminary categories from each question. With this

extraction we obtained preliminary categories based on frequency of occurrence. Categories based on frequency include every occurrence of a particular word or phrase. We then created categories based on linguistics. Categories based on linguistics include only ideas of words or phrases. Examine an example response "I chose my first choice because it is convenient and I like everything that is convenient." The word convenient occurs twice in this sentence but this participant is only expressing one idea. Therefore, a category based on frequency would include the word convenient twice and a category based on linguistics would include the idea convenient only once.

We then combined categories which overlapped or matched each other. For, example, a category titled 'professor' would be combined with categories titled 'instructor' or 'teacher.' In order to capture words that may be spelled incorrectly or words in multiple forms (e.g., professor, professors), we used a wildcard search (i.e., added * before and after a segment of a word).

We discontinued manipulating categories once we reached five to nine categories for each question or when we reached 97% coded for each question. We followed this same procedure for each of the eight qualitative questions in our study.

Results

Traditional Text Analysis

For purposes of this paper we will compare TTA and CTA by which categories were created. Our aim is to asses if by using computerized text analysis a researcher would find different categories and in turn make different assumptions than a researcher would find utilizing a traditional method. For our TTA we formed the following categories for each question (see tables 1-8).

Question 1c: "Please tell us why your first choice is the most effective for services."

Category explanation
Noticeable (e.g., hard to miss; prevalent; eye catching)
Convenience (e.g., fast; requires little effort)
Tangible (e.g., available off campus; accessible)
User friendly (e.g., easy to use; easy to access)
Trustworthy (e.g., valid; believable; assurance)
On the computer frequently (e.g., online; internet; websites)
Prefer one on one communication (e.g., human interaction)
Multiple messages (e.g., this source gives me more than one piece of information; it has everything I need)
None (e.g., I cannot think of an explanation, no answer; irrelevant answer)
Other (e.g., a method not coded for)

Table 1. Categories for question 1c

Question 1d: Of the 11 methods listed in (a) above, which method is least effective? Why?"

Category explanation
Option does not apply to me (e.g., I do not live on campus, I am not in any clubs)
Didn't know it existed (e.g., I did not know we had that source on campus)
Not noticeable (e.g., easy to miss; not prevalent; not eye catching)
Not convenient (e.g., not quick; requires a great deal of effort)
Not Tangible (e.g., not available off campus; not accessible)
Not User friendly (e.g., not easy to use; not easy to access)
Not Trustworthy (e.g., not valid; not believable; lack of assurance)
I am not on the computer frequently (e.g., online; internet; websites)
I do not prefer one on one communication (e.g., human interaction)
Lack of Multiple messages (e.g., this source does not give me more than one piece of information; it does not have everything I need)
None (e.g., I cannot think of an explanation, no answer; irrelevant answer)
Other (e.g., a response not coded for)

Table 2. Categories for question 1d.

Question 1d: "Tell us if you learned of an available service, a critical deadline or an important piece of campus information by some method other than those identified in (a) above. What was that communication method?"

Category explanation
Instructor (e.g., professor, teacher)
Advisor (e.g., academic advisor)
Direct Mail out (e.g., UNF mailed me something through postal mail)
UNF staff (e.g., secretaries, office staff)
Outside media (e.g., radio station, television)
Organizations on campus (e.g., Greeks, student athletics)
Actively searched for information(e.g., I found it myself by looking)
None (e.g., I cannot think of an explanation, no answer; irrelevant answer)
Other (e.g., a method not coded for)

Table 3. Categories for question 1e.

Question 3: "How do you prefer to have your questions answered or information provided?"

Category explanation
It Depends (e.g., I like certain methods for certain situations)
Pamphlets (e.g., packets of information from offices on campus)
Extended office hours (e.g., Saturday hours, evening hours)
Flexibility of several options (e.g., I like all of them; I like these three)
Internet Dialogue (e.g., I would like to talk to someone over the internet)
Satisfaction (e.g., these ways work for me)
None (e.g., no answer; irrelevant answer)
Other (e.g., a method not coded for)

Table 4. Categories for question 3.

Question 4c: "Please tell us why you pay most attention to your 1st choice when you want to find out what's happening."

Category explanation
Noticeable (e.g., hard to miss; prevalent; eye catching)
Convenience (e.g., fast; requires little effort)
Tangible (e.g., available off campus; accessible)
User friendly (e.g., easy to use; easy to access)
Trustworthy (e.g., valid; believable; assurance)
On the computer frequently (e.g., online; internet; websites)
Prefer one on one communication (e.g., human interaction)
Multiple messages (e.g., this source gives me more than one piece of information; it has everything I need)
None (e.g., I cannot think of an explanation, no answer; irrelevant answer)
Other (e.g., a method not coded for)

Table 5. Categories for question 4c.

Question 4d: "Of the 13 methods listed in (a) above, which method is least effective way of communicating what's happening on campus? Why?"

Category explanation
Option does not apply to me (e.g., I do not live on campus, I am not in any clubs)
Didn't know it existed (e.g., I did not know we had that source on campus)
Not noticeable (e.g., easy to miss; not prevalent; not eye catching)
Not convenient (e.g., not quick; requires a great deal of effort)
Not Tangible (e.g., not available off campus; not accessible)
Not User friendly (e.g., not easy to use; not easy to access)
Not Trustworthy (e.g., not valid; not believable; lack of assurance)
I am not on the computer frequently (e.g., online; internet; websites)
I do not prefer one on one communication (e.g., human interaction)
Lack of Multiple messages (e.g., this source does not give me more than one piece of information; it does not have everything I need)
None (e.g., I cannot think of an explanation, no answer; irrelevant answer)
Other (e.g., a response not coded for)

Table 6. Categories for question 4d.

Question 4e: "Tell us if you learned of an on campus event by some method other than those identified in (a) above. What was that communication method?"

Category explanation
Instructor (e.g., professor, teacher)
Advisor (e.g., academic advisory)
Direct Mail out (e.g., UNF mailed me something through postal mail)
UNF staff (e.g., secretaries, office staff)
Outside media (e.g., radio station, television)
Organizations on campus (e.g., Greeks, student athletics)
Actively searched for information(e.g., I found it myself by looking)
Outside individual (e.g., someone not related to UNF)
None (e.g., I cannot think of an explanation, no answer; irrelevant answer)
Other (e.g., a method not coded for)

Table 7. Categories for question 4e.

Question 5: "How can we improve the myWings portal so that you would use it more frequently as an information resource?"

Category explanation
Combination log-on site (e.g., allow for one log in to get to sites such as blackboard, UNF, etc.)
Include specific academic information (e.g., Program of study)
Customization (e.g., make it more individual specific; make it major/minor specific; filter announcements)
Appearance (e.g., make it less complex; less cluttered; add icons)
Post information specific to me (e.g., upcoming scholarships; job openings on campus)
User friendly (e.g., make it more simple, more easy to use)
Post deadlines (e.g., let me know when deadlines approach)
Update (e.g., update the website more frequently)
Pop-ups (e.g., eliminate the excess screens that pop-up)
None (e.g., I cannot think of an explanation, no answer; irrelevant answer)
Other (e.g., a method not coded for)

Table 8. Categories for question 5.

Based on these categories, we then had data coders assess each response. For responses which included more than one idea (i.e., a response which could be coded two different ways), coders were instructed to assign two or more codes. Because a response could be assigned two or more codes, our total for each category did not equal 1954 (total number of participants).

In addition to coding each response, coders created memos for each question. That is, if a coder read a response for which that coder felt was necessary to note uniqueness based on the nature of the question, that coder made a memo. For the question 5 "How can we improve myWings", for example, the responses varied so much that coders often read unique, original ideas. In this case, coders create memos for these questions (see table 9 for example).

Question	Memos		
5	Allow for customization		Tell us how myWings could be useful
	Let us make our own username		Use more color
	Add a search engine		Make labels and use pictures
	Add direct links		Put more interesting information on there
	Add a comic		Put today's announcements on there
	Make myWings the home page		Make teachers use it
	Put a raffle on there		Attach UNF email
	Put music on myWings		Make more announcements stand out
	Put a link back to the homepage		Add a help number
	Make it more like Osprey Online		Allow me to change my address
	Make an alphabetical listing for searches		Create tutorials
	Put something other than academics		Add events
	Make it faster		Make the icon more noticeable
	Add a calendar of events		Email me when there is a change in my account
	Put student update on myWings		

Table 9. Example of Memos for Questions in survey

Computerized Text Analysis

For our CTA we were able to conclude categories for each question.

Question 1c, ("Please tell us why your first choice is the most effective for services."), responses were coded "Use Email or Computer Frequently," "None/Generalizations/See Above," "Accessible," "Provides a Great Deal of Information/Inclusive," "Noticeable," "Easy/Quick/Convenient," "Reliable," "Interesting," and "Prefer One-on-One Human Interaction."

Question 1d, (Of the 11 methods listed in (a) above, which method is least effective? Why?"), responses were coded "Ruined by Weather," "Cluttered Web Pages," "Clutters Campus," "Not in Club/Organization," "Not User Friendly," "Little/No Access," "Only Utilize When Necessary," "Don't Know How to Utilize," "Ineffective," "Did Not Know Source Existed," "Unreliable," "No Time to Utilize," "Do Not Live on Campus," "Do Not Use," and "None/Generalizations/See Above."

Question 1e, ("Tell us if you learned of an available service, a critical deadline or an important piece of campus information by some method other than those identified in (a) above. What was that communication method?"), responses were coded "None/Generalizations/See Above," "Professors/Advisors/Instructors," "Method Already Listed," "Email," "Other," "Staff," "Students," "News and Other Media," "Blackboard," and "Mail."

Question 3, ("How do you prefer to have your questions answered or information provided?"), responses were coded "Blank/None," "It Depends," "Flexibility of Several Options," "Satisfaction with Answers," "Extended Office Hours," and "Timeliness."

Question 4c, ("Please tell us why you pay most attention to your 1st choice when you want to find out what's happening"), responses were coded "None/Generalization/See

Above," "Noticeable," "Use Email or Computer Frequently," "Easy/Quick/Convenient," "Inclusive/Provides a Great Deal of Information," "Accessible," "Reliable," "Prefer One-on-One Human Interaction," and "Interesting,"

Question 4d, ("Of the 13 methods listed in (a) above, which method is least effective way of communicating what's happening on campus? Why?"), responses were coded "Clutters Campus," "Ruined By the Weather," Cluttered Web Pages," "Not User Friendly," "Not in Club/Organization," "Little/No Access," "Only Utilize When Necessary," "Ineffective," "Do Not Know How to Utilize," "Unreliable," "Did Not Know It Existed," "No Time to Utilize," "Do Not Live on Campus," "Do Not Use," and "None/Generalizations/See Above."

Question 4e, ("Tell us if you learned of an on campus event by some method other than those identified in (a) above. What was that communication method?"), responses were coded "None/Generalizations/See Above," "Professor/Instructor/Advisor," "Method Already Listed," "Other," "Outside Media," "Email," "Blackboard," and "Mail."

Question 5, ("How can we improve the myWings portal so that you would use it more frequently as an information resource?"), response were coded "It Should Give More Information," "Make Easier to Access," "Market myWings," "Fewer Pop-Ups," "Make it Necessary/Only Use for Registration," "I am Satisfied," "Combination Log-On," "Technology Suggestions/Issues," "Customization," "Make it User Friendly," "None/Generalizations/See Above."

Discussion

To review, our aim in this study was to assess similarities and differences among qualitative coding methods. Specifically, we chose to use a hand coded method based on Grounded Theory and a computerized coding method using SPSS Text Analysis for Windows™ 2.0. Overall we found that our results (i.e., categories) from using a traditional text analysis were very similar to our results from using a computerized text analysis. Although we had similar findings, we found limitations in both methods.

Limitations of using TTA include a tendency for data coders to "read between the lines" when coding data. Data coders who "fill in the blanks" by reading between the lines may have been doing so to make sense of an idea which had not been clearly articulated. However, reading between the lines may also skew results. Skewing results may happen when a data coder inappropriately fills in blanks which were not intended by an author of a response.

Limitations of using TTA also include categorization assignment. As primary researchers, perhaps we did not create categories that should have been created. Thus, when data coders were coding our data, these data coders may have placed an excess amount of information in the 'other' category.

Some aspects of this TTA process were noteworthy. In addition to using the category "other", when coding data many responses, a majority of the time data coders found it necessary to code data as "none" (i.e., not relevant). The reason for having to use the category of "none" is because many responses had no value. That is, some respondents shared their opinion about a topic which was not of value. For example, when given the question "How can we improve the myWings portal so that you would use it more frequently as an information source?" one respondent stated "Bring back the old system!" This information is not of value to our research and thereby must be placed in a "none" category. As a matter of fact, a majority of the data we collected for this study fell into the category of "none".

Just as there were limitations for TTA, there were limitations for CTA. As researchers we wanted to explore differences among analysis of the exact same qualitative data using both a traditional text analysis method and a computerized text analysis method. One of the limitations we

encountered using computerized text analysis was that we still had humans in charge of the analysis. In other words, although we used a computer program and technology, we still relied on a human to control the chosen computer program and thus left room for human error.

A second limitation that we found by using computerized text analysis is that the program we used was unable to detect slang or spelling errors in our data. Although program writers produced a bank of possible synonyms, we still found that some words were not coded. Words such as "professor" were often spelled "proffessor," "taecher," and "isntructer." Some of these misspellings were better caught using a hand coding method than a computerized coding method. Similar to problems with slang and spelling errors, we found that the computerized text analysis we used was unable to detect specific syntax. We found that many of our participants formulated sentences by starting with an object, then a verb, than a subject. This particular method of sentence formation would not be picked up by computerized analysis.

Although there were limitations to our computerized text analysis, there were also benefits. One of the benefits of using a computerized text analysis was speed. After creating rules and advanced properties for categories the computer program we used was able to pick up hundreds of pieces of data in a very short amount of time (e.g., seconds). Comparatively, the hand coding method forced humans to code hundreds of pieces of data in a quite long time (e.g., days).

In conclusion, we found that results from a traditional text analysis (based on Grounded Theory) were very similar to results from a computerized text analysis (using SPSS Text Analysis for Windows™ 2.0). Although there were benefits and limitations for both data coding processes, researchers should consider a few key issues when deciding whether to use TTA or CTA. To make a decision a researcher should consider human factors, time allowed to complete analysis, and amount of data to code.

References

Evans, W. (2002). Computer environments for content analysis. In O. V. Burton (Ed.), *Computing in the social sciences and humanities* (pp. 68-75). Urbana and Chicago: University of Illinois Press.

Glaser, B. G., & Holton, J. (2004). Remodeling grounded theory. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 5(2). Retrieved July 13, 2006, from http://www.qualitative-research.net/fqstexte/2-04/2-04glaser-e.htm

Neuman, W. L., (2006). Analysis of qualitative data. *Social Research Methods: Qualitative and Quantitative Approaches* (pp.457-489). Boston MA: Allyn and Bacon.

Seale, C. (2005). Using computers to analyze qualitative data. In Silverman, D. *Doing qualitative research* (pp. 188-201). London: SAGE Publications.

Wade, N. (1984). Background paper in the science business: Twentieth century fund report on the commercialization of scientific research (pp. 30-31). Priority Press: New York.