

### III. Content Analysis

#### A. Topics appropriate to content analysis:

1. Content analysis can be used with any form of communication:
  - a. The early uses of content analysis were in studies of the nature and effects of propaganda, and this type of analysis later spread to studies of news media, education, and then communication in general.
  - b. Content analysis is usually thought of in connection with the mass communication media, but it is by no means limited to these. It may be used with any kind of written material, as well as with pictorial and sound materials.
  - c. Frequent artifacts for content analysis are books, magazines, poems, newspapers, songs, movies, videos, paintings, speeches, letters, laws, & constitutions.
2. Answering "who says what, to whom, why, how, and with what effect?"
  - a. As a mode of observation, the focus is on the what.
  - b. As an analysis of data, content analysis addresses the why and with what effect.
3. Content analysis is the equivalent in document study to survey research. It is conducive to the use of formal hypotheses, large scientifically drawn samples, and quantitative data that can be analyzed with computers and modern statistical techniques. In addition to hypothesis testing, other purposes for content analysis are:
  - a. To describe trends in communication content.
  - b. To relate known characteristics of sources to messages they produce.
  - c. To audit communication contents against standards.
  - d. To analyze techniques of persuasion.
  - e. To analyze style.
  - f. To relate known attributes of the audience to messages produced for them.
  - g. To describe patterns of communication.
4. There are three main emphases of content analysis:
  - a. Content analysis may provide information about the characteristics of

the originator's culture or about the originator.

- b. The interest may be in the message itself, such as the relative effectiveness of alternative messages, or a comparison of messages from a single source at different times or under different circumstances.
  - c. A content analysis study may tell something about the effects of the message on the target audience.
5. Examples of research problems studied through content analysis:
- a. Attempts at identifying the source of messages by analysis of sentence length and frequency of various classes of nouns (could be used to determine unknown or disputed authorship)
  - b. Studies of the content of messages by counts of words and symbols, by amount of space devoted to various topics, by presence of different types of bias, and by comparison of the frequency of occupations portrayed on television with the frequency of those occupations existing in society.
  - c. Studies of the differential content of advertising messages as related to the audience toward which the magazines were aimed.
  - d. Studies of the frequency of items in a message.

**B. Sampling in content analysis:**

1. Units of analysis are the individual units about which or whom descriptive and explanatory statements are made.
  - a. Doing content analysis becomes more complicated when the units of observation differ from the units of analysis.
  - b. It is extremely important to be clear about the unit of analysis, because sample selection depends on what the unit of analysis is.
2. Sampling techniques: Content analysis of written communication may occur at any or all of the following levels: words, phrases, sentences, paragraphs, sections, chapters, books, writers, or the social contexts relevant to the works.
  - a. The primary step is to compile a sampling frame, or a list of all units from which the sample is to be drawn.
  - b. After a sampling frame is constructed, any of the probability sampling techniques may be used in content analysis -- random, systematic, stratified, or cluster.

- c. The process of sampling does not need to end when the researcher reaches the unit of analysis.

C. **Coding in content analysis:**

1. Communications may be coded on manifest and latent content:
  - a. Manifest content is the visible, surface content of a communication. This approximates the standardization in survey questionnaires.
  - b. Latent content is the underlying meaning of the communication. This parallels the subjective component of field research.
2. Units and categories -- the content analysis procedure involves the interaction of two processes:
  - a. Reliability-specificity: Specification of the content characteristics to be measured; and
  - b. Validity-depth: Application of the rules for identifying and recording the characteristics when they appear in the texts to be analyzed.
3. Conceptualization and the creation of code categories: categories should reflect the purposes of the research, and should be exhaustive, mutually exclusive, and independent. By independent, this means that the value of one category does not determine the value of another category.
4. Relationship between theory and research:
  - a. Induction: An inductive approach entails the researcher's initial immersion in documents of some type in order to identify the dimensions, or themes, which seem meaningful to the producers of the documents. A greater reliance upon induction, especially at the onset of a content analysis, is more consistent with a qualitative approach. When content analysis is carried out in an inductive or mixed manner it resembles the process of building grounded theory advocated by Glaser and Strauss in conjunction with participant observation.
  - b. Deduction: A deductive approach is followed when the categories are suggested to the researcher by a theoretical perspective and the documents provide a convenient source of data on which to test a hypothesis. In a purely deductive approach, there is little interplay between theory building and theory testing.
  - c. Different levels of measurement may be used in content analysis -- nominal categories, ordinal ranking, interval, or ratio. Keep in mind that the level of measurement implicit in your coding scheme does not necessarily reflect the nature of your variables.
  - d. Pretest a coding scheme before using it in content analysis. You

should decide what manifest or latent contents of communications will be regarded as indicators of the different attributes composing your research variables.

5. Counting and record keeping: Most content analysis research is quantitative. Quantification may be performed by using one of the following systems of enumeration:
  - a. An appearance system that calls for searching the material for appearance of a certain attribute. The size of the context unit determines the frequency with which repeated recording units occurring in close proximity to each other are counted separately.
  - b. A frequency system in which every occurrence of a given attribute is recorded. Knowing the exact frequency with which a category appears is more valuable than merely knowing whether or not the category appears at all.
  - c. A time/space system that is based on various measures of space (column inches, front page, size of headline) or units of time (minutes a radio/tv devotes to a news item, etc.) to describe the relative emphases of different categories in the analyzed material. This measures how much attention is devoted to the category.
  - d. An intensity system, which is generally used in studies dealing with attitudes and values. Methods of quantifying for intensity are based on the construction of scales.
  
6. Coding units of analysis vs. units of observation:
  - a. The recording unit is the smallest body of content in which the appearance of a reference is counted (a reference is a single occurrence of the content element). The major recording units are:
    - 1) Single word is usually the smallest unit employed. The problem with using a single word is that there may be too many words for the researcher to manage. The single word is seldom used, except in studies dealing with readability, style, psychotherapy, and literary detection.
    - 2) The theme refers to the moral purpose, or goal of a document. Since the theme has no spatial boundaries like words, there is less consensus about where the theme begins and ends, lowering intercoder reliability.
    - 3) The character as a recording unit is limited to such documents as novels, plays, television programs, or movies. The category systems for characters usually involves ascribed or achieved statuses. The recording unit is the particular person, and the number of persons fitting into each of the categories is recorded.

The character is concrete and unambiguous, and usually the numbers are manageable.

- 4) The sentence or paragraph has the advantage of having clear boundaries. A disadvantage is that they often contain more than one topic or theme, making them not mutually exclusive for coding.
  - 5) The item is the recording unit when the entire article, film, book, or program is characterized. This unit may be too vague, and may present problems when items fall between two categories.
- b. The context unit is the largest body of content that may be examined in characterizing a recording unit.
  - c. For example, the recording unit may be a single term; but in order to describe whether the term is treated favorably, one has to consider the sentence in which the term appears (the context unit). The sentence is taken into account when recording coding the term.
7. Record the base from which the counting is done. Every observation should be coded according to one of the attributes making up a variable.

**D. Advantages of content analysis:**

1. Content analysis is economical in terms of both time and money. A staff and expensive equipment are not required.
2. Content analysis is objective, systematic, and typically quantitative. These features make it reliable and valid. One should be able to consistently come up with the same results and the results should accurately reflect the data from a communication.
3. Content analysis can handle large volumes of data; and it is easy to repeat part or all of the coding to make sure the information is coded in a consistent manner.
4. Content analysis permits longitudinal studies--one can study processes occurring over long periods of time.
5. Content analysis is able to deal with unstructured material. This means that documents that were not prepared with scientific analysis in mind (diaries, letters, songs, etc.) can be meaningfully and successfully analyzed for patterns and relationships.
6. Content analysis involves sensitivity to context and symbolic forms in communication. Situational, semantic, and political consequences can be researched. It is also possible to analyze data from a document in a context quite different from the one used in the source.

7. Content analysis is unobtrusive, and seldom has any effect on the subject being studied.

-

E. **Disadvantages of content analysis:**

1. Content analysis is limited to the examination of recorded communications. The communications may be oral, written, or graphic, but they must be recorded in some fashion to permit analysis.
2. Perhaps the most persistent criticism of content analysis is that it is a subjective procedure. It is often not possible for a researcher to explain in a research report all the subtle nuances that lead to one classification rather than another. This means that the probability of replication is low.
3. When relying on content analysis, it is important to describe procedures in detail, so that intersubjectivity can be used to assess reliability. This assessment of reliability does not determine the validity of a procedure, but is a prerequisite for validity.