United States General Accounting Office

GAO

Program Evaluation and Methodology Division

September 1996

Content Analysis: A Methodology for Structuring and Analyzing Written Material

GAO/PEMD-10.3.1

Preface

GAO assists congressional decisionmakers in their deliberations by furnishing them with analytical information. Many diverse methodologies are needed to develop sound and timely answers to the questions that the Congress poses. To provide GAO evaluators with basic information about the more commonly used methodologies, GAO's policy guidance includes documents such as methodology transfer papers and technical guidelines.

This transfer paper on content analysis describes how GAO can use this methodology in its audits and evaluations. It defines content analysis and details how to decide whether it is appropriate and, if so, how to develop an analysis plan. The paper also specifies how to code documents, analyze the data, and avoid pitfalls at each stage. Several software packages useful for GAO audits and evaluations are described.

Content Analysis: A Methodology for Structuring and Analyzing Written Material is one of a series of papers issued by the Program Evaluation and Methodology Division (PEMD). The purpose of the series is to provide GAO evaluators with guides to various aspects of audit and evaluation methodology, to illustrate applications, and to indicate where more detailed information is available.

We look forward to receiving comments from the readers of this paper. They should be addressed to Joseph F. Delfico at (202) 512-2900.

Brian P. Crowley
Assistant Comptroller General
Office of Policy

Joseph F. Delfico
Acting Assistant Comptroller General
Program Evaluation and Methodology Division

Contents

Abbreviations

AID	Agency for International Development
DOD	Department of Defense
GAO	General Accounting Office
VA	Department of Veterans' Affairs

Chapter 1

What Content Analysis Is

A Definition of Content Analysis

In content analysis, evaluators classify the key ideas in a written communication, such as a report, article, or film. Evaluators can do content analysis of video, film, and other forms of recorded information, but in this paper, we focus on analyzing words. Here is a formal definition of content analysis: it is a systematic research method for analyzing textual information in a standardized way that allows evaluators to make inferences about that information. (Weber, 1990, pp. 9-12, and Krippendorff, 1980, pp. 21-27) Another expression of this is as follows: "A central idea in content analysis is that the many words of the text are classified into much fewer content categories." (Weber, 1990, p. 12)

The classification process, called "coding," consists of marking text passages with short alphanumeric codes. This creates "categorical variables" that represent the original, verbal information and that can then be analyzed by standard statistical methods. The text passages can come from structured interviews, focus group discussions, case studies, open-ended questions on survey instruments, workpapers, agency documents, and previous evaluations.[1] Content analysis is particularly useful in GAO work because of the large quantity of written material that evaluators typically collect during a project, especially when it comes from diverse and unstructured sources.

To classify a document's key ideas, the evaluator identifies its themes, issues, topics, and so on. The result might be a simple list of the topics in a series of meeting notes. Content analysis can go further if the evaluator counts the frequency of statements, detects subtle differences in their intensity, or examines issues over time, in different situations, or from different groups.

[1]See appendix I for a brief discussion of related forms of textual analysis. Babbie (1992) and Weber (1990) give an overview of the form of content analysis we discuss in this paper.

Thus, content analysis can not only help summarize the formal content of written material; it can also describe the attitudes or perceptions of the author of that material. For example, if an evaluator wanted to assess the effects of a program on the lives of older people from their perspective, he or she could analyze open-ended interview responses to determine their outlook on life, loneliness, or security. Similarly, an evaluator could assess the effect of Voice of America broadcasts by analyzing the content of Soviet newspaper articles and radio broadcasts. (Inkeles, 1952)

The Uses of Content Analysis

Here are several ways in which GAO evaluators have successfully used content analysis techniques.

1. In Stars and Stripes: Inherent Conflicts Lead to Allegations of Military Censorship (GAO, 1988), GAO evaluators used content analysis to help assess issues of censorship, news management, and other influences on various editions of the military newspaper. Details of technique and substance from this report are used as examples throughout this transfer paper.

2. In Student Loans: Direct Loans Could Save Billions in First 5 Years With Proper Implementation (GAO, 1992c), GAO evaluators examined transcripts of focus groups discussing the difficulty of implementing a student loan program. The participants' views on whether the Department of Education could administer a direct loan program were mixed, but the evaluators were able, through content analysis, to highlight the dominant views and the reasons for them.

3. Federal Employment: How Federal Employees View the Government As a Place to Work (GAO,

1992a) reported that while the majority of the survey respondents looked favorably on working for the government, many did not. The evaluators used content analysis to assess the respondent's insightful, written comments to open-ended questions. An appendix in the report is devoted to the analysis of these comments.

4. Another excellent example of the use of content analysis appears in Women in the Military: Deployment in the Persian Gulf War (GAO, 1993c). For this study, the evaluators gave a primarily positive assessment of women's performance, using content analysis to determine that while men and women endured similarly harsh encampment facilities and conditions, both men and women considered health and hygiene problems inconsequential and their cohesion in mixed-gender units effective.

5. Among other fine examples of the use of content analyis is Veterans' Health Care: Veterans' Perceptions of VA Services and VA's Role in Health Care Reform (GAO, 1994a). The report's scope and methodology section details the analysis and summary of veterans' views that changing the VA system could, among other things, diminish or eliminate their benefits as well as harm them both emotionally and in terms of their specialized health care needs.

Other uses of content analysis in GAO reports include an analysis of transcripts of focus groups on people's ability to participate in food assistance programs (GAO, 1990), an analysis of descriptive text on the maintenance of aging aircraft (GAO, 1993a), and an analysis of open-ended discussions with Amerasian immigrants on their experiences in Vietnam, the Philippines, and the United States (GAO, 1994b).

Computerized Content Analysis

The increasing availability of written information on computer files, and the increasing number of computer programs to analyze text files, makes content analysis easier to do than ever before. Moreover, computerized programs can easily code textual data and combine them with quantitative data. The evaluator can then analyze both kinds of data with various statistical methods. However, content analysis can proceed even when written information is not available in computer files.

Some Advantages of Content Analysis

It Can Be Unobtrusive

One problem with surveys and some experimental methods is that evaluators and their informants can interact during data collection in ways other than how they would "naturally" react. For example, a content analysis of the hearing transcripts might be more useful than interviews with federal officials about what took place during public hearings on proposed environmental regulations. The officials might leave out important points, unconsciously or purposely, in order to protect themselves, but the transcripts provide the complete record. Thus, bias can be reduced during data collection. Similarly, the evaluator can eliminate from analysis survey questions that might be inappropriate because they invaded a respondent's privacy.

It Can Deal With Large Volumes of Material

Content analysis has explicit procedures and quality control checks that make it possible for only a few or a great number of evaluators to analyze large volumes of textual data. Furthermore, the explicit procedures

and quality control checks allow two or more groups of analysts to work on the same kind of data in different geographic locations, and computer software may be used to perform many of the required steps. (See appendix II.)

It Is Systematic

Content analysis can help evaluators learn more about the issues and programs they examine because it is systematic. It has structured forms that allow evaluators to extract relevant information more consistently than if they were reading the same documents only casually.

It Can Corroborate Other Evaluation Methods

When the findings from content analysis are not the main evidence in an evaluation, they can still be used to help corroborate other findings, such as responses from closed-ended surveys or from economic measures. For example, Webb and colleagues have described how investigators can use "multiple operations" to increase confidence in their findings, although we do not discuss them in this paper. (Webb et al., 1981)

Some Disadvantages of Content Analysis

Because content analysis is systematic, sufficient human resources must be committed to it and rigorously applied to it. This may mean, for some evaluation applications, that the benefits may not outweigh the cost of the resources. Moreover, while content analysis has safeguards against distortion of the evidence, evaluators must use judgment in coding the data. If the potential users of the results will be uneasy about the judgment-making process, content analysis may not be advisable. A different approach that does not convert text to categorical variables might be preferable. (See appendix I.)

How to Apply Content Analysis

GAO evaluators can use content analysis to articulate a program's objectives, describe its activities, and determine its results.

Program Objectives

Many evaluations characterize a program's objectives. For example, evaluators might compare a program's legislative objectives with those of the executive branch. To do this, they might gather written or tape-recorded information from the program's legislative history and from interviews with agency officials. In content analysis, they would then be able to compare the two kinds of documentary sources to determine whether the agency's goals conform to its legislative intent.

Program Activities

To describe a program's activities, an evaluator could perform case studies, attend agency meetings, or interview program stakeholders (for example, managers, service deliverers, or beneficiaries) and then use content analysis to examine the results. For example, GAO evaluators might ask program stakeholders open-ended questions about a program's activities and then describe them by simply tabulating the categories of activities the respondents have reported.

The extent to which program activities were accurately targeted could also be investigated. Evaluators could interview program beneficiaries and analyze their responses to assess their eligibility for the program's services. The responses could then be compared with established eligibility criteria, and the evaluators could estimate the proportion of program recipients who were truly eligible.

Program Results

When evaluators want to estimate the results of a program, they might take sample surveys, construct case studies, or examine earlier evaluation reports. When such data are quantitative, a variety of statistical procedures can be applied. (See GAO, 1992e, and Mohr, 1988.) However, to the extent that such data are textual, the evaluator can estimate program results with the help of content analysis.

Evaluators may analyze content when they are, for example, uncertain about program effectiveness criteria or when they find many diverse criteria within the program, are engaged in exploratory work, want to ensure that structured questions did not miss something, or want to clarify the meaning of close-ended questions.

The Organization of This Paper

The seven major steps in conducting a content analysis are outlined in table 1.1, along with the chapters in which we describe them.

Table 1.1: Seven Steps in Conducting a Content Analysis

Number	Step	Chapter
1	Deciding whether to use content analysis	2
2	Defining the variables	3
3	Selecting material for analysis	3
4	Defining the recording units	3
5	Developing an analysis plan	3
6	Coding the textual material	4
7	Analyzing the data	5

In chapter 2, we cover the first step: deciding whether to use content analysis by considering the kinds of questions we need to answer and the material available for evaluation. In chapter 3, we explain the planning phase in the second through fifth steps:

defining the variables we want to collect information about, defining the material to include in the analysis, defining the recording units, and developing an analysis plan.

In chapter 4, on implementing the analysis, we outline ways to code the textual material, including how to create codes and train coders. Chapter 5 covers the actual analysis and the reliability of the coding process, which affects the interpretation of the results. Chapter 6 concludes with a caution about major pitfalls to avoid.

The appendixes briefly present other methods for analyzing textual data, some of the computer software that can facilitate content analysis, and technical procedures for gauging the accuracy of content analysis.

Chapter 2

Deciding Whether to Use Content Analysis

The five major factors in considering whether to use content analysis are the objectives of the assignment, the data that are available or to be collected, the kinds of data required, the kinds of analysis required, and the resources needed. Since content analysis is often part of a broader evaluation design, the decision to use it must fit within the assignment's overall design. (GAO, 1991c) While the evaluator considers during the early stages of an evaluation design whether to include content analysis, one or more of these factors may rule out its use.

Assignment Objectives

GAO often expresses an assignment's objectives in the form of three broad categories of evaluation question: descriptive, normative, or impact questions. (GAO, 1991c) In theory, content analysis can address all three categories. In practice, descriptive and normative questions are especially amenable to content analysis; program impact questions are less commonly answered through content analysis.

Answering a descriptive question provides information about conditions or events. For example, in a report on alleged censorship of news stories in Stars and Stripes, GAO used content analysis to describe the sources and nature of articles printed in the paper's European and Pacific editions. An advisory panel of professional journalists made judgments about allegations of managing and censoring the news; GAO supplied the results of its content analysis to the panel for its deliberations.

The answer to a normative question compares an outcome to a norm, or standard. In the Stars and Stripes report, evaluators made normative comparisons between news coverage and content in the military newspaper and related stories from the Associated Press and United Press International that

had been the source for the Stars and Stripes stories. The question "To what extent does the content of news stories in Stars and Stripes indicate news management or censorship?" is normative because it implies a criterion.

Impact questions were beyond the scope of GAO's Stars and Stripes study. For example, the evaluation did not attempt to estimate the impact of a 1984 change in Department of Defense (DOD) editorial policy for the newspaper by comparing news articles before and after 1984. In another study, however, GAO evaluators did use content analysis to examine the perception of impact rather than the impact itself when they determined the views of military veterans about health care in VA hospitals. (GAO, 1994b)

Data Available or to Be Collected

Whether or not content analysis is appropriate depends on the nature of the information to be evaluated. The information can be anything written: an original document; a transcript of a speech, conversation, discussion, or oral answer to a question; or a verbal description of visual information, such as a film, video, or photograph. Documents may be government administrative records, newspaper articles or editorials, answers to questions in an interview or questionnaire, transcripts of focus group discussions, advertising copy, judicial decisions, program evaluations, descriptions of program activities, field notes, or summaries of workpapers. Some documents may already exist at the beginning of the assignment; others may have to be created through data collection during the assignment.

Kinds of Data Required

In the early stages of an assignment, evaluators choose variables of interest. For the descriptive Stars and Stripes assignment, for example, important

variables included the frequency of stories on selected issues, such as the Iran-contra affair and the presidential campaign; the percentage of stories from other sources, such as staff reporters, AP, UPI, and other wire services; and the percentage of stories that conveyed a negative DOD image. (GAO, 1988) Obviously, if documents are to be useful, they must promise to yield information on the variables of interest.

For a normative evaluation, the variables are often similar to those for a descriptive evaluation, because the only difference is the addition of a criterion in the normative evaluation. In a program impact evaluation, the kinds of data that are required include outcome variables and contextual variables that may be necessary to rule out rival explanations for the outcomes. (GAO, 1991c)

Kinds of Analysis Required

Considering data requirements goes hand in hand with analysis requirements. In many evaluations, the most important, or only, form of analysis may be a simple aggregation of quantitative data or a comparison of categorical variables. When the subject matter is textual and the evaluation questions lend themselves to numerical descriptions or comparisons, content analysis is usually a good choice. For example, in the Stars and Stripes study, a key question pertained to whether the European and Pacific editions differed in the types of stories they covered. Therefore, the evaluators classified textual data into story topic categories and displayed most of the results in simple tables that compared frequency counts for the two editions.

Had the Stars and Stripes report required a comparison of subtleties in the language of the news stories, then content analysis would probably not

have been the best methodology to use. Rather than transform the text into categories, a better approach might have been to retrieve and display comparable passages side by side. Evaluators could then systematically form conclusions about the apparent differences. (See appendix I.)

Resources Needed

In content analysis, evaluators must consider three principal types of resources: an analyst with the technical knowledge and experience to plan and direct the content analysis, personnel to do the coding, and computer capability to carry out the analysis. At least one member of the project team should know about content analysis and have experience with it. This person then takes responsibility for planning the technical aspects of the work, training the team members who will make the classifications, supervising the production of a database, and either performing or directing the statistical analysis.

Team members knowledgeable about the subject matter must carefully read the text and code its passages. Except for the very smallest textual databases, the coding process is fairly labor-intensive. For example, in a recent AID evaluation, coding 280 interviews required approximately 4 person-weeks, even with the aid of computer software. This does not include the time devoted to developing the coding system (several days), transcribing the interviews and getting them into a form suitable for computer analysis (approximately 3 person-weeks of clerical staff time), and training the coders (2 days).

The resources of personnel to do coding and computer capability to do the analysis are frequently interrelated because the coding task can be carried out with software. For most GAO content analyses,

the amount of data dictates whether analysis is to be done by computer. This means that the textual data must be suitable for computer processing and specialized programs must be available. (Appendix II reviews some of these programs.)

Chapter 3

Planning a Content Analysis

The four planning steps of a content analysis are defining the variables the evaluator wants information about, selecting the material to be analyzed, defining the recording units, and developing an analysis plan. Each step must be completed before data collection begins, although evaluators can move back and forth across the steps as they develop the evaluation design and come to grips with its practical constraints. To start the planning steps implies commitment to content analysis, but the evaluation method may still be reconsidered before resources have been committed to its implementation.

Defining the Variables

The assignment's evaluation questions lead directly to the relevant variables. In the Stars and Stripes example, we asked "To what extent does the content of news stories in Stars and Stripes indicate management or censorship of the news management?" In practice, however, defining a variable may be separated into two parts: conceptualizing the variable and specifying its categories.

Conceptualizing and Categorizing

"Conceptualizing a variable" means identifying subjects, things, or events that vary and that will help us answer the question. In the Stars and Stripes example, the two variables "news story topic" and "image of the military" were defined. News story topic was variable across the stories that appeared in Stars and Stripes, and the paper's distorted coverage of topics might indicate news management or censorship. Image of the military could also conceivably vary across the paper's stories, and imbalance in the image of the U.S. military might be another indicator of news management or censorship.

"Specifying the categories" distinguishes one subject, thing, or event from others by putting them each and severally into a limited number of categories. Thus, to completely define a variable for content analysis, we need to specify its categories. The variable's category may be either nominal or ordinal and it must be exclusive and exhaustive. Nominal variables have no intrinsic order. For example, gender can be treated as a nominal variable with two categories—male and female—but there is nothing about either category that warrants ranking one ahead of the other. Ordinal variables do have an intrinsic order. For example, attitude is often divided into categories such as greatly dislike, moderately dislike, indifferent to, moderately like, and greatly like. These categories can be ranked from top to bottom or bottom to top.[1]

Categories must be mutually exclusive and exhaustive. If they overlap, then information may be erroneously classified. Likewise, if the categories do not cover all possible classes of information, then a variable may be misclassified or not recorded at all.

News story topic in the Stars and Stripes example was a nominal variable that had five categories: acquired immunodeficiency syndrome, Iran-contra, strategic issues (such as Intermediate Nuclear Forces and the Strategic Defense Initiative), the 1988 presidential campaign, and other. Each news story could thus be conceptually labeled as fitting into one of these categories. The first four categories corresponded to politically sensitive topics, so they seemed relevant to the evaluation question. The fifth category, "other," ensured that all stories would be labeled.

Military image was also a nominal variable but it had four categories: negative, neutral, positive, and mixed.

[1]See Babbie (1992, ch. 5) for a general treatment of conceptualization and measurement and Weber (1990, sec. 2) and Babbie (1992, ch. 12) for discussions of categories.

Each news story about the U.S. military was placed into one of the categories. If the variable had had only the three categories negative, neutral, and positive, it would have been not nominal but ordinal. The category "mixed" was included because without it some stories would not have been classified. This fourth category helped ensure that the categories were mutually exclusive and exhaustive.

Determining the Number of Categories

What dictates the number of categories for a variable? Some variables seem to have an intrinsic set of categories. For example, a week can have seven categories (the seven days) or two (weekdays and weekend). For news story topic, the list of possible categories is virtually endless, so the evaluator must use judgment and be guided by the evaluation question.

In the Stars and Stripes assignment, the evaluators needed evidence to show the extent, if any, of news management or censorship. Studying all possible categories of news story was not feasible, so they chose only those for which they could determine some editorial manipulation.

The practical limit to the number of categories that can be handled is important. Both the coding process and the analytical tools available may suggest upper limits. And, certainly, the interpretation of results can become very complicated when categories are numerous. Generally speaking, the categories assigned to each variable should not exceed seven in the final steps of the analysis but may include more in the coding process because later, after the results of coding are known, evaluators can combine some categories. They may not, however, expand them.

Some ordered variables have a natural middle or neutral point. For those that do, selecting an uneven number of categories allows coders to determine a middle ground. For example, for observations about attitude, the five categories greatly dislike, moderately dislike, indifferent to, moderately like, and greatly like are better than the four categories greatly dislike, moderately dislike, moderately like, and greatly like. This is because the latter scale unrealistically forces all attitudes into either negative or positive categories.

Selecting Material for Analysis

To select textual material to include in the content analysis, evaluators may find it easiest to think first about a population of documents. For some assignments, this population may already exist, as in the Stars and Stripes evaluation. For other assignments, evaluators have to collect data into a database. This happened when GAO evaluators used focus groups to obtain responses to food assistance programs on Indian reservations. (GAO, 1990)

Defining a Document

A document should be physically separable, minimally sized, and self-contained textual information. A letter is a document. Each daily edition of Stars and Stripes is a document. A file folder is not a document because it contains within it smaller items that are physically separable, some of which are self-contained. A book is somewhat ambiguous as a document. Most books could be considered documents, but an edited book in which each chapter had separate authors might better be thought of as an aggregate of documents. A transcription of an open-ended interview would probably be defined as a document. However, if the scope of the evaluation were limited to responses to just one interview question, then a transcription of just the pertinent

answer might be the document. Thus, evaluators have latitude in defining a document. The guiding principle is to let the evaluation's purpose and needs determine the definition.

Choosing a Sampling Method

Sampling is necessary when a document population is too large to be analyzed in its entirety. Two broad options are available, probability sampling and nonprobability sampling. Probability sampling may be the right choice if the evaluation question implies the need to generalize from the sample to the population and if the procedures required for probability sampling are practical under the circumstances. Nonprobability sampling, sometimes called judgment or purposeful sampling, may be the right choice if generalization is not necessary or if probability sampling procedures are not practical. Examples of probability sampling and nonprobability sampling are GAO (1992d) and Patton (1990, pp. 169-83), respectively.

In some assignments, multistage sampling is appropriate. For example, in a study of federal personnel actions, one might first select a probability sample of personnel folders—an aggregate of self-contained documents—and then, in the second stage, sample "action" documents within the folders.

Sampling a document's segments may also be useful. For example, in a study of recommendations from GAO reports, we might probabilistically select one recommendation (that is, the recommendation itself plus its supporting material) from each of several reports. Weber (1990) recommends that documents be sampled in their entirety in order to preserve semantic coherence. However, the sampling of segments may be a good strategy when a document contains substantial amounts of material not relevant

to the study or when it is desirable to draw information from a large number of lengthy documents.

The Stars and Stripes content analysis used sampling. Since two editions of the paper had been published—dailies in Europe and the Pacific—a reasonable population of documents would be all issues of Stars and Stripes published in the decade ending in 1988. (The Congress had made its study request in 1987.) During the decade 1978-88, each edition contained 28 pages, so the document population was much too large to be studied in its entirety.

To reduce the textual material to manageable proportions, the evaluators chose a nonprobabilistic sample of documents. Specifically, they selected all issues of both the Pacific and European editions that were published in March 1987.[2] For content analysis, they chose only news stories. For comparison purposes, they also identified all AP and UPI stories that dealt with DOD and the U.S. military and with sensitive topics otherwise cited in the allegations of censorship.

Defining the Recording Unit

Once evaluators have defined the variables and selected the textual material, their next major task is to define the recording units. A recording unit is the portion of text to which evaluators apply a category label. For example, the Stars and Stripes news story was the focus of analysis in that the evaluation objective was to draw conclusions about whether the stories had been subject to news management or censorship. Therefore, the news story became the recording unit; each news story was categorized by

[2]The rationale for choosing this particular month was that it fell 6 months after the Pacific editor in chief had been appointed and before the congressional inquiry about censorship in the paper.

topic, and each news story about the U.S. military was categorized by image. In general, six recording units are commonly used: word, word sense, sentence, paragraph, theme, and whole text. (Weber, 1990)

Words

When words are the recording unit, evaluators categorize each individual word. This recording unit is well-defined because we know the physical boundaries of a word. When all words have been placed in categories, a content analysis becomes simply a word count. Although word counts would probably find limited application in GAO, knowing the frequency of key words may be useful. Most content analysis software and some other specialized forms of software can automatically count individual words.[3]

Word Sense

"Word sense" is a variation on words as units. Some computer programs can automatically distinguish between the multiple meanings of a word and can identify phrases that constitute semantic units the way words constitute semantic units. The word senses can then be counted just as if they were words. Applications in GAO are probably limited.

Sentences

Sentences may occasionally be useful recording units, especially in structured material such as written responses to an open-ended questionnaire item. Although the physical boundaries of sentences are well-defined, using them as units implies human coding, because computer programs cannot automatically classify sentences as they do words and word senses.

[3]The word count is automatic because humans do not have to code the recording units before analysis. The computer programs simply make a pass through the document, keeping count of all individual words encountered.

Paragraphs

A paragraph is a structured unit above the sentence, so it can be a recording unit. Sometimes, however, a paragraph embraces too many ideas for consistent assignment of the text segment to a single category. This leads to the problem of unreliable coding (discussed in chapter 4).

Theme

Theme is probably better suited than sentences to coding open-ended questionnaires because a theme can include the several sentences that are commonly a response to such questions. Theme is a useful recording unit, if somewhat ambiguous. Holsti describes a theme as "a single assertion about some subject" (1969, p. 116). The boundary of a theme delineates a single idea; we are not restricted to the individual semantic boundaries of sentences and paragraphs. The evaluator who defines theme as a recording unit should include guidance regarding whether, at one extreme, sentence fragments can be coded or, at the other, paragraphs or multiple paragraphs. However, even with such guidance, coders necessarily use their judgment in determining the boundaries of particular theme units and may therefore be unreliable in their coding.

Whole Text

Whole text is a recording unit larger than a paragraph but still with clearly defined physical boundaries. For example, in the Stars and Stripes assessment, a whole news story was a unit of analysis. A news story has physical and other attributes that coders can ordinarily use to distinguish it easily from editorials or syndicated columns. In the extreme, an entire document may be a recording unit. Whole-text coding is almost always unreliable.

Developing an Analysis Plan

Developing a plan for an analysis is the final planning step. It links the data back to the evaluation question.

Traditionally, most content analyses have focused on the presence of variables or their frequency, intensity, or identity by space or time.

The Presence of a Variable

Analysis sometimes focuses on the mere presence of a variable in a document. For example, in examining the roles and performance of women in the military, GAO evaluators conducted a number of focus groups and treated the transcript for each group as a document. One variable was "attitude about women's job performance," and it had two categories, positive and negative. In one part of the analysis, the evaluators simply tabulated the number of focus groups in which participants registered either positive or negative views about women's job performance. That is, a given focus group was described not by the number of positive and negative views that that group expressed but just by whether it expressed any positive or negative views.

The Frequency of a Category

Counting the number of times a category is coded is more than simply tabulating the number of documents in which the code appears. In a study of how federal employees view the government as a place to work, GAO evaluators identified 21 variables, each with two categories. For example, one variable was attitude about pay with two categories: positive and negative. The evaluators gathered answers to an open-ended question at the end of a mail-out questionnaire sent to a random sample of employees; they counted all instances in which each category was coded across all documents. (GAO, 1992b) Singleton et al. (1988) says that the frequency count is the most common method for measuring content.

Intensity

Analysis of intensity assumes ordinal categories. (GAO, 1992e) We often measure the intensity of a person's opinions or attitudes, but other kinds of intensity variables are possible. For example, in one study, coders rated the strength of association between learning outcomes and 228 different factors in 179 reviews of school learning research. Strength of association had three categories: (1) weak, uncertain, or inconsistent relationship to learning, (2) moderate relationship, and (3) strong relationship. The primary data analysis was the computation of means for groups of variables. (Wang, Haertel, and Walberg, 1990)

Space or Time

Analyzing the space or time devoted to a topic in a document is common in content analysis. For example, the newspaper space (measured in column inches) associated with a topic may reflect the importance of a topic. For television or radio, air time is a similar measure. Note that using space or time in content analysis requires more than just coding the topic. For example, in one study, evaluators first used column inches to draw conclusions about newspaper coverage of foreign news and then applied a statistical test to compare differences in coverage between newspapers that had overseas staff with those that did not. (Budd, Thorp, and Donohew, 1967, pp. 12-13)

Analysis Options

In developing a data analysis plan, evaluators depend for analysis options on the measurement level of the variables—nominal, ordinal, or interval (or ratio). (GAO, 1992e) When evaluators choose nominal variables, they commonly tabulate category frequencies, but other possibilities exist. (Reynolds, 1984) With ordinal variables come other possibilities. (Hildebrand, Laing, and Rosenthal, 1977) Interval, or

ratio, variables—which may be used in conjunction with variables coded from qualitative information—afford many possibilities for data analysis and are well covered in many statistical textbooks. (Moore and McCabe, 1989)

Chapter 4

Coding the Textual Material

Coding means to mark recording units—that is, textual passages—with short alphanumeric codes that abbreviate the categories of variables and that carry other information as well. In this chapter, we assume that all the textual material for content analysis is in computer-readable form. Figure 4.1, for example, shows a fragment of a response from an interview into which the code "[costshn-36" has been inserted above line 33. The left bracket signals to the particular software employed here that the characters that follow it are a code. The letters "costshn," standing for "cost-sharing negative," are a coded way of saying that the lines beginning with line 33 include a negative statement about cost-sharing. The number 36 means that the coded passage that begins on line 33 ends on line 36.

Figure 4.1: A Coded Passage

```
   32    should bear part of cost simply to prove
[costshn-36
   33    commitment.  But contribution level should be tied
   34    to formula; our training activities are not easy
   35    to tie to long-term sustainability; in this
   36    activity our target group can't really pay.
```

Once the document has been marked, the evaluator can relatively expeditiously analyze it by, for example, counting the codes. Counting the instances in the document of the code in figure 4.1 would tell the analyst something about the document's negativity toward cost sharing.

Creating Codes

Codes are simply abbreviations, or tags, for segments of text. Before evaluators can code a document, they must first create a code for each variable's categories. To minimize error, a code should be an abbreviated

version of a category. In figure 4.1, for example, the variable is "attitude toward cost-sharing," and it has three categories: negative, neutral, and positive, labeled n, 0, and p. When coders identify a textual statement about cost-sharing, they can easily insert the correct code because the choices are "costshn," "costsh0," and "costshp."

Many coding schemes are possible, depending on software constraints. Software usually limits the type of characters that can be used, the total number of characters in the code, and upper case versus lower case alphabetic characters.[1]

Evaluators should define their codes in a coding manual that they prepare for training coders and for their use during actual coding. The manual should at minimum contain the list of codes and what they mean and overall coding guidance.

Coding Options

Textual material can be coded directly on the computer or it can be coded manually and transferred clerically to electronic media. With the latter option, a coder works with hard-copy documents and simply marks the passages with a pencil or colored marker. Training requirements are minimal.[2]

Some content analysis software programs make it relatively easy to code directly from the computer keyboard. A document is displayed on the screen, and the coder enters codes directly into the text. The

[1]For a more general discussion of code types and coding schemes, see Miles and Huberman (1994, ch. 4).

[2]Coding can be separated into two parts: (1) the judgmental task of applying the codes to the textual material and (2) the clerical task of entering the codes into the computer. The same person need not do both tasks.

possible disadvantage of computer coding is that the coders must be trained to use the software.[3]

At least two methods are used for transferring manual codes to computer files. With some content analysis programs, the manual codes are entered from the keyboard just as if they had been entered directly in the first place. Entering manual codes electronically means that computer entry time is added directly to the manual coding time. Some programs such as Textbase Alpha, however, offer a shortcut data entry procedure for material that has been manually coded.

The best choice between coding options depends upon the material to be coded, hardware and software availability, and the experience and preference of the coders. If coders feel more confident working with hard-copy documents, then manual coding followed by one of the shortcut computer entry procedures might be preferable. If they are comfortable with direct computer entry, then overall time may be saved.

Coder Selection and Training

Coding is generally quicker and more accurate and credible the more expertise coders have in the subject of the material being analyzed. For example, in coding documents pertaining to Medicare claims, the coder's knowledge of medical terminology and practices would probably be useful.

Coders are trained to accurately apply the codes in training sessions that inform them about the purpose of the content analysis, the nature of the textual material, and the coding scheme. This explanatory information is then followed by practice with real or

[3]Direct computer coding might be awkward when a coder must move from page to page in a document in order to make a judgment about code assignment. In this circumstance, the coder might feel more confident working manually with hard-copy documents.

simulated text. Coding accuracy may require several sessions with intersession feedback to the coders.

The Potential for Coding Error

The four interrelated potential sources of coding inaccuracy in most applications of content analysis are (1) deficiencies in the documents, (2) ambiguity in the judgment process, (3) coder bias, and (4) coder error. (Orwin, 1994) For example, a poorly written document may lead to a coder's making ambiguous decisions, or ambiguity in the judgment process may set the stage for coder bias.

Deficiencies in the Documents

If a document is vague, the coder may become uncertain and make mistakes. Deficiencies in the original documents cannot usually be remedied, but coding conventions can help achieve coder consistency. For example, ambiguity in a Stars and Stripes article about weaponry may lead a coder to doubt whether to code it as a "strategic issue" or to not code it because it is really about tactics. In this case, the evaluators would do well to establish coding conventions in the coding manual and to address them during coder training.

Ambiguity in Making Judgments

In all but the most straightforward of variables, coders have to exercise judgment, and judgment opens the door for error. For example, in a study of the evaluation reports from international aid projects, coders used a five-point scale to rate the extent to which the objectives of the various aid projects had been met. At first, short phrases defined the points on the scale; at the highest level, for example, objectives were "fully achieved" or "almost fully achieved." Practice coding sessions revealed inconsistencies among the coders, so some coders suggested that a numerical scale would be better—objectives were

"90- to 100-percent achieved"—but some inconsistency still occurred. When a third scale provided both word and numeric definitions, the result was coder consistency at the necessary level. The changes to the coding instrument and the training probably both contributed to this improvement.

Coder Bias

It is hard to imagine a topic about which a coder would have no preconceptions. As Orwin notes, "Ambiguities in the judgment process and coder bias are related in that ambiguity creates a hospitable environment for bias to creep in unnoticed" (1994, p. 142). Training helps coders stay on guard against unintentional bias, and the trainers may be able to spot coders whose bias is intentional. It also helps if documents are assigned to coders randomly.

Coder Error

Coders are bound to occasionally apply the coding criteria incorrectly or just write down the wrong code. Such error can be systematic, tending to favor or disfavor certain categories, or merely random. Wise choices in constructing category labels can help avoid such mistakes, as can proper training.

Intercoder Reliability

Consistency is often referred to as "intercoder reliability." It means the degree to which different coders assign the same codes to segments of text. Much inconsistency can generate misleading data. In many circumstances, evaluators can make numerical estimates of intercoder reliability and use the results to judge the readiness of coders to proceed from training to actual coding (see appendix III). To check intercoder reliability during practice, either coders should examine the same documents or else a subset of the documents should be the same for all coders.

Systematic Error

Even when coders are relatively consistent from one to another, coding can still produce systematic error: the coders as a group tend to make the same errors in assigning category codes to segments of text. In general, gauging the extent of systematic error is more difficult than checking intercoder reliability because it implies that someone knows the "true" codes for text segments. No one in fact has such knowledge. However, evaluators may be able to detect gross levels of systematic error during training and then redefine the variables' categories and modify the coding manual.

Selecting and Managing Documents

Using All the Documents

Even though the population of documents may seem conceptually clear, assembling them for coding generally has three problems: missing documents, inappropriate documents, and uncodable documents. There may be a discrepancy between the supposed population of documents and those actually located. For example, in an evaluation of international development projects in existence over a 10-year period, the documents sought were project evaluation reports, but reports could not be found for all the projects. When documents are missing even after a persistent search, evaluators should note the probable reasons before proceeding with the content analysis. When substantial numbers of documents are missing, the content analysis must be abandoned.

An inappropriate document is one that does not match the definition of document required for the analysis. Almost inevitably, upon inspection, some

documents prove inappropriate for the content analysis. For example, in the international development study, some reports that had been labeled and indexed as project evaluation reports did not actually fit that description. Inappropriate documents should be discarded but a record should be kept of the reasons.

Some documents might match the requirements of the analysis but turn out to be uncodable. For example, missing pages or ambiguity of content raise such severe doubts about the quality of the data that it would be better not to include such documents in the analysis.

Once the set of working documents has been determined, the person in charge of coding should record each document in a log. Each document should be given a unique number, and as the coding proceeds, the following minimal information should be recorded: the coder it was assigned to, the date it was coded, and unusual problems.

Using a Sample of Documents

When the documents to be coded are a sample of a population, the sample should be chosen from the working population identified in the procedures above. See chapter 3 on selecting material for analysis for some of the sampling considerations.

Applying Codes

In manual coding on hard-copy documents, the coder simply marks the boundaries of the recording unit and writes the code in the margin of the document, as in figure 4.1. It is often helpful and speedier to use different colored pens for each variable. The procedure is similar when using a computer, but the details depend on the software. Coders can link brief comments to a recording unit in coding manually and

with some software. Such comments may be useful during data analysis to give a rationale for the code, to make cross-reference to another passage in the document, to flag the coder's uncertainty, and so on.

Codes That Overlap

In either manual or computer coding, two codes can overlap: the recording unit for one variable overlaps the recording unit for another variable. Figure 4.2 excerpts an interview in which one objective was to find out what local officials thought about their central agency's actions. The figure shows two coded variables: "weaknesses in the agency's strategies" and "consequences of agency actions." Weaknesses in the agency's strategies had three categories: inconsistency (coded here as "in"), micromanaging, and other. Consequences of agency actions also had three categories: inefficiency (coded here as "in"), vulnerability to fraud, and other. The code "[weakin-92" indicates that the passage between lines 88 and 92 identifies inconsistency as a weakness of the central agency. The code "[consin-93" between lines 89 and 93, and therefore overlapping the first code, indicates that a consequence of agency action—in this case, inconsistency—is inefficiency.

Figure 4.2: Overlapping Codes

```
[weakin-92
88          The agency needs to be more consistent in its
89          strategies and priorities.  It often appears that
[consin-93
90          they latch onto whatever fad is in fashion (e.g.,
91          AIDS, working with teenagers, etc), adopt a
92          strategy, and then alter it the following year.
93          This causes confusion and inefficiencies.
```

Nested Codes

A code is nested within another when the recording unit for one variable completely envelopes the

recording unit for another variable. Figure 4.3 shows a coded portion of an interview for two variables. One is "view about time spent on financial management" with three categories: excessive, about right, and insufficient. The other is "causes for project delays" with four categories: financial management problems, insufficient staff, supply shortages, and other. Code "[fimgtex-424" indicates that a passage between lines 416 and 424 expresses the view that time spent on financial management is excessive. Code "[delayfm-424" indicates that a passage between lines 417 and 424 attributes project delays to financial management problems. The second passage is nested within the first; as may be seen from the figure, nesting is a special case of overlapping.

Coders may find it useful to account for structured data associated with a document. When coding responses to open-ended interview questions, for example, they might want to crosstabulate the coded text variables with demographic variables such as gender, age, and ethnicity. The demographic variables can be added easily by hand or with computer software.

Figure 4.3: Nested Codes

```
[fimgtex-424
416             Financial management can take an excessive amount
417             of time.  The grantee has experienced serious
[delayfm-424
418             delays in getting financial requests approved
419             when an error is made either by the grantee in
420             its submission or by the agency in processing the
421             request.  For example, when the exchange rate
422             changes after the submission, the agency requests
423             that the grantee recalculate the budget and
424             resubmit its request.
```

Using a Computer to Code

This section assumes that the documents to be coded are available in a word processing format such as WordPerfect and that coding proceeds with the computer program called Textbase Alpha. Textbase Alpha was designed for the analysis of qualitative data, but it was not specifically oriented toward traditional content analysis. However, it is simple to use, and it performs the basic content analysis tasks. (The distinction between qualitative analysis programs and content analysis programs is described in appendix I.) There are seven steps to coding such documents.

1. Edit the documents with the word processor. While content analysis programs ordinarily have a text editor function, these are usually primitive; some analysis programs require that margins have particular settings and other special formats. With Textbase Alpha, a feature called "prestructured coding" can be used to some advantage. Suppose a document contains a series of paragraphs, each a response to an open-ended question on a mail-out interview. Pressing a Textbase Alpha function key automatically codes the paragraphs with appropriate labels such as Question 1, Question 2, and so on, so that they can be retrieved or counted by their labels. For prestructured coding to work, the first word of each paragraph must be the label and the paragraph must have a hanging indent.

2. Create an ASCII file with the word processor. It is usually necessary to strip away the word processor's formatting codes by saving the file as an ASCII file. Content analysis programs can import ASCII files. In the Textbase Alpha example, WordPerfect must be used to create the ASCII file.

3. Start the content analysis program and import the ASCII data files. Content analysis programs follow

more-or-less standard procedures for starting and importing files. Some programs require that text lines be numbered; some do this automatically and other require a separate step. In the Textbase Alpha program, the coder imports the ASCII files. Lines are numbered by clicking on a Textbase Alpha menu choice.

4. Attach codes to text segments. This involves marking the boundaries of each segment and inserting a code. With most programs, a segment starts at the beginning of a line and ends at the line's end. If a sentence starts or ends in the middle of a line, the whole line is marked. With some programs, codes are first inserted manually and then keyed into the computer. With all programs, this two-stage process is an option. Textbase Alpha is unusual in that a segment can begin or end in the middle of a line. Boundaries are marked according to cursor position, and codes are entered in a data entry box at the bottom of the screen. The coder simply moves the cursor through the text, stopping where necessary to attach codes to segments.

5. Analyze the data. Coding in effect creates a database of categorical variables. All content analysis programs have some ability to manipulate and display them. Usually the database can also be exported for further analysis with a statistical program. Textbase Alpha can calculate code frequencies for all documents or for selected documents. Individual documents can be labeled with up to 15 variables like socioeconomic factors, coder name, date, and so on. It can also count words without coding them. (We discuss data analysis at some length in chapter 5.)

6. Print the results. The printouts for most programs have limited flexibility. However, the results can usually be exported to a word processor for editing.

In our Textbase Alpha example, the results of analysis can be either printed or written to a WordPerfect file that can be opened later.

7. Export the results. Most content analysis programs can create ASCII files so that the results can be exported either to a word processing program for editing and subsequent incorporation into a report or to a statistical program for further analysis. Some programs can export files specifically for standard statistical packages such as SPSS. Textbase Alpha can construct files for display and for statistical analysis in programs such as SPSS.

Chapter 5

Analyzing the Data

The essence of content analysis is coding—that is, providing a bridge from words to numbers. Once that has been achieved, data analysis follows the usual forms of analysis. This chapter is a brief overview of the analysis tasks, since relevant statistical methods are widely available.

Preparing for Data Analysis

The basic analytic task in content analysis is to count the occurrence of codes, whether all occurrences of a given category (for example, all occurrences of Stars and Stripes articles that portray a negative image of the military) or only certain subcategories of occurrences (for example, separate counts of such articles in the Pacific and European editions). Planning the counting task in advance avoids duplicative and unnecessary effort. However, using computer programs to do the counting lessens the burden and helps the analysis evolve (assuming, of course, that the appropriate variables have been coded).

The choice of software is important because programs differ substantially. A form of analysis that might be easy to implement with one program can be awkward or even impossible with another. (Appendix II gives a brief summary of this variation.) Evaluators should consult with someone who is familiar with several types of software before choosing one and may find it advisable to use more than one computer package.

Estimating Reliability

When several coders code the documents, their consistency is important. If the coders differ substantially, then the results of the content analysis become questionable. Chapter 4 outlined steps for minimizing unreliability. Another important step is to assign selected documents to several coders at once

so that estimates of reliability can be made (see appendix III).

Counting a Code's Frequency

Drawing inferences from the frequency of codes is the simplest and often the most useful form of data analysis. Drawing conclusions in the Stars and Stripes assignment, evaluators counted the number of articles that presented a negative image of the military and compared the number to the number of wire service articles with negative images. Because the numbers were sufficiently large, the evaluators used percentages. The analysis showed that 47 percent of the wire service stories portrayed a negative image but that the European edition had only 35 percent and the Pacific edition 27.

Finding Associations

Beyond simply counting, evaluators might look for an association between two or more variables. In the Stars and Stripes assignment, the frequency of news articles on various topics was compared between the Pacific and European editions. In the language of content analysis, the variable "topic" was compared to the variable "edition." Topic had the subcategories military, Iran-contra, AIDs, strategic treaty, and presidential campaign.

The final Stars and Stripes report contained a table similar to table 5.1, with which we can examine the association between topic and edition. If the data were to show that knowledge of one variable provides us with knowledge about the other, we would then say that the variables were associated. For example, suppose we have a bin containing 100 randomly chosen articles from the Pacific edition and 100 from the European edition. If we randomly select one article from the bin, and if it is about Iran-contra, does knowledge about that topic tell us which edition the

article appeared in? If the answer is yes, the two variables are associated.

Table 5.1: Frequency of Stars and Stripes Stories on Selected Topics

	Pacific edition		European edition	
Topic	**Number**	**Percent**	**Number**	**Percent**
U.S. military	71	52.2	144	55.8
Iran-contra	33	24.3	45	17.4
AIDS	17	12.5	33	12.8
Strategic treaty	10	7.4	20	7.8
Presidential campaign	5	3.6	16	6.2
Total	**136**	**100.0**	**258**	**100.0**

Table 5.1 shows that the percentage of articles on Iran-contra was somewhat greater in the Pacific edition; the percentage of articles on the presidential campaign was somewhat greater in the European edition. The remaining categories do not show much difference. Thus, there may be a weak association between topic and edition. That is, topic only is somewhat predictable from edition, or edition only is somewhat predictable from topic.

A table like this may disclose a relatively strong relationship between variables, but often the relationship is ambiguous. By subjecting the data to a statistical analysis, moderate or weak associations can readily be established. Because both variables are unordered—that is, they are nominal variables—we could compute a statistic like Cramer's V with statistical software.[1] Cramer's V ranges from 0, indicating no association, to 1, indicating perfect association. The data in table 5.1 yield a value for V turns of 0.09, a very modest degree of association.

[1]Sometimes the statistic is called Cramer's C—as, for example, in Seigel and Castellan (1988).

Reporting the Methodology and Results

The methodology and results of a content analysis should be reported the way they are for other evaluations. The methodology should be described in sufficient detail that readers will have a clear understanding of how the work was carried out and its strengths and limitations. For example, the report should reveal

- the evaluation question addressed;
- the nature of the material analyzed;
- the variables coded and the coding categories;
- whether documents were sampled and, if so, how;
- the recording units;
- the coding procedures and copies of coding instruments;
- the statistical analysis techniques; and
- limitations that would prevent another from using the information correctly.

The verbal conclusions from the content analysis should be backed up by tables and statistical summaries. Where it is applicable, evaluators should include statements about the statistical precision of the findings.

Chapter 6

Avoiding Pitfalls

Evaluators planning a content analysis should be aware of some pitfalls ahead of them. The ready availability of relevant material can lead to aimless and expensive fishing expeditions motivated by the hope of turning up something interesting. Quantifying documentary information may produce important and interesting data, but mere counting for the sake of counting is likely to produce precise but meaningless or trivial findings. Below are some steps to take to avoid the pitfalls.

Planning

Be Clear About the Questions

The evaluation questions drive the study. If they are ambiguous or not suited to the users' needs, even a well-implemented method will produce findings of doubtful value. To be clear about the questions means to state them as specifically as possible so that the answers will be useful to decisionmakers. One exception to this rule—probably the only exception—is when the main purpose of the study is for evaluators to learn systematically about a substantive area in preparation for doing a main study. When this is the goal, the findings may not be directly useful to decisionmakers, but they should be a stepping stone to subsequent studies designed to serve policy needs.

Consider the Broad Options

Content analysis is only one approach to drawing conclusions from textual data. Other options that allow for the retrieval and manipulation of actual segments of text are briefly discussed in appendix I. The textual methods referred to there may be better suited to answering some evaluation questions than content analysis.

Define the Variables Carefully

The need for careful definitions of the variables, including the specification of their categories, cannot be overstated. Pitfalls abound: defining variables that cannot be used to answer the evaluation questions, defining variables that are so ambiguous as to defy reasonable categorization and interpretation, specifying categories that are not mutually exclusive and exhaustive, and specifying categories ambiguously so that coders can work only capriciously. Faulty definition is one of the main contributors to unreliability in the coding process.

Defining the variables should begin early because the definition may require a restatement of the evaluation questions. The possibility of redefinition should extend into the implementation phase, because training coders constitutes a test of the categories and may reveal problems in making the connection between the variables' definitions and the assignment of codes.

Define Recording Units Carefully

The selection of recording units is based upon the nature of the variables and the textual material to be coded. For a given variable, different recording units can produce different findings. Therefore, considerable thought must go into the decision on recording units. Later, the coders must understand the recording units and apply them in a way such that the reliability of the coding process is maintained. When the recording units have obvious physical boundaries, as whole text, paragraphs, and words do, the coder's task is relatively easy. When the theme is a recording unit, as it often is in an evaluation, extra precautions must be taken to avoid unreliability.

Develop an Analysis Plan

The steps in content analysis are deceptively simple and may therefore tempt the evaluator to postpone

serious thought about data analysis until coding has been completed. This would be a mistake. In designing and implementing a content analysis, evaluators will come to several decisions that bear on whether the analysis will be possible. These decisions—most notably, defining the variables, defining the recording units, and choosing the software—should not be made until after a preliminary data analysis plan has been developed. Otherwise, the evaluator may arrive at the time for data analysis and find some important options foreclosed.

Plan for Sufficient Staff and Time

Content analysis can be time-consuming. A coding manual must be prepared and, probably, revised several times. Coders must be trained and given time to practice coding until their reliability is satisfactory. These two steps alone can easily take a couple of months. The time required for the final coding process depends upon the amount of material to be coded, the number of variables, the number of coders, and the judgment required for coding decisions. Careful definition of variables will help keep the need for judgment to a minimum but, in most analyses, some variables will be complex and subtle and coding decisions will take time.

Coding

Produce a Coding Manual

A good coding manual is indispensable. Avoid the temptation to save time by not producing one or by producing only the skeleton of one. The time spent in being complete will be more than repaid by making the coders' task easier and faster and, especially, by ensuring coding of the highest quality.

Train the Coders Thoroughly

Good training is essential. Even experienced coders need to learn about the aims of the evaluation, the material to be coded, and the coding system. They may also need training in the software. Inexperienced coders will additionally need guidance in good coding practice—keeping proper records, adopting tactics for avoiding errors, knowing when to seek advice, and so on. All coders need practice in applying the coding system to examples of the material to be coded.

Pretest the Coding System

Pretests can be carried out in conjunction with training. Pretests with the persons who will do the final coding affords the opportunity to fix problems by redefining variables, especially the categories. Coders-in-training can give direct feedback on the difficulties they have with the coding system. There is no substitute. Pretests also provide a means for making preliminary estimates of reliability. Indeed, actual coding should not begin until reliability is satisfactory.

Develop Management Procedures

A single person should be given overall responsibility for the document coding. The best choice is usually someone who has coding experience and who will also perform some of the coding as a head coder. This person should develop detailed procedures for keeping track of documents, assigning them to coders, and maintaining a log of the process. Usually the head coder also provides the first level of troubleshooting: responding to queries from coders, resolving ambiguities about categories, and making at least preliminary decisions to remove problematic documents from the database.

Analyzing and Reporting the Data

Cross-Check Preliminary Results	Things are not always what they seem. Try to verify findings by using related variables or slightly different analysis methods. This is also a time to check on the reliability of the coding process.
Apply Statistical Tests	In some circumstances, statistical tests of significance may be appropriate. Use them to rule out chance as an explanation for the results.
Make External Comparisons	Compare the content analysis results to other forms of evidence, either in the same evaluation or from the literature on the topic.
Do Not Overstate the Conclusions	Remember the origins of the data and the assumptions they are based on. Confidence in the answers to evaluation questions and the forcefulness of the implications derived from them must fit the data and the methodology. Sometimes confidence is high but, at other times, the conclusions must be carefully qualified.

Appendix I

Analysis of Qualitative Data

Content analysis applies to textual information in the form of words. An analyst can classify text into categories as described in chapter 1. The categories are treated like numerical data in subsequent statistical manipulations. The statistical analysis permits the analyst to draw conclusions about the information in the text. This is the traditional form of content analysis.

Content analysis, as defined in this paper, can be viewed as being one among a number of methods for analyzing textual data. Under the title of qualitative data analysis, Tesch (1990), describes many possibilities for analyzing textual data. A number of those alternatives classify text into categories but do not give numerical labels to the categories in preparation for statistical manipulation. (See for example, Miles and Huberman (1994) and Strauss and Corbin (1990).) Analysis in these other qualitative approaches typically involves graphic manipulation and display of text segments in the form of either codes or actual words rather than statistical manipulation. Content analysis is usually confined to statistical analysis.

We might want to address some of the evaluation questions with textual data. These questions are best answered with content analysis and other forms of qualitative analysis. To a degree, software programs such as AQUAD can be used in either situation (Tesch, 1992). AQUAD was designed for the style of qualitative analysis that retains the text segments intact. It basically offers the ability to cut and paste coded segments of computerized documents. Its ability to count codes also gives it some content analysis capability.

In designing an evaluation that will use qualitative data, consideration should be given to a variety of

approaches, including but not limited to content analysis. As always, the methods the analyst chooses should be matched to the evaluation questions.

Appendix II

Software for Content Analysis

This appendix describes computer software that may be useful to content analysis. The list of programs here is by no means complete, and it is purely descriptive, not a GAO endorsement of any program. The descriptions focus on features of the software that are necessary or optional for use in content analysis; they do not refer to other features that are not relevant to content analysis.

The content analyst must carry out several of these six functions:

1. Edit: generate and edit recorded information, including the creation of ASCII files.

2. Code: mark recording units and attach category codes.

3. Search: identify specific words, phrases, and categories.

4. Count: count the number of specific words, phrases, or categories in each recording unit.

5. Retrieve: retrieve specific words, phrases, or categories.

6. Export: create a computer file for analysis by statistical packages.

Therefore, the software in table II.1 is described in this appendix primarily in regard to these functions. The table is organized so that the software with the greatest number of features is at the top, the least at the bottom.

Table II.1: Software Features Relevant to Content Analysis[a]

Software	Edit	Code	Search	Count	Retrieve	Export
askSam	+	+	+	+	+	—
Textbase Alpha	—	+	+	+	+	+
AQUAD	0	+	+	+	+	—
TEXTPACK PC	—	—	+	+	0	+
Micro-OCP	—	—	+	+	—	—
WordCruncher	0	—	+	+	—	0
WordPerfect	+	—	+	—	0	0

[a]The software feature is adequate or better = +. The feature is somewhat limited but not totally absent = —. The feature is absent = 0.

askSam

askSam was designed not for content analysis but as a general purpose database manager that can handle structured and unstructured qualitative and quantitative data.[1] This description of its features is based on askSam version 2.0a for Windows.

askSam has been used in several GAO projects that involved the analysis of large amounts of textual information, including (1) transcripts of focus group discussions; (2) structured interviews consisting of 100 questions asked of 200 persons, several of the questions being open-ended; (3) a COBOL database transformed into an askSam database consisting of thousands of records, each including one open-ended

[1]Other free-form database managers include Concordance and ideaList. See Côté and Diehl (1992) for a review.

free text field; and (4) an automated version of the GAO open recommendations report.[2]

Text to be coded could be prepared on a word processor and converted to an ASCII file and then imported to askSam. However, askSam can import information directly in a variety of formats such as dBase and WordPerfect (5.x and 6.0). The program's built-in word processor is relatively flexible and can be used to enter data.

Text passages can be coded from within askSam's word processor by text-editing. That is, while the text is displayed on the screen, a code is typed in at the beginning of the passage and a single character is placed at the end of the passage. A form of automatic coding is also available; a selected character that appears in the raw text, a colon for example, can serve as a code, or field character. The text that follows that code, on the same line, can be analyzed as a coded passage.

The program has strong search capabilities for words (including codes) and phrases. Words and phrases can be counted, thus providing the basis for content analysis. The full texts for all instances of a code can also be retrieved and displayed on the screen or printed. There is no simple way to export the results of code counts to statistical programs for further analysis.

askSam's great versatility makes it harder to learn and somewhat more awkward to use than some of the

[2]The GAO applications mentioned here were performed with earlier versions. A number of GAO applications of askSam have been performed in conjunction with GAO's Questionnaire Programming Language (QPL). Procedures for converting QPL data files, containing the results of focus groups or open-ended interview questionnaires, for example, are given in GAO (1991b), pp. 156-63. The document also describes some of the analytical steps that may be carried out on the converted files.

more specialized programs such as AQUAD and Textbase Alpha.

Textbase Alpha

Textbase Alpha was developed for the qualitative analysis of data from interviews. Although not designed for content analysis, it has some numeric analysis features, and it can produce an output file that SPSS can use directly for categorical data analysis.

Text to be coded is prepared on a word processor and converted to an ASCII file. A separate data file is created for each document. Supplementary data, such as identifiers and demographic variables may be added at this time.

In coding, the analyst moves the cursor to mark the beginning and end of a recording unit and then keys the code so that it appears in a special data entry box at the bottom of the screen. The program also includes a prestructured coding feature in which the paragraph format of the text (prepared in the word processor) leads to a form of automatic coding. This may be especially useful for handling the responses to interviews whose paragraph-like structure corresponds to a series of questions.

Textbase Alpha has flexible procedures for text retrieval by code. A search may be made across all documents or only selected ones (for example, only Hispanic respondents if ethnicity has been added as a demographic variable). The results of searching text passages are saved in an ASCII file, which can be viewed on screen or imported into a word processor for editing.

The frequency of some or all codes can be counted, with the results also stored in an ASCII file. The

program will also count all or selected words in the textual material, and the count can be made for all or selected documents.

The program can construct an SPSS file in which each document corresponds to an SPSS case. Demographic variables and codes become SPSS variables.

AQUAD

Like Textbase Alpha, AQUAD was developed primarily for the analysis of qualitative data in circumstances in which there is no intent to transform the results to numbers. However, AQUAD has several features that make it useful for content analysis.

Textual material is prepared on a word processor and converted to ASCII files for processing by AQUAD. Each document constitutes one file. For example, if 10 interviews were conducted, 10 ASCII files would be prepared.

Coding in AQUAD can be performed with the textual material displayed on the screen as on a word processor. The cursor is moved to the line where the passage to be coded begins, and the code is entered. The code carries three kinds of information: the line where the segment begins, the line where it ends, and the category label. If the analyst prefers to mark the codes on hard copy first, AQUAD provides a shortcut by which they can be entered into the database.

Even though it was not designed as a content analysis program, AQUAD can be used to count code frequencies and to retrieve the coded passages in their entirety.

TEXTPACK PC

TEXTPACK PC was designed for analyzing open-ended survey questions but over the years it has been extended to a variety of applications such as content analysis and literary and linguistic analysis.

In Version V, Release 4.0, for MS/DOS, the text to be coded is prepared on a word processor, which also produces an ASCII file that the program can read. All documents are included in a single file. TEXTPACK PC transforms that file to others in TEXTPACK format for use in the actual analysis. The program has minimal text-editing capability; editing is best done with a word processor.

In coding, the analyst specifies a code "dictionary" of words, sequences of words, and word roots (that is, the beginnings of words). The dictionary is created in the form of an input file for TEXTPACK PC, and the coding is automatic in that the computer looks for and counts the matches of "words" in dictionary and character sequences in the text file. Unlike Textbase Alpha and AQUAD, the recording units that are counted are limited to words, phrases, or word roots in the text. TEXTPACK PC also performs a simple word frequency count (that is, without counting sequences or word roots) without the necessity of creating a code dictionary.

The text retrieval feature identifies and displays words in context. A dictionary file is used to specify the "words" to be searched. Results are displayed in standard KWIC format with identifying information so that each occurrence can be traced back to its location in the text.

A frequency count of codes, produced as described above, can be saved to a file in a form that SPSS and SAS.

Micro-OCP

Micro-OCP is the microcomputer implementation of a mainframe concordance program known as OCP, or Oxford Concordance Program. A concordance is an alphabetical list of words showing the context of each occurrence of each word. It makes word lists with frequency counts, indexes, and concordances from texts in a variety of languages and alphabets. Although designed especially for literary analysis in which individual words are the recording units, the program can be used to perform content analysis by using a somewhat limited form of coding.

As with most other programs, the textual material would ordinarily be generated by a word processing program and converted to ASCII format for importation to Micro-OCP. To perform a content analysis, the analyst also requires a "command" file, which can be developed with a word processor or Micro-OCP. The command file is, in effect, a set of instructions that tells Micro-OCP what it is to do with the textual material.

Text passages can be coded with a word processor by inserting code characters at the beginning of a passage, but there is no way to mark the end of a passage. It is therefore possible to count the occurrence of codes, but the ability to retrieve a coded passage is limited, except when words are the recording units.

Different kinds of text passages can be marked (Micro-OCP calls the markings "references") for later use in the analysis. For example, when the textual material is composed of answers to a series of interview questions, all responses to question 1 could be marked "Q1," those to question 2 "Q2," and so on. By appropriate use of Micro-OCP commands, a given content analysis could then be limited to responses to question 1, for example.

Micro-OCP searches for words and brings back the results in one of three basic forms: a word list, an index, or a concordance. Typical content analysis applications are producing (1) a word list of codes, along with the frequencies of the codes, (2) a concordance of selected words as a preliminary to other forms of analysis, (3) a concordance of codes as a crude way to retrieve partial text passages, and (4) an index of selected words or codes to provide the basis for a second-stage "look-up" of words or codes in the text. Used in these ways, Micro-OCP can provide a rudimentary form of content analysis.

WordCruncher

WordCruncher indexes text files and retrieves and manipulates data from them for viewing or analysis.[3] WordCruncher is primarily designed to display the text associated with words or word combinations (that is, the context). It also provides a count of the number of instances of each word and a way of creating a free-standing thesaurus, facilitating the development of categories for a content analysis.

Before analysts use WordCruncher for content analysis, they generate the text material and code it in a word processor. (Under some circumstances, WordCruncher generates second- and third-level codes automatically.) The codes consist of two parts: a reference symbol and a reference label (such as "question10"), which identify the location of words in the text.

Once the text has been coded, WordCruncher is used to produce an index—a list of words along with their frequencies. Then, when the analyst highlights a word and presses the enter key, the program finds each instance of the word and displays its context.

[3]Other text-indexing and retrieval software includes Folio Views and re:Search. See Côté and Diehl (1992) for a review.

WordPerfect

A word processing program, such as WordPerfect, is indispensable for carrying out a content analysis. It can be used to create a textual database for later use with other programs, to edit an existing database, to attach codes necessary for content analysis, and to convert from a word processor format to ASCII format. Virtually all word processors can perform these tasks and their editing capabilities are usually much superior to the primitive editing features found in most specialized content analysis programs.

Some word processors have powerful search features that are useful during the early stages of content analysis. WordPerfect has QuickFinder, which searches for words and phrases within files and across files. The analyst can then scroll through the text to find the words and phrases that QuickFinder has highlighted. Used in this way, the program can be helpful in defining variables and categories and in deciding what material to code.[4]

QuickFinder File Indexer is an enhanced search utility included in WordPerfect 5.2 and later versions. An index of all words in a file or files is created and saved as a basis for all searches. Using the index greatly increases the speed of the search.

QuickFinder allows the analyst to specify quite complex word patterns through the use of search modifiers. Thus, the analyst can search for files containing

- each one of a set of words (Boolean AND);
- any one of a set of words (Boolean OR);
- one word but not another;
- particular word forms (using "?" and "*" as wild-card characters);

[4]Many other file-indexing packages (such as Isys, Magellan, and ZyIndex), independent of word-processing packages, are available. See Côté and Diehl (1992) for a review.

- phrases (words next to each other);
- two words within n number words of each other; and
- two words in the same line, sentence, paragraph, page, or section (between two hard pages).

Appendix III

Intercoder Reliability

An important measure for judging the quality of a content analysis is the extent to which the results can be reproduced. Known as intercoder reliability, this measure indicates how well two or more coders reached the same judgments in coding the data. Among the variety of methods that have been proposed for estimating intercoder reliability, we discuss three.

A simple and commonly used indicator of intercoder reliability is the observed agreement rate. The formula for this is

$$P_o = \frac{n_a}{n_o}$$

where P_o = observed agreement rate,
n_a = number of agreements, and
n_o = number of observations.

Table III.1 gives an example from Krippendorff (1980). Coders A and B have each assigned category labels 0 or 1 to a total of 10 recording units. They agree in 6 out of 10 cases, so

$$P_o = \frac{6}{10} = 0.6$$

Table III.1: Codes Applied by Two Coders

	Recording unit									
Coder	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
A	0	1	0	0	0	0	0	0	1	0
B	0	1	1	0	0	1	0	1	0	0

Although this indicator is simple, the observed agreement rate is not acceptable because it does not

account for the possibility of chance agreement. This is important because even if two coders assign codes at random, they are likely to agree at least to some extent. The expected agreement rate arising from chance can be calculated and used to make a better estimate of intercoder agreement.

The chance agreement rate is fairly easy to compute when the data are redisplayed as in table III.2. Each pair of observations from coders A and B will fall into one of four cells: (1) A and B agree that the code is 0, (2) A codes 0 and B codes 1, (3) A codes 1 and B codes 0, and (4) A and B agree that the code is 1. If we count the number of instances of each pair, the results can be displayed as in table III.2.

Table III.2: Observed Co-occurrences of Codes

	Coding by B		
Coding by A	**0**	**1**	**Total**
0	5	3	8
1	1	1	2
Total	**6**	**4**	**10**

The following formula gives the chance agreement rate:

$$P_c = \frac{1}{n^2} \sum_{i=1}^{c} n_{i.} n_{.i}$$

where P_c = chance agreement rate,
$n_{i.}$ = observed row marginals (from table III.2),
$n_{.i}$ = observed column marginals (from table III.2), and
n = number of observations.

Using the numbers in table III.2, the chance agreement rate is

$$P_c = \frac{(8)(6)+(4)(2)}{10^2} = 0.56$$

Now the observed agreement rate of 0.6 does not look so good because, by chance, we could have expected an agreement rate of 0.56.

The chance agreement rate is accounted for in a widely used estimate of intercoder reliability called Cohen's kappa (Orwin, 1994). The formula is

$$K = \frac{P_o - P_c}{1 - P_c}$$

where K= kappa,
P_o= observed agreement rate, and
P_c= chance agreement rate.

With the data in table III.2, kappa is

$$K = \frac{0.6-0.56}{1-0.56} = 0.09$$

Kappa equals 1 when the coders are in perfect agreement and equals 0 when there is no agreement other than what would be expected by chance. In this example, kappa shows that the extent of agreement is not very large, only 9 percent above what would be expected by chance.

Kappa is a good measure for nominal-level variables, and it is computed by standard statistical packages such as SPSS. Seigel and Castellan (1988) discuss kappa, including a large-sample statistic for significance testing. Kappa can be improved upon

when the variables are ordinal, interval, or ratio. Krippendorff (1990) provides very general, but more complicated, measures. Software programs for computing such variables have been developed in some design and methodology groups within GAO.

Bibliography

Babbie, E.R. Survey Research Methods. Belmont, Calif.: Wadsworth, 1973.

Babbie, E.R. The Practice of Social Research, 6th ed. Belmont, Calif.: Wadsworth, 1992.

Berelson, B. Content Analysis in Communication Research. New York: Free Press, 1952.

Budd, R.W., R.K. Thorp, and L. Donohew. Content Analysis of Communications. New York: Macmillan, 1967.

Carmines, E.G., and R.A. Zeller. Reliability and Validity Assessments. Beverly Hills, Calif: Sage, 1979.

Côté, R.C., and S. Diehl. "Searching for Common Threads." Byte, 1992, pp. 290-305.

Eisner, M. "Long-term Dynamics of Political Values in International Perspective: Comparing the Results of Content Analysis of Political Documents in the USA, GB, FRG and Switzerland." European Journal of Political Research, 18 (1990), 605-21.

Fox, D. "Techniques for the Analysis of Quantitative Data." In The Research Process in Education. New York: Holt, Rinehart, and Winston, 1969.

Graesser, A.C., S.E. Gordon, and L.E. Brainerd. "QUEST: A Model of Question Answering." Computer Mathematics Applications, 23:6-9 (1992), 733-45.

Hildebrand, D.K., J.D. Laing, and H. Rosenthal. Analysis of Ordinal Data. Newbury Park, Calif.: Sage, 1977.

Holsti, O.R. Content Analysis for the Social Sciences and Humanities. Reading, Mass.: Addison-Wesley, 1969.

Inkeles, A. "Soviet Reactions to the Voice of America." Public Opinion Quarterly, 16 (1952), 612-17.

Jobson, J.D. Applied Multivariate Data Analysis. Vol. 2. Categorical and Multivariate Methods. New York: Springer-Verlag, 1992.

Kaplan, A., and J.M. Golden. "The Reliability of Content Analysis Categories." In H.W. Lasswell et al. (eds.), The Language of Politics: Studies in Quantitative Semantics. New York: George Stewart, 1949.

Kolbe, R.H., and M.S. Burnett. "Content-Analysis Research: An Examination of Applications with Directives for Improving Research Reliability and Objectivity." Journal of Consumer Research, 18 (1991), 243-50.

Kovach, C.R. "Content Analysis of Reminiscences of Elderly Women." Research in Nursing and Health, 14 (1991), 287-95.

Krippendorff, K. Content Analysis: An Introduction to Its Methodology. Newbury Park, Calif: Sage, 1980.

McTavish, D.G., and E.B. Pirro. "Contextual Content Analysis." Quality and Quantity, 24 (1990), 245-65.

Mauch, M.C. "Belief Schemata and the Use of Discrepant Base Rates in Judgment Making." Ph.D. dissertation. Catholic University, Washington, D.C., 1986.

Miles, M.B., and A.M. Huberman. Qualitative Data Analysis: An Expanded Sourcebook, 2nd ed. Thousand Oaks, Calif.: Sage, 1994.

Mohr, L.B. Impact Analysis for Program Evaluation. Chicago: Dorsey Press, 1988.

Moore, D.S., and G.P. McCabe. Introduction to the Practice of Statistics. New York: W.H. Freeman, 1989.

Mosteller, F., and D. Wallace. Inference and Disputed Authorship: "The Federalist." Reading, Mass.: Addison-Wesley, 1964.

North, R.C., et al. Content Analysis: A Handbook with Applications for the Study of International Crisis. Evanston, Ill.: Northwestern University Press, 1963.

Orwin, R.G. "Evaluating Coding Decisions." In H. Cooper and L.V. Hedges, The Handbook of Research Synthesis. New York: Russell Sage Foundation, 1994.

Patton, M.Q. Qualitative Evaluation and Research Methods, 2nd ed. Newbury Park, Calif: Sage, 1990.

Ramallo, L.I. "The Integration of Subject and Object in the Content of Action: A Study of Reports Written by Successful and Unsuccessful Volunteers for Field Work in Africa." In P.J. Stone et al. (eds.), The General Inquirer: A Computer Approach to Content Analysis in the Behavioral Sciences. Cambridge, Mass.: MIT Press, 1966.

Reynolds, H.T. Analysis of Nominal Data, 2nd ed. Newbury Park, Calif.: Sage, 1984.

Robinson, W.S. "The Statistical Measure of Agreement." American Sociological Review, 22 (1957), 782-86.

Scott, W.A. "Reliability of Content Analysis: The Case of Nominal Scale Coding." Public Opinion Quarterly, 19 (1955), 321-25.

Seigel, S., and N.J. Castellan, Jr. Nonparametric Statistics for the Behavioral Sciences, 2nd ed. New York: McGraw-Hill, 1988.

Shapiro, D.H., and D.E. Bates. "The Measurement of Control and Self-Control: Background, Rationale, and Description of a Control Content Analysis Scale." Psychologia, 33 (1990), 147-62.

Simonton, D.K. "Lexical Choices and Aesthetic Success: A Computer Content Analysis of 154 Shakespeare Sonnets." Computers and the Humanities, 24 (1990), 251-64.

Singleton, R., Jr., et al. Approaches to Social Research. New York: Oxford University Press, 1988.

Spiegelman, M., et al. "The Reliability of Agreement in Content Analysis." Journal of Social Psychology, 37 (1953), 175-87.

Strauss, A., and J. Corbin. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park, Calif.: Sage, 1990.

Tesch, R. Introductory Guide to Textbase Alpha. Desert Hot Springs, Calif.: Qualitative Research Management, 1989.

Tesch, R. Qualitative Research: Analysis Types and Software Tools. New York: Falmer Press, 1990.

Tesch, R. AQUAD User's Manual. Desert Hot Springs, Calif.: Qualitative Research Management, 1992.

U.S. General Accounting Office. HUD's Evaluation System: An Assessment, PAD-78-44. Washington, D.C.: 1978.

U.S. General Accounting Office. Stars and Stripes: Inherent Conflicts Lead to Allegations of Military Censorship. GAO/NSIAD-89-60. Washington, D.C.: 1988.

U.S. General Accounting Office. Food Assistance Programs: Recipient and Expert Views on Food Assistance at Four Indian Reservations, GAO/RCED-90-152. Washington, D.C.: 1990.

U.S. General Accounting Office. Designing Evaluations, GAO/PEMD-10.1.4. Washington, D.C.: 1991a.

U.S. General Accounting Office. QPL Reference Manual, Version 3.0, GAO/HRD Technical Reference Manual 6. Washington, D.C.: 1991b.

U.S. General Accounting Office. Federal Employment: How Federal Employees View the Government as a Place to Work, GAO/GGD-92-91. Washington, D.C.: 1992a.

U.S. General Accounting Office. Quantitative Data Analysis: An Introduction, GAO/PEMD-10.1.11. Washington, D.C.: 1992b.

U.S. General Accounting Office. Student Loans: Direct Loans Could Save Billions in First 5 Years With Proper Implementation, GAO/HRD-93-27. Washington, D.C.: 1992c.

U.S. General Accounting Office. Using Statistical Sampling. GAO/PEMD-10.1.6. Washington, D.C.: 1992d.

U.S. General Accounting Office. Aircraft Maintenance: FAA Needs to Follow Through on Plans to Ensure the Safety of Aging Aircraft, GAO/RCED-93-91. Washington, D.C.: 1993a.

U.S. General Accounting Office. Developing and Using Questionnaires, GAO/PEMD-10.1.7. Washington, D.C.: 1993b.

U.S. General Accounting Office. Women in the Military: Deployment in the Persian Gulf War, GAO/NSIAD-93-93. Washington, D.C.: 1993c.

U.S. General Accounting Office. Veterans' Health Care, GAO/HEHS-95-14. Washington, D.C.: 1994a.

U.S. General Accounting Office. Vietnamese Amerasian Resettlement, GAO/PEMD-94-15. Washington, D.C.: 1994b.

Vogt, W.P. Dictionary of Statistics and Methodology. Newbury Park, Calif.: Sage, 1993.

Wang, M.C., G.D. Haertel, and H.J. Walberg. "What Influences Learning? A Content Analysis of Review Literature." Journal of Educational Research, 84:1 (1990), 30-43.

Webb, E.J., et al. Nonreactive Measures in the Social Sciences, 2nd ed. Boston: Houghton Mifflin, 1981.

Weber, R.P. Basic Content Analysis, 2nd ed. Newbury Park, Calif: Sage, 1990.

Yin, R.K. "Evaluation: A Singular Craft." Paper presented at the annual meeting of the American Evaluation Association, Seattle, Washington, 1992.

Glossary

ASCII File	A type of personal computer file used to exchange information between applications. Constructed in accordance with specifications of the American Standard Code for Information Interchange.
Categorical Variable	Distinguishes among subject, timing, and event by putting them into a finite number of categories.
Code	A short alphanumeric term that refers to the category of a variable and often the location of a text passage. To code is to mark a text segment with a code.
Coder	A person who analyzes textual material and applies codes to text segments.
Intercoder Reliability	The degree of coding consistency between two or more coders.
Nominal Variable	A categorical variable in which the categories have no inherent order.
Ordinal Variable	A categorical variable in which the categories have an inherent order.
Qualitative Data Analysis	A broad range of techniques, such as content analysis, for analyzing nonnumerical information, usually textual material but sometimes pictures, audio recordings, videos, and so on.
Recording Unit	A portion of text that a category label is applied to.

Papers in This Series

This is a flexible series continually being added to and updated. The interested reader should inquire about the possibility of additional papers in the series.

The Evaluation Synthesis. GAO/PEMD-10.1.2.

Content Analysis. GAO/PEMD-10.1.3.

Designing Evaluations. GAO/PEMD-10.1.4.

Using Structured Interviewing Techniques. GAO/PEMD-10.1.5.

Using Statistical Sampling. GAO/PEMD-10.1.6.

Developing and Using Questionnaires. GAO/PEMD-10.1.7.

Case Study Evaluations. GAO/PEMD-10.1.9.

Prospective Evaluation Methods: The Prospective Evaluation Synthesis. GAO/PEMD-10.1.10.

Quantitative Data Analysis: An Introduction. GAO/PEMD-10.1.11.

United States
General Accounting Office
Washington, D.C. 20548-0001

Official Business
Penalty for Private Use $300

Address Correction Requested

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100