CHAPTER I

INTRODUCTION

Need for the Study

Evaluation practitioners often analyze narrative information in order to determine its relevance to particular study questions. The general approach best suited to this activity is called content analysis, a highly structured and systematic technique.

However, because of the inherent complexities of developing appropriate analytical constructs and the sheer volume of information that must often be analyzed, content analysis is a very difficult process to perform well. Consequently, evaluation practitioners often use impressionistic or overly simplistic methods when a content analysis of narrative information is required. As a result, their analyses often suffer from low reliability and even lower validity. They could do much better analyses by using certain content analysis techniques. The problem is, up until now, the techniques have been very difficult to use.

Fortunately, recent advancements in microcomputer technology have presented new opportunities for effectively and efficiently coping with the inherent complexities of content analysis. First, the growing availability of relatively inexpensive microcomputers has put the power of the computer age on millions of desks in American offices. Second, these machines are usually equipped with an arsenal

of general purpose programs for word processing, data base management, and the like. While many general purpose programs appear to be adaptable to content analysis methods, they were not designed with this use in mind. As a result, they are not accompanied with suggestions for how they might best be used as content analysis tools. To remedy this situation, advice and demonstrations of how practicing evaluators can use general purpose programs to implement accepted content analysis procedures on narrative information are needed.

Narrative information includes documents like books, requests for proposals, plans, and reports; or responses to open-ended questionnaire or interview questions. Structured content analyses of documents are often very difficult, tedious, and time consuming to complete, particularly as the size of the documents increases. Because of this, such analyses are usually performed impressionistically, but often with very useful results. Examples include reviews of books, movies and news conferences. While computers can aid in the analysis of large bodies of narrative information like these, it is not a substitute for the critical, insightful, and sometimes colorful analyses and syntheses done by human experts.

On the other hand, content analyses of responses to open-ended questions are typically performed more systematically--usually with specific questions, variables, and possible responses in mind. Such systematic analyses are much easier to at least partially implement on a computer than the typically ambiguous designs used to analyze large documents. As a result, evaluation practitioners are more likely to immediately benefit from using computerized techniques

focused on content analysis of responses to open-ended questions than from using computerized techniques focused on content analysis of large documents.

Common reasons for using open-ended questions are because the full range of responses to the questions is often not known or a wide range of responses is expected. Such questions also allow people to put the responses in their own words rather than be forced to choose from a fixed list of responses. In addition, open-ended evaluation questions are usually tailored to the unique characteristics of a particular study. As a result, new analytical constructs for describing or interpreting the responses must often be developed. In practice, analytical constructs are operationalized by way of the category systems into which the narrative information is coded. If computerized techniques can be used to help practicing evaluators develop these category systems, the reliability and validity of the studies in which they employ the techniques should improve immensely.

Furthermore, survey questions are usually asked of a large number of people. When these questions are open-ended, the volume of narrative responses that must be analyzed can become quite large. As a result, computerized techniques to help code this large volume of responses more reliably and validly should be a boon to practicing evaluators as well.

In summary, a new opportunity exists for practicing evaluators to use powerful, microcomputer-implemented content analysis techniques in their work. This opportunity results from the following conditions:

1. Evaluators often collect narrative information that is best understood through the use of content analysis.
2. Because of the difficulties in obtaining reliable and valid results from content analyses, evaluation practitioners can benefit from using procedural techniques developed by content analysis experts.
3. Because general purpose microcomputer programs are currently available that can be adapted to content analysis uses, evaluation practitioners can now use computerized techniques that were previously available only to a limited number of researchers with access to very expensive and specialized resources.
4. Because evaluation practitioners are already accustomed to analyzing responses to open-ended questions systematically, they are likely to immediately benefit from using computerized techniques focused on this type of content analysis problem.
5. Two potential areas for improving the reliability and validity of content analyses include the use of computerized techniques to help develop category systems and code large volumes of data.

Purpose and Objectives of the Study

In order to address the needs and opportunities listed above, the purpose of this study is to advance the body of knowledge about how evaluation practitioners can use microcomputer programs to

improve the reliability and validity of content analyses of responses to open-ended survey questions. This purpose can be met by accomplishing the following four objectives:

1. Describe conceptual and operational relationships between evaluation, content analysis, and microcomputers.

Because the fields of evaluation and content analysis evolved from a common heritage of scientific inquiry, they contain many operational similarities. On the other hand, fundamental differences between the conceptual orientations of each field also exist. A description of key evaluation and content analysis concepts allows for the fundamental similarities and differences between them to be identified. These concepts can then be used as organizers for presenting (a) overviews of evaluation and content analysis, (b) a general model for conducting an evaluation effort compatible with content analysis methods, and (c) how microcomputers can be used to help analyze responses to open-ended survey questions. These conceptual and operational relationships are discussed in Chapter 2.

2. Determine the effects of using specialized output from microcomputer programs on the reliability and validity of developing a category system for a large set of responses to an open-ended survey question used in a simulated evaluation effort.

3. Determine the effects of using specialized output from microcomputer programs on the reliability and validity of coding the set of responses in terms of the final category system used.

This dissertation study includes two procedurally overlapping experiments. The focus of the first experiment is on improving the

reliability and validity of developing a content analysis category system. The focus of the second experiment is on improving the reliability and validity of coding the set of responses in terms of an established category system. The setting for the experiments is a simulated evaluation effort that includes analyzing responses to an open-ended survey question. The mock survey was used in an evaluation of a controversial accountability system for a medium sized public school district. Participants in the experiments--volunteer students enrolled in a number of College of Education classes--were first asked to develop a category system for a set of responses to the open-ended question. Midway through the simulation and after the official category system was adopted by the hypothetical project director, the participants were asked to code all of the collected responses in terms of the final categories. The basic methodology used in the experiments is presented in Chapter 3. The results of the experiments are summarized in Chapter 4.

Three complex content analysis activities presented briefly in Chapter 3 are discussed in detail in the appendices. These activities were not performed by the experimental participants, but they still were vital to the success of the experiments. Instead, they were performed by the researcher and two different panels of experts. The activities included: (a) developing a pool of responses to the simulation's open-ended survey question by the researcher and one panel of experts (Appendix A), (b) processing all the individual participants' content analyses by the researcher (Appendix B), and (c) developing a category hierarchy by the researcher and the second

panel of experts (Appendix C).

The accounts contained in the appendices offer researchers interested in conducting studies similar to this one enough detail to adapt the activities to their own purposes. Those readers only interested in the techniques used to help analyze responses to open-ended survey questions can forego the more complex discussions found in Appendices A, B, and C.

4. Identify the limitations and benefits of the study and suggest a program of future research based on the relationships between evaluation, content analysis, and microcomputers; and the results of the study.

The use of microcomputers for content analyses of responses to open-ended survey questions has the potential to make an important contribution to evaluation practice. Chapter 5 is used to (a) present an interpretation of the research results, (b) identify the limitations of the study, (c) identify the benefits of the study, and (d) propose a program of future research based on the relationships between evaluation, content analysis, and microcomputers; and the results of the study.