ABSTRACT
         As educators become more interested in the use of
qualitative methods, a major concern is the management of the large
amounts of textual data during analysis. The efficiency and possibly
the effectiveness of manual systems, i.e., color-coded index cards
and wall-mounted systems, are reduced as the amount of textual data
increases. An emerging family of text management programs, such as
ASKSAM (Access Stored Knowledge via Symbolic Access Method), may help
address this concern by automating aspects of the analysis and, thus,
facilitating the process. A case study example that compared results
achieved through the use of ASKSAM and through a file card analysis
indicated that the data analyzed via the computerized approach
yielded essentially the same result as had previously been obtained
through the traditional analytic approach. These results suggest that
the use of a computer-based text management program can increase the
efficiency of the data analysis process without significantly
altering the interpretations made from the data. (YLB)

Computer-Based Text Management to Facilitate the
Analysis of Qualitative Data

Ronald L. Jacobs, Ph.D.
Robert D. Aron, M.B.A.

TRAINING & DEVELOPMENT
RESEARCH REPORT

Graduate Program in Training & Development

The Ohio State University

October 1987

FOREWORD

The use of computers to analyze qualitative data has become
convenient only in the past few years. Only a few studies using the
various packages and analytic techniques have been reported, however,
and the approaches need to be better detailed for practitioners.
There is at least one caution to be observed in these applications.
The cognitive processes of deciding what categories of content are
significant enough to look at, what key words/phrases
are important enough to search for, or how to synthesize similar but
not identical statements remain the province of the researcher.
Although artificial intelligence may hold some promise of aiding in
the judgmental aspects of analysis, current technology is primarily
focused on mechanics. The glamour of computerization of mechnics
should not overshadow the source of the decisions, nor should it
obscure where the responsibilty for drawing the right conclusions
rests.

In this paper, a description of one approach to analyzing
qualitative data is described. It is of interest in this research
that data analyzed via the computerized approach yielded essentially
the same result as had previously been obtained through traditional
analytic approaches. On one hand it is reassuring that there appears
to be reliability in the computer analytic approach. On the other
hand, although the study was not specifically designed to assess the
effects of computerized analysis on precision or level of detail, it
is at least somewhat disappointing that there did not appear to be

1

increased refinement of the judgments made. That the increased ease and speed with which sorts and searches can be done may improve the quality of analysis, as well as increase efficiency, is a potential effect that deserves future study.

Oliver W. Cummings, Director
Research, Methodology & Evaluation
Center for Professional Education
Arthur Andersen & Company
St. Charles, Illinois 60174

ABSTRACT

This paper reports the results of using a text management
program to analyze a set of qualitative data. Qualitative data are
generally in the form of descriptive language originating from field
observations, interview transcripts, and document analyses. A
continuing concern among many training and development professionals
is their ability to manipulate large amounts of this text data. The
efficiency and possibly the effectiveness of manual systems, i.e.,
color-coded index cards, wall-mounted systems, are reduced as the
amount of text data increases. The results here suggest that an
emerging family of text management programs, such as ASKSAM, may h lp
address this concern by automating aspects of the analysis and,   Js,
facilitating the process.

The role of inquiry has been of continuing concern and growing importance for many educators in training and development (Bratton, 1983; Jacobs, 1985). This concern has been prompted, in part, by efforts to more precisely identify the education and training needs of constituent groups, evaluate the effects of on-going programs, and investigate the appropriate roles of practitioners and researchers (Hall, 1979). It is within this context that issues regarding the use of qualitative methods and the analysis of qualitative data have been raised. The specific focus of this paper is on qualitative data analysis, which is the process of categorizing and restructuring descriptive data into more understandable and meaningful themes (Bogdan & Biklen, 1982).

Qualitative data analysis is typically iterative in that the analysis occurs concurrently with data collection, and continues until all the data have been collected. Qualitative data are generally in the form of descriptive language originating from field observations, interview transcripts, and document analyses. A continuing specific

concern of many inquirers is their ability to manipulate and sort

through large amounts of this textual information (Drass, 1980). A

number of different text-management approaches have been proposed,

including the use of color-coded index cards, wall-mounted category

systems, or coded file boxes, etc. Patton (1983) and others, however,

have noted that these manual systems are cumbersome, and their

efficiency and effectiveness are greatly reduced as the amount of data

increases.

Recent computer software advances may provide some answers to

this text-management concern by automating aspects of the data

analysis process. The purpose of this paper is to describe

computer-based text management; relate the development of

text-management programs to qualitative data analysis; present a

case-study example of how one program, ASKSAM*, was used to analyze a

set of qualitative data; and, discuss using a computer-based text

management system. The point of this paper is that while qualitative

data analysis relies primarily on the observations, interpretations,

and perceptions of the inquirer, the use of a computer-based text

management system can potentially facilitate the process.

----

*ASKSAM (Access Stored Knowledge via Symbolic Access Method) is a

commercially-available software product from Seaside Software, Corpus

Cristi, Texas 78412. Our interest in this program is only within the

context of improving the analysis of qualitative data. ASKSAM

represents a unique class of text-management programs that may also

have research utility. Other similar programs may have been used for

our application.

Computer-Based Text Management

Text management is the manipulation of textual material for the purpose of working further with that material. Computer-based text management is the accomplishment of that text manipulation by partially automating the process. Word processing and data base programs may be used for computer-based text management, though these programs have limited usefulness for qualitative data analysis. To be useful in this application, a text-management program must have a flexible word-search capability, which will allow the inquirer to locate various forms of the words or phrases of interest, instead of the exact words. Word processing programs may have some word or phrase-search capability, but the sophistication of this feature is restricted by the intent of the program. Data-base programs may have a more sophisticated word-search capability than word processing programs, but they are also limited to searching for exact word descriptors only.

There has been a progression of advancements that have enabled specialized, text-management programs to be used for qualitative data analysis. Initially, the languages and programs that possessed the desired capabilities were available only for large computers. The increased storage capacity of small computers has allowed for the subsequent development of minicomputer and microcomputer versions of the large-computer languages. For example, Drass (1980) reported the use of LISPQUAL in the analysis of a set of field notes. This program was written in the LISP language, which is now available in several versions for the microcomputer. And Enrals (1984) described the use of a microcomputer version of the PROLOG language in the analysis of

6

8

census data taken from England in 1851. While these programs have the

necessary combination of structure and flexibility required for

qualitative data analysis, their use as a general inquiry tool has

been limited. In each of the cited studies, the inquirer developed

the program from the respective language for each data analysis use.

This required the inquirer to have some prerequisite programming

skills in each of the languages.

The text-management program of particular interest here,

ASKSAM, is part of an emerging group of programs that provide the

combined capabilities required for qualitative data analysis on a

small computer, while at the same time, not requiring expert

understanding of a computer language. The program features a simple

command-like language, built-in search functions based on Boolean

logic (i.e., AND/OR logic), and a content analysis capability that

will count the frequency of words or phrases. As such, these type of

programs have the capability to allow greater flexibility in the

identification of word descriptors. This flexibility is particularly

useful during the different stages of the qualitative data analysis

process. As will be shown, the program used in this case-study,

ASKSAM, will not identify the emergent themes or structures within the

data. That capability is limited to a few experimental programs

having some degree of artificial intelligence. ASKSAM, however, does

have the capability to facilitate the identification of particular

text data which support the generation of themes by the inquirer. The

program enables the coding of text data and displays them in their

textual contexts. Each of these are important text-management

activities related to qualitative data analysis.

7

Qualitative Data Analysis

The use of a computer-based text management program is founded
on the premise that all qualitative data analysis procedures share two
common phases: a mechanical phase and an interpretive phase (Drass,
1980). The mechanical and interpretive phases are progressively
interwoven to form a descriptive fabric from apparently meaningless
threads of information. A computer-based text management program
would primarily provide for greater efficiency in the mechanical
phase, though the interpretive phase would also be affected. An
understanding of each phase might be helpful in order to fully
understand the role of a computer-based text management program.

Mechanical phase . Much routine work is required to manage
the data under analysis. This work is separate from the more
judgmental or interpretive phase of qualitative data analysis, though
the boundaries between the two are many times difficult to
distinguish. During the mechanical phase, however, data must be
stored, counted, retrieved, revised, and presented in various forms.
These activities may vary in specificity, but they share a common
element in that they are generally non-interpretive in nature. Skrtic
(1985) provides an example of the types of activities that were
documented during the mechanical phase of data analysis:

"Detailed table of contents were developed for all relevant
documents. They served as a guide to the information
contained within them. Each document was assigned a
number; relevant sub-parts within documents were noted by
page number. This coding system was used to cross
reference documentary data with card categories and to

8

relat. both of these data sources to the outline of the

case study" (p. 197).

The mechanical phase, therefore, readies the data for

subsequent judgments, interpretation, and model-building by the

inquirer. As Drass (1980) states, "Together with the interpretive

phase, they (the activities during the mechanical phase) make the data

analysis possible" (p. 334).

Interpretive phase . Qualitative data analysis is

essentially interpretive work. The inquirer must use a variety of

analytic techniques to begin to identify and generate significant

patterns and themes from the data. Crucial to the interpretive phase

is the assumed presence of an underlying structure to the data, the

meaning of which can be understood only through some theoretical

framework, brought to the process by the inquirer. The meaning of the

data is established and confirmed through further checking and

cross-checking of additional data and data sources.

Thus, the mechanical and interpretive phases are highly

related and mutually supportive. The accomplishment of the

interpretive aspects of data analysis cannot occur without reliance on

the mechanical manipulation of the data. As Drass (1980) states, ". .

. interpretive work requires data manipulation of the mechanical

operations yet, at the same time, the interpretive work will dictate

the type and, in many cases, the content of mechanical operations" (p.

337). The following statements may provide greater understanding of

how the two phases are interwoven during the inquiry:

    *Knowledge of a phenomenon is generated in situ

    (interpretive), but physically broken down into

9

thematic parts for analysis (mechanical).

*Knowledge of a phenomenon is built through
components of meaning (interpretive) as data are
categorized (mechanical).

*The Knowledge base must be readily accessible
(mechanical) so as to aid the triangulation
and auditing of knowledge; thus, confirming the
understanding of meaning (interpretive).

Computer-Based Text Management Case-Study

The subject of the case-study is a set of critical incidents
that were reported as part of a study to analyze the interpersonal
skill requirements of supervisors (Jacobs, 1986). These data were
used for two reasons. First, the analysis of critical incidents is,
in many respects, representative of the analysis approaches used with
most other forms of qualitative data. That is, whatever factors that
may exist in the critical incidents to explain their underlying
meanings is determined through an intensive process of searching for
recurring themes and ideas, confirming these themes through the
analysis of additional critical incidents and data sources, and
eventually naming the themes.

Second, the use of qualitative data that had originally been
analyzed by one of the present authors provided a unique opportunity
for comparing the manual and computer-based approaches to data
analysis. It was hoped that the author's insights about the original
setting might provide additional understanding about the differences
between the two approaches. As such, additional information (i.e.,
Personal Notes, Methodological Notes, and Theoretical Notes) were also

10

analyzed with the critical incidents, which were represented as Field Notes (See Table 1). These categories of organizing qualitative data, as suggested by Corsaro (1981), provided the opportunity for all aspects of the inquiry setting to be included in the analysis.

The following is a summary of how ASKSAM was used to analyze the data in the case-study example:

1. All text data (Field Notes, Personal Notes, Methodological Notes, and Theoretical Notes) were entered in free form, similar in the way the data might be organized by the inquirer for a personal record. Each Note was entered as a separate data record. Headings and other clerical data (i.e., Scene, Activity, Participants) were entered as structured fields.

2. Broad themes and ideas were identified through an initial analysis of the text data. For example, after reading the Theoretical Note and Field Notes, the authors hypothesized that supervisors seemed to use their offices as a setting to convey information to subordinates. Thus, the concepts of supervise and office were believed to have some significance that might justify further analysis. In that sense, ASKSAM was used here to confirm the impressions documented in the Theoretical Note by identifying the location and frequency of the words within their various textual contexts.

3. A command sequence was written to locate forms of the words supervisor and office as they appeared together in each data record. In a procedural sense, the sequence accomplished this task by first sorting all the data records using a code to distinguish the data records of interest from other records that might have been

11

13

stored on the same disk. Then the sequence reviewed the remaining

sets of records, searching for the words of interest within each

record. Table 2 presents the specific command sequence used in this

case study. Of importance, the inquirers were able to locate other

forms of the word supervisor by placing an asterisk (*) after the

prefix, supervis . As a result, other forms of the word, such as

e.g., supervise , supervisory , supervising , would also be located if

they as well appeared in the same data record with office .

    4. The results of the query were extracted and presented in

bold letters within their respective data record (See Table 3). The

results showed that the pairing of the words supervisor and office

occured wihin two Field Notes and the Theoretical Note. One

additional variation of supervisor was also located, supervisor's .

The results as presented could also have been reviewed on the display

screen or stored within a special file on the data disk for later

retrieval. The results of this analysis confirmed the relatively high

frequency of supervisor's references to their office (two of four

Field Notes), and presented all the textual contexts in which the

references occurred.

    Further interpretations of the data could have been

accomplished at this point by the inqu ers through the generation of

new memes and categories. That process, however, was beyond the scope

of this research. A theme-name code could have been inserted at these

data points to allow for the subsequent sorting and assembling of all

the text-data under that theme name, such as e.g., "Using the office

as a setting for supervisor-subordinate communications." These codes

could also facilitate the displaying of the inferential trail, which

12

14

would be helpful when conducting an external audit of the inquiry (Halpern, 1983). Additional questions that might have guided subsequent data analyses and collection activities include: Under what specific conditions do supervisors use their offices? What symbolic meaning does the office represent to supervisors? Does use of the office mediate the effectiveness of the communication between supervisor and subordinates?

5. Of note, the command sequence could have been written to allow the results from this query to be related to the results of subsequent or previous queries. In this way, many different text relationships can be investigated.

Discussion

The results of this case-study example were, in most respects, similar to those found when the data were originally analyzed using 3 x 5 file cards. In the original analysis, the supervisor-office relationship was also noted, and was subsequently integrated with information generated from other critical incidents. These data were collapsed into a more general supervisory characteristic, which was labeled, supervisor as counselor (Jacobs, 1986). One can only presume that if all 120 critical incidents had been included in the computer-based analysis, the results would also have been similar.

Of note, however, was the relative ease by which the command sequence was written and the efficiency by which the text-data were analyzed. The simple command-sequence generated for this case study would not differ in form even if more text-data had been entered. In addition, the original study required the analysis of approximately 120 critical incidents, which at that time required several half-day

13

sessions to sort and categorize. In the present example, the co-author of this paper analyzed all data records using the computer-based approach in several minutes. It is not possible to compare fully these two approaches because of the differences in the amount of text-data analyzed. Our experience, however, suggests that the use of a computer-based text management program can increase the efficiency of some aspects of qualitative data analysis to a significant degree.

The results of the case-study example suggest that the use of a computer-based text management program can increase the efficiency of the data analysis process without significantly altering the interpretations made from the data. The results presented here are similar to those reported by Drass (1980), who re-analyzed a set of field notes that were originally presented in Boys in White by Becker et al. (1961). The use of a text-management program, however, raises additional concerns that should be considered for future research on this topic. These concerns, and suggestions about them, are the following:

*Computers present information in a linear fashion only. Some important insights may not be apparent when analyzing data using a computer as they might when using a manual system. Manual systems may allow a more wholistic review of the text-data. Computer-based approaches may be limited in this way. We suggest that computer-based and manual data analysis approaches be integrated so that the advantages of each may be utilized.

*Over-reliance on a computer-based approach to

14
16

perform more of the data analysis functions than was
intended. ASKSAM, and other similar programs, are to be used
only to facilitate the mechanical aspects of the process.
Use of the program is not intended to replace, in any
fashion, the interpretive aspects of the process. We suggest
that inquirers explicate in advance the role of the
computer-based and manual data analysis approaches. This
documentation is consistent with recent advances in
ascertaining the rigor of qualitative research through
an external auditing process.

*Additional research and experimentation using computer-based
text management programs is essential. Confidence in the
reliability of these approaches has not been fully
established. The use of naive investigators using both
methods to determine whether the categories that emerge
would be parallel is desirable for future research. We
suggest that until specific guidelines are generated from
future research efforts, any computer-based approach to
qualitative data analysis should be used only for comparison
and confirmation purposes.

Conclusion

There is a growing interest in the use of qualitative methods
among educators. One major concern about the use of qualitative
methods is the management of the large amounts of text-data during
analysis. Qualitative data analysis is essentially interpretive work,
though much mechanical manipulation is also required. The mechanical
phase readies the data for subsequent judgments and interpretations

15

derived from the data. Until recently, the mechanical phase of

qualitative data analysis was accomplished mostly through various

manual systems. Interest in the use of a computer-based text

management program represents a logical extension in attempts to

increase the efficiency of the qualitative data analysis process.

These attempts may as well lead to a more thorough and imaginative

exploration of qualitative data.

16

Table 1
Qualitative Data Analyzed using ASKSAM*

==

Field Notes: An employee has been taking an inordinate amount
                of sick-leave days. The supervisor questioned
                her and explained that when she is not at work,
                other employees have to pick up the slack caused
                her absence. The supervisor asserted the
                importance of maintaining good service to the
                customer. The supervisor also advised the
                employee about the regulations requiring a
                physician's letter documenting the extent of the
                illness.

                A customer calls the postmaster and complains that
                the carrier omitted her house. The supervisor
                responded by calling the carrier into his office
                for a meeting. The supervisor listened to the
                carrier's explanation. The supervisor reasserted
                postal service policy regarding the intentional
                omission of a delivery box and suggested the
                carrier inform him of any problems in the future.

                An employee is assigned to unload a mail truck alone.
                The supervisor assessed the situation, and asked
                whether the employee required help because of the
                amount of bags in the truck and an impending deadline.
                The employee responded that he could handle the job,
                and acknowledged the concern of the supervisor.

                An employee's productivity has decreased noticeably in
                the past few weeks. The supervisor discussed the
                situation with the employee to find out the nature of
                the problem. The supervisor listened to the employee's
                discussion of a personal problem and then suggested
                some alternatives so that her work performance might
                be less affected.

Personal
    Note: Respondents have difficulty understanding my request
                for information. I have had to provide an example for
                almost every person. Generating critical incidents
                seems to be a difficult task.

Methodolo-
  gical Note: I realize that I am collecting only effective critical
                incidents. All of the stories seem to have a
                successful resolution. I must attempt to probe the
                respondents more deeply.

Theoretical
    Note: Many of the supervisors have reported using their
                offices to convey information to subordinates.
                This could be an important area to explore because
                it may affect supervisor-employee relationships.

*From Jacobs (1986)   19

Table 2
ASKSAM Command Sequence

```
:extract  {caps}  {PRINT}
{sort Page[}
RDA200-* RDA200-* {AND} supervis* supervis* {AND} office
{IMAGE ALL LINE =}
{ROW 1 COL 36 "FIELD NOTES}
{ROW 2 JUSTIFY 15 "Scene:" COL 17} Scene[
{ROW +1 JUSTIFY 15 "Activity:" COL 17} Activity [
{ROW +1 JUSTIFY 15 "Participants:" COL 17} Participants[
{ROW +1 JUSTIFY 15 "Date:" COL 17} Date[ {COL 34 "Page:"} Page[
{ROW +1 NOT LEADING 5}
```

Table 3

Results of ASKSAM Analysis

Scene: A U.S. Postal Service facility
Activity: Human Relations Critical Incidents
Participants: **Supervisors**

FN: An employee's productivity has decreased noticeably in the past few weeks. The **supervisor** discussed the situation with the employee in the **supervisor**'s **office** to find out the nature of the problem. The **supervisor** listened to the employee s discussion of a personal problem and then suggested some alternatives so that her work performance might be less affected.

FN: A customer calls the postmaster and complains that the carrier has omitted her house. The **supervisor** responded by calling the carrier into his **office** for a meeting. The **supervisor** listened to the carrier's explanation. The **supervisor** reasserted postal service policy regarding the intentional omission of a delivery nbox and suggested the carrier should inform him of any problems in the future.

TN: Many of the **supervisors** have reported using their **offices** to convey information to subordinates. This could be an important area to explore because it may affect supervisor-employee relationships.

21

REFERENCES

Becker, H., Hughes, E., & Strauss, L. (1961). Boys in white .
        Chicago: University of Chicago Press.

Bogdan, R., & Biklen, S. (1982). Qualitative research for
        education: An introduction to theory and methods .
        Boston: Allyn and Bacon, Inc.

Bratton, B. (1983, November). Instructional/training design
        competencies and sources of information about them.
        (DID/AECT Occasional Paper). Association for Educational
        Communications and Technology.

Corsaro, W. (1981). Entering the child's world: Research
        strategies for field entry and data collection in a
        pre-school setting. In J. Green and C. Wallat (Eds.),
        Ethnography and language in educational settings .
        Norwood, NJ: Ablex Publishing Co.

Drass, K. (1980). The analysis of qualitative data: A computer
        program. Urban Life , 9 (3), 332-353.

Ennals, R. (1984). Beginning micro-prolog . New York: Harper & Row
        Publishers.

Hall, B. (1979). Participatory research: Breaking the academic
        monopoly. In J. Niemi (Ed.), Viewpoints on adult education
        research . Information Series, No. 171. Columbus, Ohio:
        ERIC Clearinghouse on Adult, Career, and Vocational
        Education.

Halpern, E. (1983). Auditing naturalistic inquiries: The
     development and application of a model . Unpublished
     doctoral dissertation, Indiana University.

Jacobs, R. (1985). A rationale for using qualitative methods in
     performance technology. Performance & Instruction Journal ,
     24 (5), 20-23.

Jacobs, R. (1986). Use of the critical incident technique to analyze
     the interpersonal skill requirements of supervisors.
     Journal of Industrial Teacher Education , 23 ,(2), 56-61.

Patton, M. (1980). Qualitative evaluation methods . Beverly Hills:
     Sage Publications.

Skrtic, T. (1985). Doing naturalistic research in educational
     organizations. In Y. Lincoln (Ed.) Organizational theory
     and inquiry: The paradigm revolution . Beverly Hills:
     Sage Publications.

Ronald L. Jacobs, Ph.D., is an assistant professor and coordinates the Training and Development graduate specialization, Department of Educational Policy and Leadership, 160 Ramseyer Hall, The Ohio State University. Office telephone number: (614) 292-5037. Home telephone number: (614) 451-4709.

Robert D. Aron is a doctoral student in the areas of training and development and adult education, Department of Educational Policy and Leadership, The Ohio State University.