CHAPTER 3

Beyond Description

An Integrative Model of Content Analysis

As noted in Chapter 1, this book takes the view that content analysis is (or should be) a research technique that conforms to the rules of science. Most closely related to the technique of survey research, it uses messages rather than human beings as its level of analysis. Issues that apply include those of validity (internal and external), reliability, sample representativeness, the principle of maximum information (Woelfel & Fink, 1980), and objectivity (or intersubjectivity). Before proceeding on a discussion of exactly how content analysis may be conducted to achieve these standards, a basic background on the ground rules and terminology of the scientific method is in order.

The Language of the Scientific Method

Whether explicitly stated or not, a primary goal of most scientific studies is to identify ***causal relationships***. That is, we hope to discover at least one causal agent (X) that leads to at least one outcome (Y). Establishing a cause-and-effect relationship is rarely (many would say never) achieved in social and behavioral scientific study; it is almost impossible to meet all three criteria for causality: (a) a relationship, (b) time ordering (such that X precedes Y in time), and (c) the elimination of all alternative explanations. The second criterion—time ordering—requires either a study that has two or more measurement points over time (a ***longitudinal*** study) or an experiment. The third criterion—accounting for all alternative explanations—is generally impossible to fully achieve. However, the task is to do the most complete job possible, identifying and measuring as many ***control*** variables as possible.[1]

Given that the discovery of true causality is essentially an unattainable goal, we do not refer to "cause" and "effect." Rather, we refer to each "presumed cause" as an ***independent variable*** and each "presumed effect" as a ***dependent variable***. A ***variable*** is a definable and measurable construct that varies, that is, it holds different values for different individual cases or units. For example, we may predict that gender is related to interrupting behavior, such that males will interrupt conversations more often than will females. Each unit (person, in this case) will hold a certain value on the independent variable (male or female) and a certain value on the dependent variable (e.g., 4 times in 5 minutes, 12 times in 5 minutes). These values must vary across units or there exists no variable for study; for instance, if all persons in the study are male, no comparison may be made with females, and therefore "gender" does not exist in the study as a variable; we could not assess the impact of gender on interrupting behavior. In this example, gender is the independent variable and interrupting behavior the dependent variable. In a ***hypothesis,*** we predict that one's gender affects one's interrupting behavior; the converse is clearly impossible in this case.

The main empirical, quantitative methods available to the social or behavioral scientist to investigate ***hypotheses*** or ***research questions*** about possible relationships between independent and dependent variables are the ***experiment*** and the ***survey***. An experiment is an investigation in which at least one independent variable is manipulated or controlled. A survey is a study in which an attempt is made to measure all variables—independent and dependent—as they naturally occur. Note the simplicity of these definitions. Despite common expectations, an experiment does ***not*** have to take place in a laboratory setting (although many do). And a survey does ***not*** have to consist of a questionnaire (although most do). An experiment could be conducted in a "real" setting, such as a workplace. For example, a researcher might randomly assign employees in an organization to types of working conditions (e.g., face-to-face work teams vs. "virtual," online work teams), and then outcome variables could be measured.

Most surveys do involve the use of a ***questionnaire,*** that is, a set of questions that are presented to a respondent either as a self-administered paper-and-pencil booklet, as an online set of questions, or as an interview. However, many experiments also use a questionnaire, especially to measure dependent variables and control variables. And a survey that does ***not*** use a questionnaire is quite possible—it might involve direct observation of behavior: for instance, observing and tabulating a child's play behavior. Although the "self-report" nature of the study has been eliminated, it's important to note that the observation process also relies on subjective human reportage—in this case, a human observer-coder of others' behavior. As we shall see, this involvement of humans in the measurement process is of great concern in content analysis.

The relative advantages and disadvantages of the two research approaches —experiment and survey—are clear. An experiment generally enjoys a high degree of control and certainty about the validity of the independent variable(s) but is often artificial in its execution (i.e., higher on validity and lower on generalizability, which is sometimes called ***external validity***). A survey is more true to life and tends to be more generalizable so long as a ***random sample*** is employed, but its measures are more suspect, especially when they rely on self-report questionnaire responses (i.e., higher on generalizability or external validity and lower on validity). Most scholars agree that the "best" approach is one of ***triangulation,*** that is, testing for a hypothesized relationship among variables with a variety of methods—experiments, surveys, and other, more qualitative methods. The various methods' strengths and weaknesses tend to balance out, and if all the various methods reveal similar findings, the support for the hypothesis is particularly strong.

Content analysis as a research method is consistent with the goals and standards of ***survey*** research. In a content analysis, an attempt is made to measure all variables as they naturally or normally occur. No manipulation of independent variables is attempted. Some type of random sampling of the units of data collection is typical, making the findings generalizable to a larger grouping or ***population*** of messages. Note that the units of data collection are simply different from those of the typical survey (i.e., messages rather than persons). And the questionable validity of the measures in a survey also applies to the content analysis. Just as the self-report nature of most surveys calls into question the objectivity and validity of their measures, so, too, the involvement of human decision makers in the content analysis process calls into question the validity of the coding or dictionary construction. In short, the content analysis enjoys the typical advantages of survey research and usually suffers its drawbacks as well.

How Content Analysis Is Done: A Flowchart for the Typical Process of Content-Analytic Research

As a form of scientific endeavor, content analysis ought to be conducted in line with procedures appropriate to good science. Box 3.1 contains a flowchart of the typical process of content analysis, with nine steps outlined. The model follows the common steps for research in the scientific method, applying appropriate terminology for content analysis whenever needed. Subsequent chapters will explore the steps: Steps 1 through 3 are treated in Chapter 5; Step 4, Coding, is included in Chapter 6; Step 5, Sampling, is the subject of Chapter 4; Steps 6 through 8 are addressed in Chapter 7; And Step 9, Tabulation and Reporting, is dealt with in Chapter 8. An important distinction apparent in the flowchart is the splitting off of human coding from computer coding at two junctures. At this point, it's important to understand the differences between the two.

Box 3.1 A Flowchart for the Typical Process of Content Analysis Research

1. *Theory and rationale:* ***What*** content will be examined, and ***why?*** Are there certain ***theories*** or perspectives that indicate that this particular message content is important to study? Library work is needed here to conduct a good literature review. Will you be using an integrative model, linking content analysis with other data to show relationships with source or receiver characteristics? Do you have ***research questions?*** ***Hypotheses?***

↓

2. *Conceptualizations:* What ***variables*** will be used in the study, and how do you define them ***conceptually*** (i.e., with dictionary-type definitions)? Remember, you are the boss! There are many ways to define a given construct, and there is no one right way. You may want to screen some examples of the content you're going to analyze, to make sure you've covered everything you want.

↓

3. *Operationalizations (measures):* Your measures should match your conceptualizations (this is called ***internal validity***). What ***unit of data collection*** will you use? You may have more than one unit (e.g., a by-utterance coding scheme and a by-speaker coding scheme). Are the variables measured well (i.e., at a high ***level of measurement,*** with categories that are ***exhaustive and mutually exclusive***)? An ***a priori*** coding scheme describing all measures must be created. Both face validity and content validity may also be assessed at this point.

Human Coding ↙ ↘ Computer Coding

4a. *Coding schemes:* You need to create the following materials:

a. *Codebook* (with all variable measures ***fully*** explained)

b. *Coding form*

↓

Human Coding

4b. *Coding schemes:* With computer text content analysis, you still need a codebook of sorts—a full explanation of your ***dictionaries*** and method of applying them. You may use standard dictionaries (e.g., those in Hart's program, *Diction*) or originally created dictionaries. When creating custom dictionaries, be sure to first generate a frequencies list from your text sample and examine for key words and phrases.

↓

Computer Coding

5. *Sampling:* Is a census of the content possible? (If yes, go to #6.) How will you ***randomly sample*** a subset of the content? This could be by time period, by issue, by page, by channel, and so forth.

↓

6. *Training and pilot reliability:* During a training session in which coders work together, find out whether they can agree on the coding of variables. Then, in an independent coding test, note the ***reliability*** on each variable. At each stage, ***revise*** the codebook or coding form as needed.

↓

7a. *Coding:* Use at least two coders, to establish intercoder reliability. Coding should be done independently, with at least 10% overlap for the reliability test.

↓

Human Coding

7b. *Coding:* Apply dictionaries to the sample text to generate per-unit (e.g., per-news-story) frequencies for each dictionary. Do some spot checking for validation.

↓

Computer Coding

8. *Final reliability:* Calculate a reliability figure (percent agreement, Scott's *pi*, Spearman's *rho*, or Pearson's *r*, for example) for each variable.

↓

9. *Tabulation and reporting:* See various examples of content analysis results to see the ways in which results can be reported. Figures and statistics may be reported one variable at a time (univariate), or variables may be cross-tabulated in different ways (bivariate and multivariate techniques). Over-time trends are also a common reporting method. In the long run, relationships between content analysis variables and other measures may establish criterion and construct validity.

Human Coding Versus Computer Coding

Human coding involves the use of people as coders, with each using a standard codebook and coding form to read, view, or otherwise decode the target content and record his or her objective and careful observations on preestablished variables. Computer coding involves the automated tabulation of variables for target content that has been prepared for the computer. Typically, computer coding means having software analyze a set of text, counting key words, phrases, or other text-only markers. The term *CATA* has been adopted to designate the popular use of "computer-aided text analysis."

Until recently, some content analysts referred to automated coding as *machine coding,* leaving open the possibility of noncomputer automated coding. Are there currently any machines other than computers capable of conducting some type of automated content analysis? Not really. However, as will be described in Chapters 4 and 6, there are a number of video and audio technologies that may assist in the coding of visual and auditory content. In all likelihood, as their potentials for fully automated coding develop, all these technologies will be firmly linked with (controlled by, or merged with) computers. However, theoretically, machine coding could be conducted by a device other than a computer.

Chapter 6 presents some examples of codebooks and coding forms for human coding and dictionaries for computer coding via computer text content analysis. More examples are available at *The Content Analysis Guidebook Online.*

Approaches to Content Analysis

This book presents the view that content analysis is best approached as one tool for testing relationships within a basic communication model. The classic Shannon-Weaver (Shannon & Weaver, 1998) model provides the raw framework of *source, message, channel, and receiver.* Based on this, Berelson (1952) proposed five purposes for content analysis: (a) to describe substance characteristics of message content (essentially what are described in Chapter 1 as content characteristics), (b) to describe form characteristics of message content, (c) to make inferences to producers of content, (d) to make inferences to audiences of content, and (e) to determine the effects of content on the audience.

The view presented in this book does not accept the notion that it is appropriate to make conclusions about source or receiver on the basis of an analysis of message content alone. Carney (1971) expresses the view that there are three uses of content analysis: Descriptive, hypothesis testing, and *facilitating* inference. This book's presentation is more consistent with Carney's view, extending it somewhat to include the consideration of *extramessage* variables, that is, variables measured on source or receiver units.

This chapter will attempt to develop an integrative model of content analysis, which recognizes that whereas content analysis itself can only *describe* message characteristics or identify *relationships* among message characteristics, its methods are integral to a full understanding of human behavior and hence essential to social and behavioral science. When combined with results from other studies that use persons as units of inquiry (Shoemaker & Reese, 1990), content analysis can provide important missing links in multistep models of human communication behavior and of audience responses to mediated communication. Thus, whereas Berelson, for example, says that we may infer source characteristics or intent from examining message content, this book argues instead for the integration of content analytic studies with surveys of sources that give concrete evidence of source characteristics or intent. The goals of science are typically presented as *description, prediction, control,* and *explanation.* Only with an integrated approach to data collection can applications of content analysis aspire to the highest goal: explanation.

To date, the common approaches to content analysis may be categorized as descriptive, inferential, psychometric, and predictive.

Descriptive Content Analysis

Many content analyses describe a particular message pool in almost archival fashion. Researchers working in this vein are careful to limit their conclusions to the content being studied, although they may clearly be motivated by a desire to infer characteristics to the source(s) or to anticipate outcomes of the messages. These analyses are attractive in their clarity and parsimony. But they sometimes seem to be targets for those who question the scientific importance or sophistication of content analysis as a method.

An example of descriptive content analysis would be the ongoing research tracking sexual content in U.S. television programming (Kunkel, Cope-Farrar, Biely, Farinola, & Donnerstein, 2001). Whereas its origins are clearly in a concern over the effects of sexual content on viewers, the project never overstates its conclusions—they are purely content based. Key findings include an identified rise in sexual content over a 2-year comparison period, notably in shows featuring teenagers.

Also, a team at Temple University is currently involved in a long-range project to describe formal features of TV as they are presently employed (Lombard, Campanella, Linder, & Snyder, 1996). Motivated by the growing body of evidence concerning the physiological and psychological processing and impact of structural features of television (e.g., pace of editing, camera angles, sound effects, use of text and graphics), their work begins to document the state of the art of television production. Their descriptive goal is clear.

Another example of a purely descriptive content analysis is the Screen-Actors-Guild-sponsored analysis of prime-time television programming, the African American Television Report (Robb, 2000). A team led by Darnell Hunt at the University of Southern California examined a sample of 384 epi-

sodes of 87 prime-time series on the six broadcast networks, identifying a total of 6,199 characters. The study found that although African Americans composed 12.2% of the U.S. population at the time of the data collection, they accounted for about 16% of the characters. This "overrepresentation" was more marked in situation comedies, with more than half (50.5%) of all Black characters appearing in sitcoms.

It needs to be understood that *descriptive* does not always mean *univariate,* that is, describing results one variable at a time. There might be—and often should be—a predicted relationship among variables measured in the content analysis. A good example is Dixon and Linz's (2000) study of television news coverage of lawbreakers. They found a significant relationship between the race of the perpetrator and whether the crime reported was a felony. Thus, the bivariate (two-variable) relationship is

Race of perpetrator → Type of crime (felony or nonfelony)

Inferential Content Analysis

The view presented in this book does not endorse most explicit inferences made strictly from content analysis results, a view consistent with early admonitions by Janis (1949). Counter to this view, Berelson's (1952) 50-year-old encouragement continues to be invoked in cases where researchers wish to make conclusions about sources or receivers solely from content-analytic studies. Yet such unbacked inferences are inconsistent with the tenets of the philosophy of science—it is important to note that they are not empirically based.

It seems that interpersonal communication-type content analyses (especially those with known receiver[s]) tend to try to infer to the source, whereas mass communication-type studies (with undifferentiated receivers) tend to attempt to infer to receivers or receiver effects or both, although this observation has not been backed by data (e.g., a content analysis of content analyses). Clearly, however, there is great interest in going beyond description of messages. As we will see, there are alternatives to nonempirical inference.

Psychometric Content Analysis

The type of content analysis that seems to have experienced the greatest growth in recent years within the discipline of psychology is that of psychometric content analysis. This method seeks (a) to provide a clinical diagnosis for an individual through analysis of messages generated by that individual or (b) to measure a psychological trait or state through message analysis.

This particular application of content analysis might be seen as going beyond simple inference in that the measures are validated against external standards. Applying the notion of criterion validity as articulated by Carmines and Zeller (1979), the technique involves a careful process of validation, in which

content analysis is linked with other time-honored diagnostic methods, such as observations of the subject's behavior (the "criterion"). Over a series of investigations, the content analysis dictionaries (sets of words, phrases, terms, and parts of speech that are counted up in a sample of the subject's speech or writing) are refined to improve their correspondence with the older diagnostic and psychographic techniques (Gottschalk, 1995; Smith, 1992). The ultimate goal is, however, to *infer* to a given source without having to apply these other diagnostic tools each and every time. But this is done only after substantial, careful validation with numerous sources.

Predictive Content Analysis

This type of content analysis has as its primary goal the prediction of some outcome or effect of the messages under examination. By measuring key characteristics of messages, the researcher aims to predict receiver or audience responses to the messages. This necessitates the merging of content-analytic methods with other methods that use people as units of data collection and analysis—typically, survey or experimental methods or both.

A good example of this type of study is Naccarato's (Naccarato & Neuendorf, 1998) combined content analysis and audience study that linked key print advertising features to audience recall, readership, and evaluations of ads. Box 3.2 tells the story of the research process, and Box 3.3 carries the process a bit further by applying the knowledge gained from the content analysis to a hypothetical creative process of new ad creation.

In a series of studies linking media presentations of violent acts and aggregate crime and mortality statistics from archival sources, Phillips (1974, 1982, 1983; Phillips & Hensley, 1984; Phillips & Paight, 1987) has established a long and distinctive record of research using simple predictive content analysis.[2] He has examined the incidence of homicides after network news coverage of championship boxing matches, the incidence of suicides after newspaper reports of suicides, and the occurrence of deaths due to car accidents following soap opera suicides. Although Phillips's attempts to draw causal conclusions have come under criticism (Gunter, 2000), his research approach has shown robust, replicable relationships between media reports and depictions of violence and real-life events.

Another type of predictive content analysis that has been gaining popularity is the prediction of public opinion from news coverage of issues (e.g., Salwen, 1986). Through a blending of content analysis and public opinion poll summarization, Hertog and Fan (1995) found that print news coverage of three potential HIV transmission routes (toilets, sneezing, and insects) *preceded* and was *significantly related* to public beliefs about those routes as expressed in polls. Key details of this innovative and sophisticated study are reported in Box 3.4.

Box 3.2 The Practical Prediction of Advertising Readership

After 20 years as an advertising professional, John Naccarato wanted his master's thesis (see Naccarato & Neuendorf, 1998) to merge theory and research with a practical application to his chosen field. In his capacity as a business-to-business ad specialist, he was accustomed to receiving reports from publishers and from other standard readership services regarding the level of readership for the ads he placed in business-to-business publications. Privately, he had always asked what he called the "why" question: *Why* did one ad perform better than another? What was it about a given ad that attracted the reader?

He settled on content analysis as a method of linking the already accessible readership data with ad characteristics. In this way, he would be able to find out if certain ad attributes bore a relationship to readership scores. If so, although causality would not be verifiable, he could at least make *predictions* from ad characteristics. Only a handful of studies had tried to do something along these lines; only a few of these analyzed print advertising, and none had examined the business-to-business context (Chamblee, Gilmore, Thomas, & Soldow, 1993; Donath, 1982; Gagnard & Morris, 1988; Holbrook & Lehmann, 1980; Holman & Hecker, 1983; Stewart & Furse, 1986; Wood, 1989).

Naccarato's needs were concrete—he wanted to find the best combination of ad variables that would predict reader response—but he did not ignore theory and past research in his collection. From persuasion theories, he derived measures of the ad's appeals (e.g., humor, logical argument, fear; Markiewicz, 1974). From earlier content analysis studies, he adapted indicators of form attributes, such as use of color, ad size, and other layout features. From practitioner recommendations found in advertising texts, he pulled variables such as use of case histories, use of spokespersons, and competitive comparisons. And from his own personal experience in advertising, John extracted such notions as the consideration of the role of charts and graphs in the ad layout. At the end of the process of collecting variables, he had a total of 190 variables.

Naccarato's codebook and corresponding coding form were lengthy (both may be found at *The Content Analysis Guidebook Online*). As a result of combining variables and eliminating variables with low reliabilities or lack of variance, the final pool of variables was reduced to 54 form and 21 content variables for inclusion in analyses.

The population of messages was defined as all ads appearing in the trade publication, *Electric Light and Power* (*EL&P*) during a 2-year period. Sampling was done by issue; eight issues were randomly selected, with all ads in each issue included in the analysis (n = 247). All the ads in *EL&P* during this time period had been studied via the publisher's own readership survey, the PennWell Advertising Readership Research Report. This self-report mail survey of subscribers measured audience recall and readership and perceptions of the ad as attractive and informative. The survey sample sizes ranged from 200 to 700, and response rate ranged from 10% to 50%.

With the unit of analysis being the individual ad, data were merged to analyze the relationship between ad characteristics and each of the four audience-centered dependent variables. Stepwise regression analyses were conducted to discover which of the 75 independent variables best constructed a predictive model.

This approach proved to be fruitful. All four regression models were statistically significant. Variances accounted for were as follows: For ad recall, 59%; readership, 12%; informativeness, 18%; attractiveness, 40%. For example, ad recall seemed to be enhanced by use of a tabloid spread, greater use of color, use of copy in the bottom half of the ad, use of large subvisuals, and advertising a service (rather than a product). Recall was lower with ads that were of fractional page or junior page size, that used copy in the right half of the ad, and that used a chart or graph as their major visual (rather than a photo).

John Naccarato's practical interest in predicting audience attraction to business-to-business ads was rewarded with some powerful findings and resulted in a caution against taking practitioner recommendations too seriously. In only a small number of instances did such recommendations match up with the study's findings of what relates to positive reader reactions. For example, books by leading advertising professionals recommend such techniques as the use of a spokesperson, humor, calls to action, and shorter copy. Yet none of these was related to any of the four audience outcomes. On the other hand, copy placement and use of fear appeals were important predictors that practitioners usually ignore.

The Integrative Model of Content Analysis

Expanding on this notion of predictive content analysis, it is proposed that a comprehensive model for the utility of the method of content analysis be constructed.

To date, Shoemaker and Reese (1996) have been the most vocal proponents of integrating studies of media sources, messages, audiences, and media effects on audiences. They have developed a model of research *domains* for typologizing mass media studies. Their individual domains are as follows:

A. Source and system factors affecting media content
B. Media content characteristics as related to audience's use of and evaluation of content
C. Media content characteristics as predictive of media effects on the audience
D. Characteristics of the audience and its environment as related to the audience's use of and evaluation of media content
E. Audiences' use of and evaluation of media content as related to media's effects on the audience

Box 3.3 Creating the "Perfect" Advertisement

Using Content Analysis for Creative Message Construction

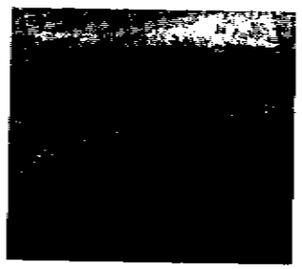
Box 3.2 shows how an integrative content analysis model can produce powerful findings with practical significance. By linking message features with receiver response, Naccarato and Neuendorf (1998) discovered specific form and content characteristics of business-to-business advertisements that led to recall, readership, and other indicators of message effectiveness. A logical next step would be to relate these findings back to the source level by constructing an ad that incorporates all of the successful predictors.

Just for fun, a sample ad has been created that does just that. The ad, shown below, is for the fictional product, *SharkArrest*. It incorporates all of the exclusively positive, significant predictors of business-to-business message effectiveness from the Naccarato and Neuendorf study into a single message.

Protect your resort with *SharkArrest*

Rows and rows of razor-sharp teeth. Five-thousand pounds of clamping pressure. These are just a few frightening attributes of sharks that can wreak havoc on a successful resort community. But now there's a solution to shark-related anxiety.

SharkArrest is a new service for the safety-conscious resort owner. The SharkArrest team of experts is specially trained to fend off sharks using advanced fish repellent techniques. The patented methods are scientifically proven and 100% environmentally & shark safe.



For more information and a free trial, call:
1-555-NOSHARKS

SharkArrest
www.sharkarrest.chomp

Form Variables

Headline placement, top: The headline, "Uninvited guests can kill your business . . .," is located at the top of the ad, making it the first distinguishable feature, from a top-down perspective. This placement relates positively to perceived ***informativeness*** of the ad.

Subject apparent in visuals: Sharks, the subject of the ad, is clearly communicated through the shark photograph in the top half of the layout. Making the subject apparent in visuals is a positive predictor of both ***readership*** and ***informativeness***.

Color: The original SharkArrest ad (not reproduced in color here) includes two colorful photos, one of a shark and another of a beach, and also some color text. Color leads to both ***recall*** and perceived ***attractiveness***.

Large size of subvisuals: The photo of the beach at the bottom is larger than a typical subvisual, which positively predicts ***recall***.

Copy placement, bottom: Two paragraphs of copy appear in the bottom half of the SharkArrest ad. This layout predicts both ***recall*** and ***attractiveness***.

Content Variables

Fear appeal: The ad uses a frightening photo of a great white shark and copy describing some threatening attributes of sharks, to strike fear in resort owners, because it makes sharks seem like a danger to both their businesses and their guests. Fear appeals positively predict both ***readership*** and ***attractiveness*** of ads.

Ad type—service: The ad is from a shark protection company that provides a service to business customers by keeping sharks away from their property. Ads for a service are significantly likely to be ***recalled*** by business-to-business ad readers.

These features could all be included in a real-life business-to-business service advertisement, thus making it a perfect ad, in light of the Naccarato and Neuendorf study results. More important, this example shows one of the many creative possibilities opened up by an integrative content analysis approach.

These authors propose combining the five domains to produce a variety of broader domains of research, with domain A-B-C-D-E as the optimal "fully elaborated model of mass communication" (p. 257). Their model clearly distinguishes between message *effects* on audiences and audiences' more active *use* of media messages.

The Shoemaker and Reese (1996) model can productively inform our discussion of integrating content analysis with additional data. Domains A, B, and C clearly address the utility of content analysis data and hold similarities to the integrative model proposed here. But key differences do exist. The Shoemaker and Reese model is in some senses more restrictive (in its particular application to mass communication research) and in other senses more expansive (in its consideration of modes of inquiry other than content analysis). The integrative model developed in this chapter is designed with the simple goal of focusing interest on the role of ***content analysis*** in answering questions via social and behavioral science investigations.

This model uses the basic Shannon-Weaver communication model (Shannon & Weaver, 1998) as a guide. That model, developed by mathematicians at

Box 3.4 Approaching Causality—Does Press Coverage *Cause* Public Opinion?

The prevailing view of news coverage by those who report it is that the news *follows* public opinion, rather than *leads* it. That is, the public agenda is first established and then news reporters simply pick up on evident trends and document them. Many scholars have questioned this point of view, and numerous media effects studies have established the potential of media messages to change the opinions of individuals. Hertog and Fan (1995) took the novel approach of tracking aggregate public opinion in light of overall news coverage on a singular topic. Using techniques originally proposed in Fan's (1988) book on computer text analysis of news coverage, they collected public opinion poll findings and content analyzed news stories on the same topic over the same period of time.

All stories pertaining to supposed AIDS transmission via sneezes, toilets, and insects, from eight news sources (four major U.S. newspapers, three news magazines, and the UPI newswire) were collected for the years 1987 to 1991. The 166 stories were human-coded for coverage of the issue—"pro" (e.g., representing the view that you *can* get AIDS from toilets) or "con" (e.g., representing the view than you *cannot*). The results of 23 National Health Interview Survey polls were available for the same period, all of which included measures of public perception of the likelihood of AIDS transmission in each of the three manners. Data were merged by time period, with each poll representing one data point.

Using Fan's (1988) ideodynamic model, the relative impacts of both pro and con AIDS transmission stories were assessed. For example, for both sneezing and insects, pro stories seemed to carry more weight than con stories, resulting in a significant change in public perception toward the erroneous pro viewpoint.

Most important, Hertog and Fan (1995) used the Granger Causality Test to examine over-time relationships between news coverage of each AIDS transmission type and public opinion. For both sneezing and toilet transmission, they found news content to predict later public opinion. Public opinion did *not* predict subsequent news content. With a relationship *and* time ordering well established, these findings come as close to establishing causality as we have seen in content-analytic research.

Bell Laboratories in 1949, was designed to describe the flow of information in a mediated system and to mathematically model conditions for optimal system operation (e.g., reduce noise). The original model consisted of the identification of the following elements: source, message, transmitter, signal, channel, noise, receiver, and destination. The model was readily adopted by social and behavioral scientists as a descriptor of the human communication process, with source "encoding," receiver "decoding," and "feedback" from receiver to source as key additions.

Despite its explication in a wide variety of studies over a 50-year period, the Shannon-Weaver model (Shannon & Weaver, 1998) has experienced little

adjustment for the changing information environment (e.g., Baran & Davis, 1995; Dominick, 1990; Schramm & Roberts, 1971; Straubhaar & LaRose, 1996). For example, current debate over what constitutes *message* and what constitutes *channel* in considering Internet Web sites (e.g., as when the site is so responsive to the user that the channel "interacts" with the receiver, creating a unique message pool and mode of presentation for each individual user) has not yet resulted in a popular revision of the model (Skalski, 2000). Generally, the Shannon-Weaver model (Shannon & Weaver, 1998) is a paradigmatic framework for most scholars studying communication activity.

The proposed integrative model of content analysis takes off where Berelson (1952) left off. Rather than engaging in overt inference making from content-analytic information alone, the integrative model calls for the collation of content analysis message-level data with other available empirical information regarding source, receiver, channel, or other contextual states. It requires that a content analysis study be examined within the framework of the basic communication model. Although this may seem rather mechanistic, it provides us with a clear picture of what components contribute to our understanding of the messages of interest, as well as the nature of the links between message variables and extramessage variables.

Evaluation With the Integrative Model of Content Analysis

The quality of the information from each component of the modeled study should be evaluated, as should the quality and strength of the *links* among components. We might think of these links between message variables and source or receiver variables in terms of how closely tied the data are. Although the strength of the ties between data sets decreases as we move from first-order to third-order linkage, all are improvements over simple description and unwarranted inference.

First-Order Linkage

The units of analysis are isomorphic (i.e., the same) for content analysis and source or receiver data. This one-to-one correspondence allows for strong relationships to be established. The one-to-one link may be a Type A, in which the precise messages analyzed in the content analysis are the ones created by the sources under study or are the ones accessed by the receivers under study. An example would be Naccarato and Neuendorf's (1998) study of print ads, in which the very ads that were content analyzed were the ads receivers responded to in a readership survey. Or the first-order link may be a Type B, in which the messages and sources or receivers are linked by a unit of analysis that

is not a message under investigation—for example, if the messages and receiver characteristics are summarized and then linked by a time period, such as *year,* as in studies of news coverage and public opinion (e.g., Domke et al., 1997; Iyengar & Simon, 1993; Watts, Domke, Shah, & Fan, 1999).

Second-Order Linkage

In this case, a link is established without a one-to-one correspondence on some unit of analysis. Such links may be anecdotal or occasional—that is, every unit in the content analysis is not matched with a unit in a source or receiver study. An example shown later in this chapter is Andsager and Miller's (1992) study of news coverage of a public issue, which they link to intermittently occurring events in the sociopolitical environment.

Third-Order Linkage

Here, there are no one-to-one or occasional correspondences of units of analysis. Rather, there is an overall *logical* link between content analysis and other studies based on the variables selected for study. Studies identified earlier as descriptive might easily fit this description. A third-order link is simply a logical link, using evidence from source or receiver studies to provide a rationale for a content analysis or using a content analysis as motivation for source or receiver studies. For example, a set of studies on alcohol advertising found that two of the most common appeals in beer and wine ads (celebrity endorsements and sex appeals, as identified in a *content analysis*) were also significantly more attractive to teens than to older adults (as discovered in *experimental studies* of teens and adults, using newly created ads; Atkin, Neuendorf, & McDermott, 1983).

The integrative approach is a simple way of analyzing the role of content analysis in the investigation of the larger framework of the communication process. Examples that have linked content analysis data with extramessage source data and extramessage receiver data are considered in turn in the following discussion.

Linking Message and Receiver Data

Often, a goal in marketing and mass media research is to demonstrate an effect (e.g., greater audience attendance) of media message characteristics. For example, the Naccarato and Neuendorf (1998) study (see Box 3.2) could be modeled in the following way, with S representing source characteristics, M/Ch representing message (within a channel) characteristics, and R representing receiver characteristics. The double-headed arrow may be viewed as leading from message to receiver (indicating effects) or from receiver to message (indicating use or voluntary exposure).[3]

Naccarato & Neuendorf (1998)

S → M/Ch ↔ R

① ↔

Content analysis of business-to-business magazine ads ↔ Readership survey conducted by publisher

Data are linked one-to-one (first-order linkage, ①); the unit of analysis is the individual ad, for which both content-analytic and survey data are collected. Note that the units of data collection are different, however (unit = ad for the content analysis, unit = person for the readership study).

In evaluating this content analysis, we might criticize the low reliabilities of a number of measures and the inclusion of a relatively small number of content (vs. form) variables. The readership survey may be criticized for sampling problems and self-report issues typical of readership studies. On the other hand, the linkage is quite sound, given the one-to-one correspondence by individual ad. As noted in Box 3.2, the researchers were able to conduct multiple regression analyses to predict readership scores from ad characteristics.

In the case of Hertog and Fan's (1995) study (see Box 3.4), the original units of sampling or data collection are not the same, but the two data sets have some shared unit of analysis. The study could be summarized as

Hertog & Fan (1995)

S → M/Ch → R

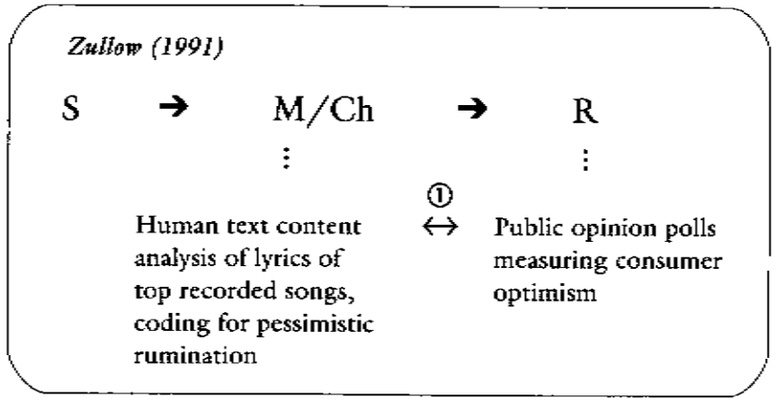
① ↔

Computer text content analysis of print news stories about HIV transmission ↔ Public opinion polls about HIV transmission

Data are linked by time period (first-order linkage); the units of analysis are time periods corresponding to 23 available National Health Interview Surveys between 1987 and 1991; the original units of data collection were the news stories for the content analysis and the individual respondents for the polls. Each study was well executed, and the tie between the two was fairly strong.

In an extraordinary study that ultimately linked pessimism in popular songs to the subsequent state of the U.S. economy (really), Zullow (1991)

also accessed publicly available data (public opinion poll findings and economic data) to link with his content analysis. This is also a wonderful example of the high level of sophistication to which content analyses may aspire. Both the coding scheme and the time-series analysis plan required a high degree of methodological and statistical expertise. Zullow found that as pessimistic rumination (i.e., negative descriptions or evaluations of an event) in popular songs increased, consumer optimism declined. Furthermore, he found that a decrease in gross national product (GNP) tended to follow. The flow from "bad vibes" songs to lowered GNP was found to occur over an average of 2 years. His study may be diagramed as

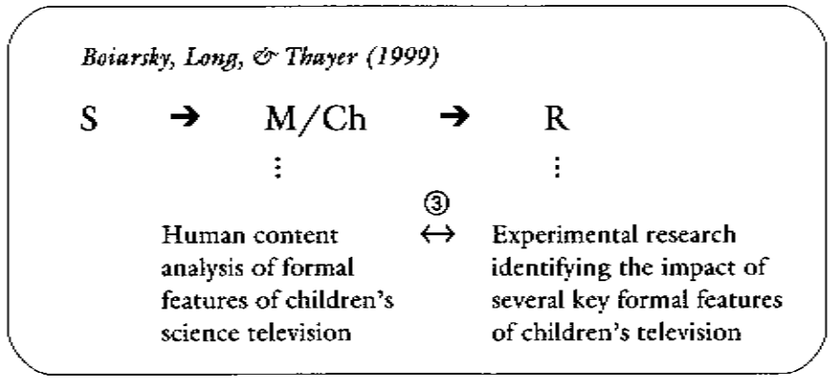


Data are linked by time period (year; first-order linkage). Each data collection was sound: The intercoder reliabilities for the coding of song lyrics were good, and the consumer polls were the highly regarded Index of Consumer Sentiment, conducted by the Survey Research Center at the University of Michigan.

Zullow (1991) has applied measures of pessimism and rumination in other contexts. In a human-coded content analysis of Democratic and Republican presidential candidate nomination acceptance speeches from 1948 to 1984, Zullow and Seligman (1990) found that the candidate who expressed more pessimistic rumination lost 9 of 10 times. Again, the study linked message characteristics and audience (receiver) responses in a clear and powerful manner.

Boiarsky, Long, and Thayer's (1999) study of children's science television provides a clear case of a ***third-order linkage,*** using content analysis to test the prevalence of key message characteristics that have previously been found to be important to the receiver's response. They chose three form characteristics that had been well studied in experimental work: content pacing, visual pacing, and use of sound effects. Past studies had found rapid visual or auditory change to increase children's attention to television programming (in some cases, resulting in enhanced learning) but on the other hand had found rapid topic switching to inhibit children's learning (p. 186). The Boiarsky team was

interested in finding out whether contemporary children's programming that ostensibly had an educational goal—science programming—used devices that would maximize children's learning. Their mixed findings indicated a high number of attention-gaining features (e.g., sound effects, quick cuts) but also rapid pacing (e.g., cuts rather than dissolves, very frequent topic shifts) that would tend to inhibit children's learning. The study's linkage with previous work may be diagramed as



Again, the contents-analytic research and experimental studies, each well conducted in their own right, are linked only loosely, by a logical third-order connection (③).

Linking Message and Source Data

In content analyses in the field of psychology, a link between source characteristics and message characteristics is often desired. For a half century, Louis Gottschalk and colleagues have been involved in developing methods of measurement of psychological dimensions (with biological roots) in children and adults, through the analysis of the content and form of their verbal behavior (Gottschalk, 1995, p. 3). The early studies, especially, provided strong links between source and message; they were designed to validate content analysis measures against more traditional evaluative procedures—self-report scales, physiological measures, and psychiatric rating scales (assessed by an expert, trained observer).

In one study, the researchers measured brain activity, cerebral glucose levels, as well as levels of hopefulness and hopelessness in verbal reports of dreams following REM sleep, non-REM sleep, or silent, waking mentation (Gottschalk, Fronczek, & Buchsbaum, 1993). They concluded that there are "different cerebral representations for hopefulness and hopelessness during each [of the three] state[s] of consciousness" (Gottschalk, 1995, p. 14). Their study could be modeled as

Gottschalk, Fronczek, & Buchsbaum (1993)

S → M/Ch → R

⋮ ⋮

① ↔

Physiological measures during one of three different states of consciousness ↔ Content analysis of verbal reports of dreams

Data are linked on the individual subject (person), a first-order linkage.

In a study of interpersonal verbal behavior, Marche and Peterson (1993) refuted a substantial body of evidence indicating that males engage in the majority of interruptions when conversing with women. Their study looked at 20-minute structured conversations of 90 dyads, with same-sex or opposite-sex composition. With good intercoder reliability (87%-95%), conversation interruptions were identified by human coders. Interruption behavior did not vary significantly by age, by gender of the interrupter, or by the gender composition of the dyad: Males did not interrupt more often than did females. The study could be modeled

Marche & Peterson (1993)

S → M/Ch → R

⋮ ⋮

① ↔

Identification of source characteristics (e.g., gender) and *controlled* context of speech event (e.g., assigned gender composition of dyad) ↔ Content analysis of interrupting behaviors

Data are linked by the individual subject (person).

In journalism, studies that link content analyses with source studies seem to be a bit rarer than those linking message and receiver data. Shoemaker and Reese's (1996) comprehensive volume on "mediating the message" is unique in its emphasis on the study of mass media sources and those source and structural factors affecting the production of media content. Notably, they rely on the Westley and MacLean (1957) model of the news-gathering process. This

model is similar to the Shannon-Weaver (Shannon & Weaver, 1998) source-message-channel-receiver model but with a notable addition—a "universe of possible messages" from which sources select (Shoemaker & Reese, 1996, p. 34). But studies that investigate how sources make this selection are infrequent. Shoemaker and Reese have summarized scores of content-analytic studies that clearly intend to infer back to source intent without source data.

The few news studies that do include source data include Farley's (1978) study of magazine publisher gender and tone of coverage of the Equal Rights Amendment, in which female publishers produced more favorable coverage. Another source-integrative study is Shoemaker's (1984) investigation of special interest and other political groups, in which she combined content analyses of *The New York Times'* coverage of 11 political groups with survey data tapping U.S. journalists' attitudes toward those groups—the more "deviant" the journalists judged the groups, the less legitimately the publication portrayed the groups (Shoemaker & Reese, 1996, p. 90). Some studies of news coverage have measured source *organizational* variables rather than characteristics of individual authors or gatekeepers. For example, in a study of the impact of corporate ownership on news reportage, Pasadeos and Renfro (1988) compared the content of the *New York Post* before and after its purchase by media mogul Rupert Murdoch. They found that Murdoch ownership signaled a greater amount of page space devoted to visuals and a more sensational coverage style.

A content analysis that included a perfunctory survey of entertainment sources is Smith's (1999). Her study of character portrayals in female-focused films of the 1930s, 1940s, and 1990s (introduced in Chapter 1) included unobtrusive measures of the gender of key "sources" of the films—writers, directors, producers, and editors—as identified in film encyclopedic sources. Her findings included the identification of a somewhat surprising impact of female involvement behind the scenes, such that greater female creative control was related to a higher level of stereotypically feminine portrayals of women characters. Smith's combination of unobtrusive measurement of a key source variable (gender) and content analysis could be modeled in the following manner.

Smith (1999)

S → M/Ch → R

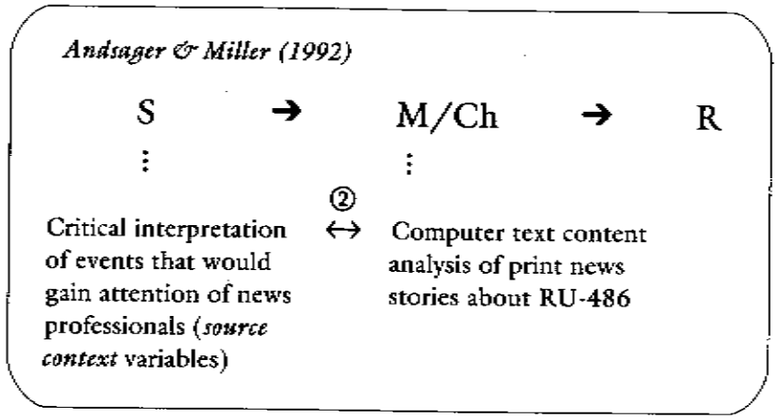
⋮ ⋮

① ↔

Unobtrusive tapping of genders of those in creative control of film ↔ Human content analysis of characteristics of female-focused films and characters within those films

Data are linked by film (a strong first-order linkage). The survey of sources was as valid as the encyclopedic sources used, and the content analysis was generally quite competent but suffered from some low reliabilities.

In a mass media study exemplifying a somewhat more tenuous source-message link, Andsager and Miller (1992) explored a connection between newspaper coverage of RU-486, the so-called abortion pill, and events occurring in the environment that were likely to affect news coverage. Each of 998 news stories appearing in major newspapers between 1982 and 1994 was computer content analyzed. Several important framings of the abortion pill were discovered in a cluster analysis of 125 key terms—information, women's health, and policy. Producing a time line for each of these three clusters, the researchers identified peaks in the types of coverage. They provided proposed explanations for these peaks based on time-matched events. For example, a peak in women's-health-type coverage of RU-486 in 1989 coincided with the announcement of a University of Southern California clinical trial of the pill and a new campaign by the National Organization for Women to legalize RU-486 in the United States. Another peak, in 1993, co-occurred with the Food and Drug Administration's approval of testing RU-486 in breast cancer prevention trials. As Andsager and Miller point out, "the incorporation of time-sequencing plots adds to understanding of relationships among a variety of concepts involved in an issue over time. They also aid in interpreting what events and issues shape coverage" (p. 9). The study may be diagramed as



Data are linked by time sequencing (but with only a second-order linkage [②]: that is, without a one-to-one correspondence of units of analysis).

Developing New Linkages

To date, content analysis studies that engage in some sort of first-order or second-order linkage between message and source or message and receiver are the exception rather than the rule. And studies that link all three—source, message or channel, and receiver—are rare. Some collected research reports, such as the 1972 *Surgeon General's Report on Television and Social Behavior* (Comstock & Rubinstein, 1972) have at least *addressed* all three. In the five-volume report on media violence, the editors combined several television content analysis and effects studies with a chapter on source (e.g., interviews with producers, writers, and directors; Baldwin & Lewis, 1972).

Researchers should be encouraged to add source or receiver data collection to their content analysis studies whenever possible. And although not forgetting theory as the primary motivator for any study, researchers might be alert to the potential for adding a content analysis to already existing findings regarding sources or receivers. For example, Solomon and Greenberg (1993) studied choices made by professional television property masters in their selection of furniture, clothing, and other props for TV commercials. Their survey of 25 working professionals found evidence of high consensus in choices of props for characters of a particular social class and gender. A content analysis could confirm how widespread is this "collective selection among individuals responsible for constructing the 'worlds' present in television commercials" (p. 17).

Sometimes, findings regarding the *effects* of a certain type of message may just be sitting there, waiting for a content analysis to add to the knowledge base. For example, Chen and Rada (1996) conducted a meta-analysis of experimental studies on the utility of hypertext, synthesizing the quantitative findings of 23 studies. They found a significant positive relationship between nonlinearity of structure and effective user performance and a significant positive relationship between the presence of graphical maps and user performance. A logical next step would be for a researcher to see how prevalent these important hypertext characteristics are in the information pools of CD-ROMs and Web sites, for instance. The characteristics of nonlinearity and graphical representation have been found to be effective for users; now we may ask, are they being put to use in hypertext applications? Content analysis may bring closure to this issue and to others that could benefit from content-analytic data.

Notes

1. A *control* variable helps assess whether an alternative explanation to a true relationship between X and Y may exist. For example, we may wish to test whether an individual's television viewing (X) leads to the individual's aggressive behavior (Y), which we could diagram as X ---> Y. There may be reason to believe that the level of aggressive behavior in the home (Z) may be related to both X and Y, most likely in one of the following ways: (a) X ---> Z --->Y or (b) X <--- Z ---> Y. That is, perhaps (a) television exposure leads to aggression in the home, which in turn leads to an individual behaving aggressively, or (b) a climate of aggression in the home leads to both increased TV viewing and an individual in the home behaving more aggressively. In either case, X does

not directly affect Y, and any relationship found between X and Y is what we call *spurious*. Also, in either case, Z constitutes an *alternative explanation* for a relationship between X and Y. We might include a measure of Z in our study as a control variable. If, after including Z as a statistical control, a relationship between X and Y still holds, then Z may be eliminated as an alternative or competing explanation.

2. More recently, Phillips's work has examined mortality rates as related to other, nonmedia factors, such as living in or visiting New York City (Christenfeld, Glynn, Phillips, & Shrira, 1999), the symbolic meaning of an individual's initials (Christenfeld, Phillips, & Glynn, 1999), and whether a person's birthday has recently occurred (Phillips, Van Voorhees, & Ruth, 1992).

3. As a point of comparison, the typical nonintegrative (i.e., wholly descriptive) content analysis might appear like this.

Dixon & Linz (2000)

S → M/Ch ↔ R

⋮

Content analysis of
television news coverage
of lawbreakers

In this model, no data have been linked from either the sources of the messages or the receivers of the messages.

CHAPTER 4

Message Units and Sampling

This chapter introduces the reader to the initial decisions necessary in content analytic research. Various types of units are considered, showing the range of choices in selecting the unit(s) for a given study. There is discussion of proper random sampling techniques, including the standard simple random sampling, systematic random sampling, cluster sampling, stratified sampling, and multistage sampling. Issues of access to sampling frames, message archive documentation (Resource 1), the use of the NEXIS database for text collection (Resource 2), and the management of the medium (e.g., the use of computer programs to unitize and mark-up video content) are explored.

Units

In content analysis, a *unit* is an identifiable message or message component, (a) which serves as the basis for identifying the population and drawing a sample, (b) on which variables are measured, or (c) which serves as the basis for reporting analyses. Units can be words, characters, themes, time periods, interactions, or any other result of "breaking up a 'communication' into bits" (Carney, 1971, p. 52).

As indicated in Chapter 1, these types of units are called units of sampling, units of data collection, and units of analysis. They are not always the same; for example, Lombard et al. (1996) have routinely used a random sampling of time periods, dates, and television channels to obtain a good representative sample of television programming. From this body of content, they analyze certain variables for each episode. For other, more microscopic variables, each randomly selected 10-second time interval within the episode ("timepoint") is used as the unit of data collection, and other units of data collection are also