

The
Content
Analysis
Guidebook

Kimberly A. Neuendorf
Cleveland State University

PUB 2002



Sage Publications

International Educational and Professional Publisher
Thousand Oaks ■ London ■ New Delhi

Defining Content Analysis

Content analysis is perhaps the fastest-growing technique in quantitative research. Computer advances have made the organized study of messages quicker and easier . . . but not always better. This book explores the current options in the analysis of the content of messages.

Content analysis may be briefly defined as the systematic, objective, quantitative analysis of message characteristics. It includes the careful examination of human interactions; the analysis of character portrayals in TV commercials, films, and novels; the computer-driven investigation of word usage in news releases and political speeches; and so much more. Content analysis is applicable to many areas of inquiry, with examples ranging from the analysis of naturally occurring language (Markel, 1998) to the study of newspaper coverage of the Greenhouse Effect (Miller, Boone, & Fowler, 1992) and from a description of how the two genders are shown on TV (Greenberg, 1980) to an investigation of the approach strategies used in personal ads (Kolt, 1996). Perhaps, one of the more surprising applications is Johnson's (1987) analysis of Porky Pig's vocalics from a clinical speech therapy standpoint. He examined 37 cartoons, finding that the per-cartoon stuttering ranged from 11.6% to 51.4% of words uttered, and certain behaviors were associated with the stuttering (e.g., eye blinks, grimaces). If you are unfamiliar with the range of content analysis applications, Chapter 9 presents an overview of the major areas of study—the main “contexts” of content analysis research.

The various techniques that make up the methodology of content analysis have been growing in usage and variety. In the field of mass communication research, content analysis has been the fastest-growing technique over the past 20 years or so (Riffe & Freitag, 1997; Yale & Gilly, 1988). Perhaps, the greatest explosion in analysis capability has been the rapid advancement in computer text content analysis software, with a corresponding proliferation of online archives and databases (Evans, 1996). There has never been such ready

access to archived textual messages, and it has never been easier to perform at least basic analyses with computer-provided speed and precision. This book will explore the expansion and variety of the techniques of content analysis.

In this chapter, we will follow the development of a full definition of content analysis—how one attempts to ensure objectivity, how the scientific method provides a means of achieving systematic study, and how the various scientific criteria (e.g., reliability, validity) are met. Furthermore, standards are established, extending the expectations of students who may hold a prior notion of content analysis as necessarily “easy.”

Is Content Analysis “Easy”?

Is It Something That Anyone Can Do?

There seem to be certain common misconceptions about the method of content analysis: Conducting a content analysis is substantially easier than conducting other types of research, content analysis is anything a scholar or student says it is, and anyone can do it without much training or forethought. It’s also widely assumed that there is little reason to use content analysis for commercial or nonacademic research. Unfortunately, these stereotypes have been reinforced by academic journals that too often fail to hold content analyses to the same standards of methodological rigor as they do other social and behavioral science methods, such as surveys, experiments, and participant observation studies. Based on more than 20 years of involvement in over 100 content analyses, I would like to dispel common myths about this method before providing a full working definition.

Myth 1: Content analysis is easy.

Truth: Content analysis is as easy—or as difficult—as the researcher determines it to be. It is not necessarily easier than conducting a survey, experiment, or other type of study.

Although content analysis must conform to the rules of good science, each researcher makes decisions as to the scope and complexity of the content-analytic study. An example of a very limited—and quite easy—content analysis is shown in the summary graph in Figure 1.1, indicating how many prime-time network TV shows have dealt with medical issues over a period of 38 years. The unit of the analysis is the individual medically oriented TV program, with three simple variables measured: (a) length of show in minutes, (b) whether the show is a drama or a comedy, and (c) the year(s) the program was aired. The raw data analyzed were listings in a readily accessible source that catalogs all TV shows on the major networks since 1948 (Brooks & Marsh,

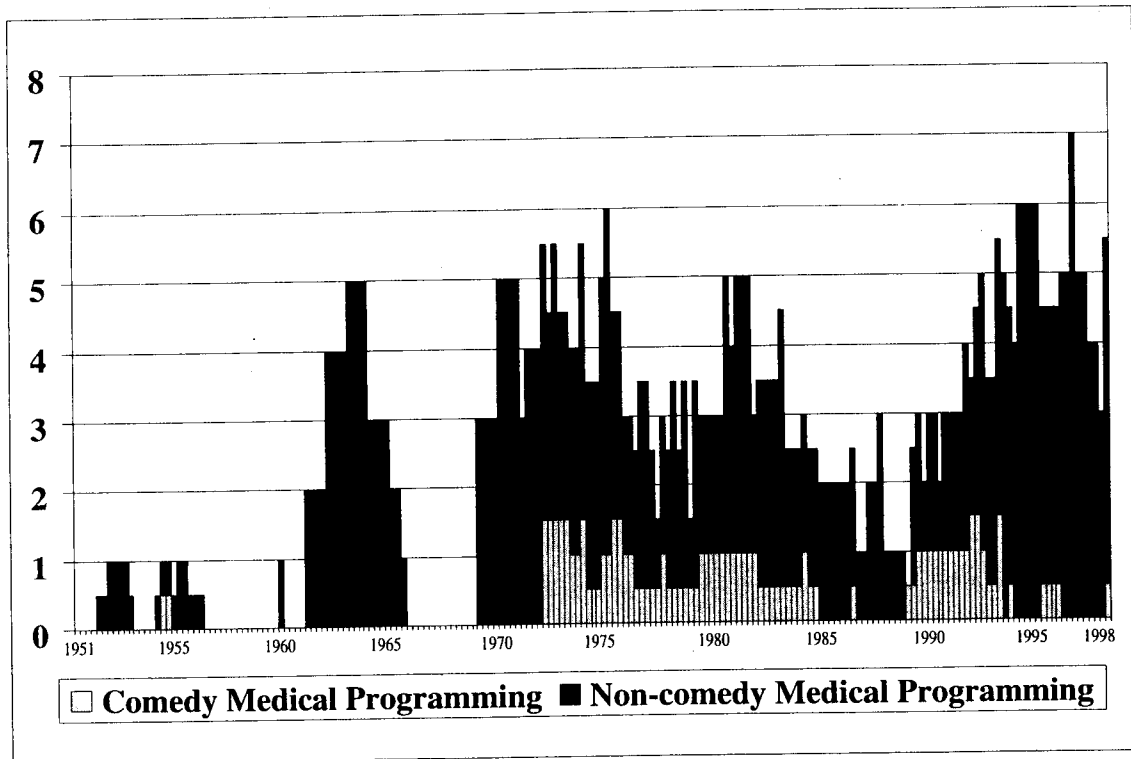


Figure 1.1. Medical Primetime Network Television Programming, 1951 to 1998 (number of hours per week)

1999). Figure 1.1 reports the findings by quarter year in a basic bar graph, indicating weekly total hours of prime-time network TV medical programming. By any assessment, this analysis would be considered easy. Correspondingly, its findings are limited in breadth and applicability. The interpretations we can make from the figure are basic: Over a 40-year period, medical shows have filled only a small portion of the prime-time period, averaging only about 4 hours per week. This has varied little over the period of study.

To make more of the findings, we must dig into the data further and examine the nature of the programs represented in the bar graph. Then, we may identify essentially two eras of TV health-related shows—the 1960s world of physician-as-God medical melodramas (e.g., *Ben Casey*, *Dr. Kildare*) and the 1970s-1990s era of the very human medical professional (e.g., *St. Elsewhere*, *ER*). Comedic medical shows have been rare, with the most successful and enduring among them being *M*A*S*H*. The 1990s included a potpourri of novel medical genres, ranging from documentary-form shows, such as *Rescue 911*, to historical dramas, such as *Dr. Quinn, Medicine Woman*, to science fiction (e.g., *Mercy Point*). Notice that these more interesting findings go beyond the content analysis itself and rely on qualitative analyses. The very simple content analysis has limited utility.

Near the tougher end of the easy-to-difficult continuum might be an ambitious master's thesis (Smith, 1999) that examined the gender role portrayals

of women in popular films from the 1930s, 1940s, and 1990s. The sampling was extremely problematic, given that no valid lists (i.e., sampling frames) of top box office hits are available for years prior to 1939. For many years after that date, all that are available are lists of the top *five* films. The researcher made the analysis even more difficult by deciding to measure 18 variables for each film and 97 variables for each primary or secondary character in each film. Some of the variables were untried in content analysis. For example, psychologist Eysenck's (1990) measures of extraversion (e.g., sociable, assertive, sensation-seeking), typically measured on individuals by self-report questionnaire, were applied to film characters, with not completely successful results. One hypothesis, that female portrayals will become less stereotypic over time, resulted in the measurement and analysis of 27 different dependent variables. With four active coders, the study took 6 months to complete; it was one of the more difficult master's theses among its contemporaries and much more difficult than many surveys and experiments.

The multifaceted results reflected the complexity and breadth of the study. The results included such wide-ranging points as (a) across the decades (1930s, 1940s, 1990s), there were several significant trends indicating a decrease in stereotypical portrayals of women in films; (b) average body shape for women varied across the decades at a near-significant level, indicating a trend toward a thinner body shape; (c) screen women who exhibited more traditional sex-role stereotyping experienced more negative life events; (d) female characters who exhibited more male sex-role traits and experienced negative life events tended to appear in films that were more successful at the box office; and (e) screen women were portrayed somewhat more traditionally in films with greater female creative control (i.e., in direction, writing, producing, or editing) (Smith, 1999).

Myth 2: The term *content analysis* applies to *all* examinations of message content.

Truth: The term does not apply to every analysis of message content, only those that meet a rigorous definition. Clearly, calling an investigation a content analysis does *not* make it so.

There are many forms of analysis—from frivolous to seminal—that may be applied to the human production of messages. Content analysis is only one type, a technique presented by this book as systematic and quantitative. Even in the scholarly literature, some confusion exists as to what may be called a content analysis. On a number of occasions, the term has been applied erroneously (e.g., Council on Interracial Books for Children, 1977; DeJong & Atkin, 1995; Goble, 1997; Hicks, 1992; Thompson, 1996), and at times, studies that warrant the term do not use it (e.g., Bales, 1950; Fairhurst, Rogers, & Sarr, 1987; Thorson, 1989).

A complete reader should analyze (1996) type qualitative of some o sages. Ba (2000), th

Rhetoric

For th and imag character form, me so much detailed i compete pal origin bear of i widely ap forms of

Narrat

This tion foc tions, ar character compos tent rea of this t 1968), dispatch initial si rative (

Discou

Th descrip words ' technic

A complete review of all the types of message analysis that compete with or complement content analysis is beyond the scope of this volume. But the reader should become aware of some of the main options for more qualitative analyses of messages (Lindlof, 1995). One good starting point is Hijmans's (1996) typology of qualitative content analyses applied to media content (according to the definitions presented in this book, we would not include these qualitative procedures as content analysis). She presents accurate descriptions of some of the main qualitative analytic methods that may be applied to messages. Based on descriptions by Hijmans (pp. 103-104) and by Gunter (2000), they are as follows:

Rhetorical Analysis

For this historically revered technique, properties of the text (both words and images) are crucial. The analyst engages in a reconstruction of manifest characteristics of text or image or both, such as the message's construction, form, metaphors, argumentation structure, and choices. The emphasis is not so much on *what* the message says as on *how* the message is presented. There is detailed reading of fragments. There is an assumption that the researcher is a competent rhetorician. This technique has a *very* long history, with its principal origins in Greek philosophy (Aristotle, 1991), and is the legitimate forebear of many of today's academic disciplines. Rhetorical analysis has been widely applied to news content, political speech, advertising, and many other forms of communication (McCroskey, 1993).

Narrative Analysis

This technique involves a description of formal narrative structure: Attention focuses on characters—their difficulties, choices, conflicts, complications, and developments. The analyst is interested not in the text as such but in characters as carriers of the story. The analysis involves reconstruction of the composition of the narrative. The assumption is that the researcher is a competent reader of narratives. One of the most complex and interesting applications of this technique is Propp's exhaustive analysis of Russian fairy tales (Propp, 1968), which establishes common character roles (e.g., hero, helper, villain, dispatcher), an identifiable linear sequence of elements in the narrative (e.g., initial situation, absention, interdiction), and particular functions in the narrative (e.g., disguise, pursuit, transfiguration, punishment).

Discourse Analysis

This process engages in characteristics of manifest language and word use, description of topics in media texts, through consistency and connection of words to theme analysis of content and the establishment of central terms. The technique aims at typifying media representations (e.g., communicator mo-

tives, ideology). The focus is on the researcher as competent language user. Gunter (2000) identifies van Dijk's *Racism and the Press*, published in 1991, as a clear example of a large-scale discourse analysis. According to Gunter, van Dijk analyzes the "semantic macrostructures," or the overall characteristics of meanings, with regard to ethnic minorities in the news media (p. 88), concluding that minority groups are depicted as problematic.

Discourse analysis has been a popular method for analyzing public communication, with analyses ranging from the macroscopic to the very microscopic. Duncan (1996) examined the 1992 New Zealand National Kindergarten Teachers' Collective Employment Contract Negotiations and identified two discourses—"Children First" and "For the Sake of the Children." Both discourses were evident in arguments used by each side in the labor negotiations, in arguments *for* teacher pay and benefits by the teachers' representatives, and in arguments *against* such expenditures by employers and government reps. Duncan's article presents numerous direct quotes from the negotiations to support her point of view. Typical of this method, she points out that her analysis "is *one* reading of the texts, and that there will be numerous other readings possible" (p. 161).

Structuralist or Semiotic Analysis

The focus here is on deeper meanings of messages. The technique aims at deep structures, latent meanings, and the signifying process through signs, codes, and binary oppositions. Interpretations are theoretically informed, and assertions are made on central themes in culture and society. Rhetorical or narrative analysis can be preliminary to this process. The assumption is that the researcher is a competent member of the culture. (See also Eco, 1976.)

Semiotics has been a valuable technique for examining cultural artifacts. Christian Metz's (1974) classic text, *A Semiotics of the Cinema*, applies the wide range of semiotic techniques to the specific medium of narrative film. He provides a "syntagmatic" analysis of the French film, *Adieu Philippine*, indicating the *structure* of the film in shots, scenes, sequences, and the like. He also offers a detailed semiotic analysis of the self-reflexive "mirror construction" of Federico Fellini's semiautobiographical film, *8-1/2*.

Interpretative Analysis

The focus of this technique is on the formation of theory from the observation of messages and the coding of those messages. With its roots in social scientific inquiry, it involves theoretical sampling; analytical categories; cumulative, comparative analysis; and the formulation of types or conceptual categories. The methodology is clearly spelled out, but it differs from scientific inquiry in its wholly qualitative nature and its cumulative process, whereby the analyst is in a constant state of discovery and revision. The researcher is assumed to be a competent observer.

Many of the systems of analysis developed by these methods are empirical and detailed and in fact are more precise and challenging than most content analyses (e.g., Propp, 1968). With only minor adjustment, many are appropriate for use in content analysis as well (e.g., Berger, 1982, 1991).

In addition to these qualitative message analysis types reviewed by Hijmans (1996), several others deserve mention.

Conversation Analysis

Conversation analysis is a technique for analyzing naturally occurring conversations, used by social scientists in the disciplines of psychology, communication, and sociology (Sudnow, 1972). The procedure has been described as a “rigorously empirical approach which avoids premature theory construction and employs inductive methods . . . to tease out and describe the way in which ordinary speakers use and rely on conversational skills and strategies” (Kottler & Swartz, 1993, pp. 103-104). Most typically, it relies on transcribed conversations. The technique generally falls within the rubric of ethnomethodology, scholarly study in which the precise and appropriate methods used emerge from within the process of study, with the clearly subjective involvement of the investigator. Examples of its applications have included an analysis of doctor-patient interaction (Manning & Ray, 2000) and an in-depth analysis of a notorious interview of Vice President George Bush by television reporter Dan Rather as they jockeyed for position in order to control the flow of a “turbulent” interview (Nofsinger, 1988/1989).

Critical Analysis

Critical analysis, often conducted in a tradition of cultural studies, has been a widely used method for the analysis of media messages (Newcomb, 1987). The area of film studies provides a good example of a fully developed, theoretically sound literature that primarily uses the tools of critical analysis (e.g., Lyman, 1997). For example, Strong’s (1996) essay about how Native Americans are “imaged” in two mid-1990s media forms—Disney Studio’s *Pocahontas* and Paramount’s *The Indian in the Cupboard*—is influenced heavily by her own roles as mother, musician, American raised during a period when “playing Indian” was a childhood rite of passage, and anthropologist long interested in White America’s representations of Native Americans. She acknowledges these various roles and perspectives, provides precise details to back her assertions (including many lines and song lyrics from the movies), and gives summative statements that bring the details into line with cultural frameworks. For example, she concludes that “Disney has created a New Age Pocahontas to embody our millennial dreams for wholeness and harmony, while banishing our nightmares of savagery without and emptiness within” (p. 416).

Normative Analysis

Some analyses are explicitly normative or proscriptive. For example, a guide to *Stereotypes, Distortions and Omissions in U.S. History Textbooks: A Content Analysis Instrument for Detecting Racism and Sexism* (Council on Interracial Books for Children, 1977), compiled by 32 educators and consultants, provides checklists for history textbook coverage of African Americans, Asian Americans, Chicanos, Native Americans, Puerto Ricans, and women. For each group, an instrument is presented with criteria for parents and teachers to use when examining children's history texts. For instance, in the Native American checklist, the following criteria are included: "The myth of 'discovery' is blatantly Eurocentric," "War and violence were not characteristic of Native nations," "The Citizenship Act of 1924 was not a benevolent action," and "The BIA [Bureau of Indian Affairs] is a corrupt and inefficient bureaucracy controlling the affairs of one million people" (pp. 84-85). The guide is certainly well intended and a powerful tool for social change. It does not, however, fit most definitions of content analysis.

Similarly, in their article, "Evaluation Criteria and Indicators of Quality for Internet Resources," Wilkinson, Bennet, and Oliver (1997) offer a list of 125 questions to ask about a Web site. Their goal is to pinpoint characteristics that indicate accuracy of information, ease of use, and aesthetic qualities of Internet material. The work is a normative prescription for a "good" Web site. Although they call their proposal a content analysis, it does not meet the definition given in this book.

In another case of normative recommendations for message content, Legg (1996) proposes that commercial films are an important venue for the exploration of religion in American culture, and she provides tips to religious educators for using movies in teaching. She contends that "in forms like contemporary films we can see the very pertinent questions with which our culture is really wrestling" (p. 401) and urges religious educators not to limit their use of film to explicitly religious films, such as *The Ten Commandments* or *Agnes of God*. Equally useful might be explorations of manifestations of good and evil in *Batman* or a discussion of dimensions of friendship, aging, Southern ethos, prejudice, and family in *Driving Miss Daisy* (p. 403). Such detailed analyses have obvious utility; however, this process does not attempt to achieve objectivity, as does a content analysis.

Myth 3: Anyone can do content analysis; it doesn't take any special preparation.

Truth: Indeed, anyone can do it . . . but only with training and with substantial planning.

While the person who designs a content analysis must have some special knowledge and preparation, a central notion in the methodology of content analysis is that *all* people are potentially valid "human coders" (i.e., individuals who make judgments about variables as applied to each message unit). The

coding scheme must be so objective and so reliable that, once they are trained, individuals from varied backgrounds and orientations will generally agree in its application.

Clearly, however, each coder must be proficient in the language(s) of the message pool. This may require some special training for coders. To analyze natural speech, coders may actually need to learn another language or be trained in the nuances of a given dialect. Before coding television or film content, coders may have to learn about production techniques and other aspects of visual communication. To code print advertising, coders may need to learn a bit about graphic design. All this is in addition to training *with the coding scheme*, which is a necessary step for all coders.

For analyses that do not use human coders (i.e., those that use computer coding), the burden rests squarely on the researcher to establish complete and carefully researched dictionaries or other protocols. Because the step of making sure coders can understand and reliably apply a scheme is missing, the researcher needs to execute additional checks. Chapter 6 presents some notions on how this might be done.

Myth 4: Content analysis is for academic use only.

Truth: Not.

The vast majority of content analyses have been conducted by academics for scholarly purposes. However, there has been growing interest among commercial researchers and communication practitioners in particular applications of content analysis. A law firm hired a respected senior professor to conduct content analyses of news coverage of their high-profile clients, to be used as evidence in conjunction with a change-of-venue motion (i.e., excessive and negative coverage may warrant moving a court case to another city in order to obtain a fair trial; McCarty, 2001). In response to criticisms, a Southern daily newspaper hired a journalism scholar to systematically document their coverage of the local African American community (Riffe, Lacy, & Fico, 1998). The marketing research unit of a large-city newspaper has begun the process of systematically comparing its own coverage of regional issues with that provided by local television news. Organizational communication consultants sometimes include a content analysis of recorded messages (e.g., e-mail, memos) in their audit of the communication flow in the organization. And the clinical diagnostic tools of criteria-based content analysis have been used in nonacademic settings by psychologists and legal professionals.

A Six-Part Definition of Content Analysis

This book assumes that content analysis is conducted within the scientific method but with certain additional characteristics that place it in a unique position as the primary message-centered methodology.

Box 1.1 Defining Content Analysis

Some of the main players in the development of quantitative message analysis present their points of view:

Berelson (1952, p. 18): Content analysis is a research technique for the objective, systematic, and quantitative description of the manifest content of communication.

Stone, Dunphy, Smith, & Ogilvie (1966, p. 5, with credit given to Dr. Ole Holsti): Content analysis is any research technique for making inferences by systematically and objectively identifying specified characteristics within text.

Carney (1971, p. 52): The general purpose technique for posing questions to a "communication" in order to get findings which can be substantiated. . . . [T]he "communication" can be anything: A novel, some paintings, a movie, or a musical score—the technique is applicable to all alike and *not* only to analysis of literary materials.

Krippendorff (1980, p. 21): Content analysis is a research technique for making replicable and valid inferences from data to their context.

Weber (1990, p. 9): Content analysis is a research method that uses a set of procedures to make valid inferences from text.

Berger (1991, p. 25): Content analysis . . . is a research technique that is based on measuring the amount of something (violence, negative portrayals of women, or whatever) in a representative sampling of some mass-mediated popular art form.

Riffe, Lacy, & Fico (1998, p. 20): Quantitative content analysis is the systematic and replicable examination of symbols of communication, which have been assigned numeric values according to valid measurement rules, and the analysis of relationships involving those values using statistical methods, in order to describe the communication, draw inferences about its meaning, or infer from the communication to its context, both of production and consumption.

This book: Content analysis is a summarizing, quantitative analysis of messages that relies on the scientific method (including attention to objectivity-intersubjectivity, a priori design, reliability, validity, generalizability, replicability, and hypothesis testing) and is not limited as to the types of variables that may be measured or the context in which the messages are created or presented.

Content analysis is a summarizing, quantitative analysis of messages that relies on the scientific method (including attention to objectivity-intersubjectivity, a priori design, reliability, validity, generalizability, replicability, and hypothesis testing) and is not limited as to the types of variables that may be measured or the context in which the messages are created or presented.

Box 1.1 presents some alternative definitions of content analysis for comparison's sake. More details on this book's definition are presented in the discussion that follows.

1. Content Analysis as Relying on the Scientific Method

Perhaps, the most distinctive characteristic that differentiates content analysis from other, more qualitative or interpretive message analyses is the attempt to meet the standards of the scientific method (Bird, 1998; Klee, 1997);